CHILDREN LISTEN: PSYCHOLOGICAL AND LINGUISTIC ASPECTS OF LISTENING DIFFICULTIES DURING DEVELOPMENT

EDITED BY: Mary Rudner, Birgitta Sigrid Sahlen, Viveka Lyberg Åhlander and K. Jonas Brännström

frontiers Research Topics

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

CHILDREN LISTEN: PSYCHOLOGICAL AND LINGUISTIC ASPECTS OF LISTENING DIFFICULTIES DURING DEVELOPMENT

Topic Editors:
Mary Rudner, Linköping University, Sweden
Birgitta Sigrid Sahlen, Lund University, Sweden
Viveka Lyberg Åhlander, Åbo Akademi University, Finland
K. Jonas Brännström, Lund University, Sweden

Table of Contents

Editorial: Children Listen: Psychological and Linguistic Aspects of Listening Difficulties During Development

Birgitta Sahlén[1]*, K. Jonas Brännström[1], Viveka Lyberg Åhlander[1] and Mary Rudner[2]

[1] Logopedics, Phoniatrics and Audiology, Department of Clinical Science, Lund University, Lund, Sweden, [2] Department of Behavioural Sciences and Learning and Linneaus' Center HEAD, Linköping University, Östergötland, Sweden

Editorial on the Research Topic

Children Listen: Psychological and Linguistic Aspects of Listening Difficulties During Development

The goal for this Research Topic was to advance the scientific state of the art by collecting empirical and theoretical contributions relating to listening in children. Empirical articles that apply methods including behavioral, psychophysical, and neuroimaging approaches to the study of any aspect of listening in children were welcomed. The plethora of the articles included in the present topic illustrate the complexity and the broad areas of research necessary to understand listening in children. The many avenues of research in the field suggest the need for continuous development to a coherent theoretical model that can be used to test predictions about listening and listening effort in children. In the following, we briefly summarize the 24 contributions.

EFFECTS OF NOISE AND CHILDREN'S OWN PERCEPTIONS OF THE LEARNING ENVIRONMENT

Listening in context, i.e., the intentional act of focusing attention on a particular source of auditory information in a specific multimodal setting is crucial for linguistic, cognitive, and social development. At the same time, listening often takes place to the accompaniment of background noise, and even low levels of background noise have been found to reduce listening comprehension in children. The preschool learning environment is considered to be particularly noisy. Few studies have, however, reported on preschoolers' own perceptions of their learning environment. McAllister et al. in a comprehensive interview study, explore how preschool children in Finland and Sweden, describe the preschool environment in relation to noise, voice, and verbal communication. Results were similar across countries; preschool children are well aware of high noise levels, they blame other children for making noise and shouting and for impaired communication and effects on hearing. Interestingly, they seem less aware of effects of noise on their own voice. Astolfi et al. reported on the relationship between acoustical measurement of classrooms and first graders' perceived well-being and noise disturbance. Children are less happy with themselves and have less fun with increasing levels of noise; feelings and perceptions that may have a serious impact on motivation and learning in the classroom. Prodi et al. investigated the effects of different types of noise, age, and gender on 11–13 year old children's speech intelligibility and sentence comprehension. Classroom noise was found to have the worst effect on both tasks. A developmental effect was seen, which depended on the task and listening condition in both tasks. Girls were

more accurate and quicker to respond in most listening conditions. It is evident that dynamic models are needed to capture the complex interaction of task demands and individuals' capacity, perceived effort and motivation in the classroom. Listening in background noise uses cognitive resources and has an inevitable effect on listening effort. Theoretical frameworks of listening effort suggest that reverberation may have similar impact on listening effort and fatigue as noise. Interestingly, findings by Picou et al. suggest that increased moderate reverberation has no effect on listening effort or fatigue. This finding together with findings showing that the association between behavioral measurements of listening effort and participants' own ratings of perceived listening effort are weak emphasize the need for further testing of theoretical assumptions.

THE ROLE OF PERCEPTUAL, LINGUISTIC, AND COGNITIVE SKILLS FOR LISTENING AND ACADEMIC SUCCESS IN CHILDREN WITH HEARING LOSS (HL) AND AUDITORY PROCESSING DISORDER—INDICATIONS FOR INTERVENTION AND THERAPY

Children with poor perceptual, linguistic and cognitive skills and who are without the correct support are at an even greater risk of listening difficulties than those with typical development (TD). Bilingual children with weak school language (L2) are often considered particularly vulnerable to background noise. We should, however, be careful explaining vulnerability to noise by bilingualism. Andersson et al. highlight the need to look beyond bilingualism and to consider explanations to academic struggle. In their comprehensive study of Swedish school children bilingualism alone predicted 38% of the variance in language scores. With information added on parental education, school characteristics, and enrolment in the school's recreation center the unique contribution of bilingualism was reduced to 9%. In the classroom setting, a vulnerable group of children comprises those with auditory processing disorder (APD). These children appear to have normal hearing sensitivity but still have listening difficulties. There is, however, a high co-existence of APD with other disorders affecting language, reading and attention, and large variation in the presentation of the difficulties. It is therefore essential to identify subgroups to inform clinical intervention for the individual child. Sharma et al. identified four different clusters of children with suspected APD. Differences in working memory capacity, phonological processing, and non-verbal intelligence were the main skills that characterized these clusters. The need for assessing a large range of skills in these children is thus evident according to the authors. Further examples of groups of children that encounter specific challenges in noisy environments are children with hearing loss (HL), with cochlear implants (CI) and/or hearing aids, children with developmental language disorder (DLD). Children in these groups can have excellent speech recognition in quiet, but still experience unique challenges when listening to speech in

noisy environments. Von Koss-Torkildsen et al. investigated how speech-in-noise (SiN) perception relates to individual differences in cognitive and linguistic abilities in children with HL and typically developing (TD) children with the Hearing in Noise Test (HINT). For the full sample, language ability explained a significant amount of variance in HINT performance beyond speech perception in quiet and, that language ability was a significant predictor of HINT performance for children with CI, Hearing aids, and DLD, but not for children with TD. The authors, as most other authors in this topic, conclude that technologies that support audibility together with language-specific early interventions to help improve children's capabilities to handle noisy classroom environments are crucial for outcome. Several other contributions address children with hearing loss. For example, Socher et al. explored why children with HL often perform more poorly compared to their hearing peers, on tests of socio-pragmatic skills. In their study, significant differences between participants with HL and children with TD were found on a measure connected to theory of mind. Further, a measure of verbal fluency was correlated with three sub-measures of pragmatic language ability. Thus, children with a better developed semantic network may be able to use language in a more flexible way for communication, which is of great importance when the source signal is degraded as for children with HL. Lexical intervention may thus promote vocabulary growth and comprehension to support interaction and learning in children with HL. This is also emphasized by Wass et al. who found that receptive vocabulary was the most influential predictor of reading comprehension in 29 11–12-year-old Swedish children with profound HL using CI. Education should thus, focus both on broadening and deepening of the children's vocabularies and comprehension of spoken language. That optimal classroom acoustics help children perceive also the minute details of language and thus promote understanding is argued by

Kirkhorn Rødvik et al., who explored perception and production of speech in children with CI compared to TD children. They found that for the participants with CIs, consonants were mostly confused with consonants with the same voicing and manner and that voiced consonants were more difficult to perceive than unvoiced consonants. As is commonly reported, vowels were perceived more easily compared to consonants. Authors conclude that classroom acoustics with high reverberation times can easily hamper language comprehension due to masking effects.

Deroche et al. examined the production and perception of lexical tones (F0) in Mandarin speaking children with cochlear implants (CI). They found that children with CI relied more on durational cues than F0-contours to produce and perceive lexical tones than their peers with normal hearing. This indicates a link between production and perception also in children with CI who have poorer access to auditory feedback during production.

Further, predictions of language development are studied, for example by Ching et al., who investigated to what extent cognitive ability at 5 years of age predicted language development from 5

to 9 years of age in a population-based sample of children with HL who participated in the Australian Longitudinal Outcomes of Children with Hearing Impairment (LOCHI) study. Digit span score at 5 years was a significant predictor of receptive and expressive language at 9 years, even when non-verbal IQ and 5-year-old receptive vocabulary were accounted for. The authors argue that these findings shed light on the unique role of early verbal working memory in predicting the development of language and vocabulary skills in children who use hearing aids. Further, Jing et al. investigated the association of rhyme awareness, a common index of phonological awareness, with vocabulary and working memory in a small group of North American children ($n = 6$) with CI. While associations were statistically significant in a larger group of children with TD ($n = 15$), only the association between rhyme awareness and working memory was significant in the children with CI. As for the production of emotional tone, Chatterjee et al. conclude that access to acoustic hearing in early childhood is important and speech prosody should be included in speech therapy. The authors compared acoustic characteristics of happy and sad vocal emotions produced by North American prelingually deaf school-aged children with CI during sentence reading with those produced by peers with TD and adults with normal hearing and postlingually deaf adults with CI. They found that all four groups differ in voice pitch between the two emotions produced, but that the difference was smallest for the children with CI.

Etiology of the hearing loss may also play a role for the development of language and cognition in children with HL. Löfkvist et al. studied the role of congenital cytomegalovirus (cCMV) infection on executive function. cCMIV is the most common cause of progressive HL and associated with behavioral anomalies. Authors did not find any significant difference in executive function between two small groups of Swedish child CI users, one with cCMV and one with genetic HL. However, they did find that pragmatic skills were reduced in the cCMV group and suggest that this may hamper academic success. Word reading and spelling in children with HL are also addressed since listening skills are not only essential for spoken language development but also for the development of reading and writing. Phonological processing skills have been considered predictive of good word decoding. In the paper by Gokula et al. the general co-existence of perceptual, cognitive, and linguistic deficits in children with word reading difficulties is highlighted. A comprehensive test battery designed to assess their auditory processing, visual attention, digit memory, phonological processing, and receptive language is used. Six percent of children with word reading difficulties have deficits across all measured tasks. The results thus emphasize the significant individual variability inherent in children with word reading difficulties and the importance of thorough and comprehensive assessments of reading skills. As for writing skills, the findings by Gärdenfors et al. conclude that spelling strategies in children with HL mostly rely on auditory input but the children with CI apply visual strategies when necessary.

MATURATION OF SPEECH PERCEPTION, PSYCHOPHYSICAL, AND NEUROIMAGING APPROACHES TO LISTENING

A range of interesting studies in this topic address maturation of masked and unmasked speech perception (i.e., Leibold and Buss, McCreery et al., Walker et al., and MacCutcheon et al.) and its relation to linguistic and cognitive development in children with TD and HL.

In their review article, Leibold and Buss summarize evidence showing that the ability to recognize masked speech develops over an extended period during maturation. Generally speaking, children have greater difficulty than adults. In steady-state noise, this difference persists until the age of about 9–10 years but when the masker is speech the difference extends into adolescence. The authors identify key challenges for future research. These include, teasing apart the factors that contribute to maturation of masked speech perception including, not least, the effect of hearing status.

Walker et al. compared developmental growth rates in speech recognition for North American children with and without HL. Children with HL showed persistent deficits in masked speech recognition until the age of 11 years but their development was parallel to that of children without hearing loss. Factors that influenced growth trajectories for masked speech recognition included stronger vocabulary skills and higher hearing aid dosage. Importantly, the authors point out the need to continue to support children with hearing loss in the academic setting as they transition to secondary education. McCreery et al. investigated the effect of hearing status on masked speech perception in North American children. They found that children with HL had poorer aided speech recognition in noise and reverberation than children with typical hearing. Children with better vocabulary and working memory had better speech recognition in noise and noise plus reverberation than peers with poorer skills in these domains. In general, the better the aided audibility the better the speech recognition in noise and reverberation. McCutcheon et al. investigated the effect of musical education on masked speech perception in South African children with TD. Authors were unable to identify any effect on either speech perception or phonological short-term memory of musical education.

Some contributions in this topic use psychophysical and neuroimaging approaches to the study of possible neural correlates to challenges of listening in children with TD and HL. Moore et al. investigated dichotic listening and neural correlates of a receptive speech task in typically developing children and children with listening difficulties. There were only subtle differences between groups but while activation in some brain regions correlated with dichotic listening for the group of typically developing children this was not the case for the children with listening difficulties. In their study on children with congenital HL, Jiang et al. suggest that the changes seen in white matter microstructures could depend on poor auditory input or cortical reorganization. Cardon and Sharma examined cross-modal reorganization of the auditory cortex in children with CI using vibrotactile stimulation. They found that children

with poorer speech perception in noise showed greater cross-modal reorganization, i.e., that their auditory cortices were more sensitive to vibrotactile stimulation than those with better speech perception in noise. Furthermore, greater cross-modal reorganization was seen in the cortex on the same side as their first CI indicating that this reorganization becomes more accentuated when auditory input is degraded.

To sum up, the 24 articles in this topic provide an important starting point for embracing very diverse aspects of listening difficulties in children. Key themes for further exploration are the effect of even low levels of background noise on perceived and actual listening comprehension in children, including but not limited to, children with special needs. This work should take into account developmental aspects. Also more intervention studies are called for. For example, can students' listening and language development be supported by teacher training aiming at fostering language learning i.e., vocabulary skills in the classroom? Further work is also needed on charting the neural correlates of listening difficulties in children. Last but not least, we believe that the complex relationship between the child's motivation, both intrinsic and extrinsic, and listening effort, measured both subjectively and objectively should be a key focus of future work as the development of more dynamic theoretical models of the interaction of these factors. Finally, we want to express our gratitude for all interesting contributions to this topic! They not only show the important advances we have

seen in this cross-disciplinary field during the last years, but they definitely also offer a great platform for future studies.

AUTHOR CONTRIBUTIONS

MR initiated the topic. All four editors communicated with topic authors and reviewers during the revision process. BS was responsible for the editorial.

ACKNOWLEDGMENTS

Processing of Acoustic Information in Lexical Tone Production and Perception by Pediatric Cochlear Implant Recipients

Mickael L. D. Deroche[1]*, Hui-Ping Lu[2], Yung-Song Lin[2,3], Monita Chatterjee[4] and Shu-Chen Peng[5]

[1] Department of Psychology, Concordia University, Montreal, QC, Canada, [2] Chi-Mei Medical Center, Tainan, Taiwan, [3] Taipei Medical University, Taipei, Taiwan, [4] Boys Town National Research Hospital, Omaha, NE, United States, [5] United States Food and Drug Administration, Silver Spring, MD, United States

Purpose: This study examined the utilization of multiple types of acoustic information in lexical tone production and perception by pediatric cochlear implant (CI) recipients who are native speakers of Mandarin Chinese.

Methods: Lexical tones were recorded from CI recipients and their peers with normal hearing (NH). Each participant was asked to produce a disyllabic word, *yan jing*, with which the first syllable was pronounced as Tone 3 (a low dipping tone) while the second syllable was pronounced as Tone 1 (a high level tone, meaning "eyes") or as Tone 4 (a high falling tone, meaning "eyeglasses"). In addition, a parametric manipulation in fundamental frequency (F0) and duration of Tones 1 and 4 used in a lexical tone recognition task in Peng et al. (2017) was adopted to evaluate the perceptual reliance on each dimension.

Results: Mixed-effect analyses of duration, intensity, and F0 cues revealed that NH children focused exclusively on marking distinct F0 contours, while CI participants shortened Tone 4 or prolonged Tone 1 to enhance their contrast. In line with these production strategies, NH children relied primarily on F0 cues to identify the two tones, whereas CI children showed greater reliance on duration cues. Moreover, CI participants who placed greater perceptual weight on duration cues also tended to exhibit smaller changes in their F0 production.

Conclusion: Pediatric CI recipients appear to contrast the secondary acoustic dimension (duration) in addition to F0 contours for both lexical tone production and perception. These findings suggest that perception and production strategies of lexical tones are well coupled in this pediatric CI population.

Keywords: lexical tone, cochlear implant, cue trading, speech production, children

INTRODUCTION

Cochlear implants (CIs) are medical devices that are surgically inserted in the cochlea of patients with severe-to-profound sensorineural hearing loss to provide auditory sensation by electrically stimulating the auditory nerve. Even though CI devices help to improve speech perception by patients, the device technology has its limitations. One constraint is that CI devices are equipped with speech-coding strategies that are temporal-envelope based (Shannon, 1983; Zeng, 2002), and their audio signals are delivered with poor spectral resolution. With this limitation, speech and other sound information involving complex harmonic pitch (fundamental frequency or F0) –critical for functions such as the perception of prosody (i.e., speech intonation and lexical tones), talker-sex, melody identification, and speech perception in noise – is poorly processed by CI patients (Qin and Oxenham, 2003; Wu and Yang, 2003; Fu et al., 2004; Kong et al., 2005; Galvin et al., 2007; Gfeller et al., 2007; Chatterjee and Peng, 2008; Cullington and Zeng, 2008; Luo et al., 2008; Peng et al., 2008, 2009; Zhu et al., 2011; Wang et al., 2013; Zhou et al., 2013; Chatterjee et al., 2015; Liu et al., 2019). For native speakers of a lexical tone language such as Mandarin or Cantonese, the aforementioned limitation hinders CI users' ability to identify contrasts in lexical tones, since F0 serves as the primary information for this task (e.g., Whalen and Xu, 1992; Liu and Samuel, 2004). This limitation is particularly challenging for pediatric CI recipients who are prelingually deaf (i.e., born deaf or become deaf before ages five or six), given that these individuals have to rely on a CI to master the lexical tone system critical for their spoken language development. The restricted access to F0 information may also affect how pediatric CI listeners utilize F0 cues along with secondary acoustic dimensions, such as duration, to identify as well as produce lexical tones.

Lexical Tone Perception

Lexical tone *perception* has been widely studied in both adult and pediatric patients with CIs. Wang et al. (2013) examined lexical tone recognition using mono-syllabic words in CI patients who were post-lingually deaf. They found that performance was much poorer than adult listeners with normal hearing (NH), and also poorer than adult individuals with severe hearing loss. They observed a negative correlation between performance and audiometric thresholds of adult listeners who are hearing impaired, particularly at 250 Hz, highlighting the critical importance of low frequencies for this task. Wang et al. (2011, 2012) found performance in this task to be correlated with complex pitch discrimination and musical instrument identification. However, adult listeners with a long CI device experience are known to alter listening strategies to perform auditory tasks. As their device does not allow a fine representation of F0 contours, post-lingually deaf adult CI users have been shown to develop alternative strategies for lexical tone recognition based on secondary (or possibly tertiary) cues. This phenomenon is referred to as *cue-trading* and has been shown in many speech perception tasks when the primary cue for the task is degraded. For instance, Peng et al. (2009, 2012) examined English-speaking CI users to distinguish questions from statements based on their contrasts in speech intonation. As the primary cue for speech intonation (F0 contour) was poorly transmitted by their devices, CI users showed greater reliance on secondary cues (intensity and duration patterns) to perform this task. Cue-trading is also observable in listeners with NH or with CI in the laboratory setting, by manipulating the type and quality of acoustic information in phonetic identification (e.g., Winn et al., 2012, 2013).

While cue-trading has been demonstrated in adult listeners, the phenomenon has been relatively under-studied in children. The literature suggests that children and adults use different sets of perceptual weights for speech recognition (Nittrouer, 1996; Hazan and Barrett, 2000). Among pediatric CI recipients who are prelingually deaf, performance in lexical tone recognition has been reported as highly variable (e.g., Ciocca et al., 2002; Peng et al., 2004; Zhou et al., 2013; Chen et al., 2014). There is, however, a consistent trend: those who perform better in lexical tone recognition tend to have longer experience with their device. This trend suggests that while cue-trading in electric hearing takes time to learn, children eventually adapt and develop novel strategies in their language. On the other hand, Chen et al. (2014) reported that while maternal education level (an indicator of socio-economic status) plays a positive role for speech recognition in children with CIs, it does not predict their lexical tone identification performance. This outcome points toward limitations inherent to the device that are not easily overcome by the development of alternative strategies or environmental factors.

Lexical Tone Production

Lexical tone *production* by pediatric CI recipients has also been investigated in several studies. Similar to findings with lexical tone recognition, considerable individual variability was observed within each study. In addition, findings among studies exhibited discrepancies, potentially related to the different protocols and methodologies adopted among those studies. Broadly speaking, two approaches have been followed. Some studies (Peng et al., 2004; Xu et al., 2004; Han et al., 2007) asked experienced listeners (typically speech pathologists or NH adult listeners who are familiar with the speech of hard of hearing) to rate how they would perceive the accuracy of the lexical tones produced by the children. Accuracy was reported as between 30 and 70% correct for the majority of children with CI, being considerably lower than the accuracy of their NH peers. Tones 1 and 4 were generally better produced than Tones 2 and 3, a pattern consistent with the developmental trend among children with NH in their acquired mastery with lexical tones in Mandarin Chinese (Li and Thompson, 1977). However, those studies warned that it is sometimes difficult for judges to make reliable assessment about the quality of lexical tone productions that are irregular over time (across repetitions). To circumvent this issue, another approach was followed in which the recordings were either analyzed acoustically and some indices were derived to reflect the production quality (e.g., Barry and Blamey, 2004) or automatically categorized by a neural network

based on F0 contours (Zhou and Xu, 2008; Xu et al., 2011; Zhou et al., 2013). This second approach thus permitted a relatively objective assessment of production quality (i.e., free from human biases). A large overlap between tonal ellipses, i.e., a lack of tonal differentiation, was reported for CI children (Barry and Blamey, 2004). Further, the outcomes of the neural network were largely consistent with NH listeners' ratings, i.e., they confirmed substantial deficits in lexical tone production by prelingually deafened children with CI that worsened as age at implantation advanced (Zhou et al., 2013).

A Link Between Perception and Production?

Outside of the lexical tone literature, it has been known for a while that perception and production are tightly linked (Bradlow et al., 1997; Houde and Jordan, 1998), including for F0 control (Elman, 1981; Larson et al., 2000). Naturally, this has led the aforementioned studies to focus largely (for human judges) or solely (for the acoustic analyses of Barry and Blamey (2004), or the neural network adapted from Zhou and Xu (2008) on the quality of the F0 contours produced. The rationale was that children with CIs would not be able to produce tones correctly unless they were able beforehand to perceive F0 sufficiently well to learn to recognize the particular F0 inflections of a given tone and eventually fine-tune their speech motor commands. To some degree, this rationale is supported by correlations between perception and production performance (Peng et al., 2004; Xu et al., 2011; Zhou et al., 2013). However, this rationale suffers from a serious limitation: considering the cue-trading phenomenon established in perception, information other than F0 must be examined. One might easily imagine that CI recipients deemphasize F0 contours and emphasize differences in intensity or duration while producing lexical tones, but such cue-trading phenomenon in production remains to be documented. This is important because regardless of the fact that all CI recipients suffer from some loss in functional spectral resolution, a fraction of these children exhibit little deficit in lexical tone production (Han et al., 2007; Zhou and Xu, 2008). Without explicit knowledge of the type of acoustic information being used for perception and those emphasized in production, we might not appreciate the roots of individual variability. If the reliance on specific acoustic dimensions in tone identification were reflected in production *by the same individuals*, this would suggest mechanistic links between the development of perception and production of lexical tones that are driven by the characteristics of the acoustic input.

Goals of the Study

The purpose of this study was (1) to examine the aspects of vocal production that Mandarin Chinese speaking pediatric CI recipients emphasize or deemphasize to convey lexical tones, (2) to compare pediatric CI recipients' lexical tone production to that of NH peers, and (3) to compare lexical tone production and perception in the same participants. The perception task followed the design of a preliminary study (Peng et al., 2017) that focused on a single, unambiguous, contrast: Tone 1 vs. Tone 4. In running

speech, Tones 2 and 3 can sometimes bear some similarity and are known to be mastered relatively later throughout development (Li and Thompson, 1977). For children as young as 6 years of age, we wished to avoid any abnormal production that would solely be the result of those tones still being learnt about. Thus, the study also focused on the production of Tones 1 and 4 exclusively, which were more likely to reflect intentional and consolidated speech motor commands.

We hypothesized that in lexical tone production, NH children would contrast the two lexical tones based on F0 contours exclusively. In contrast, pediatric CI recipients would express differences in duration or intensity patterns, while not modulating their vocal chords' vibrations well enough to contrast the high-level F0 contour of Tone 1 vs. the falling contour of Tone 4. Finally, we suspected that the perceptual weights derived for an isolated word, controlled in laboratory settings, may not necessarily generalize to perceptual strategies recruited in running speech (ecological situations), and therefore, they may not have had time to influence the speech motor commands. Thus, we expected to observe a perception-production coupling more easily in children who had more extensive experience with their device.

MATERIALS AND METHODS

Participants

Participants in this study were comprised of 40 pediatric CI recipients who all became deaf prelingually (deafness before 2.3 years of age), ranging from 6.4 to 17.2 years of age. For the most part (37 out of 40), they were implanted early between 1.1 and 4.5 years of age; only three were implanted at 5.6, 7.3, and 12.2 years of age. Consequently, the median of age at implantation was 2.5 years. These children had used their CI devices from 1.2 to 15.2 years. Note that there was no correlation between chronological age and age at implantation [$r^2 = 0.03$, $p = 0.295$], or chronological and age at profound hearing loss [$r^2 = 0.02$, $p = 0.427$], but a significant correlation between chronological age and years of CI use [$r^2 = 0.72$, $p < 0.001$]. In addition, 35 NH children (bilateral thresholds <20 dB HL at octave frequencies between 250 and 8000 Hz) were recruited. There was no significant difference in age at testing between the CI and NH groups [$t(73) = -0.4$, $p = 0.700$]. The demographics of these two participant groups are reported in **Table 1**. All participants and their parents provided written informed consent, which was reviewed and approved by the Institutional Review Board at the Chi-Mei Medical Center.

Among the CI participants, seven were implanted bilaterally, 15 were implanted unilaterally (eight on the left side, seven on the right), wearing no hearing aid on their contralateral ear. The remaining 18 were bimodal users, who wore a hearing aid on the contralateral ear (13 implanted on the left side, and 5 on the right). However, CI participants were tested (in the perception task) and recorded (in the production task) with a single CI, being the device implanted first. This CI was turned on, while using the clinically assigned settings, and any other implant or hearing aid on the contralateral ear was removed. Thirty-five participants

	Chronological age mean (std.) [min–max]	Age at implantation mean (std.) [min–max]	Duration of CI use mean (std.) [min–max]
NH children (n = 35)	10.6 (2.8) [6.8–16.5]		
CI children (n = 40)	10.9 (3.4) [6.4–17.2]	2.9 (1.9) [1.1–12.2]	7.9 (3.6) [1.2–15.2]

Numbers are given in years. All CI children had lost their hearing before age 2.3. Note that the distribution of age at implantation was skewed: only three children were implanted at 5.6, 7.3, and 12.2 years of age while the remaining 37 were implanted before 4.5 years of age.

were users with a Cochlear Nucleus device (N24, CI422, CI512, Nucleus Freedom, all using the ACE coding strategy). Four were users of a Med-El device (Pulsar, Concerto, Sonata, using the Opus2 coding strategy). One remaining participant was a user of the Advanced Bionics' HiRes90k device, using the Fidelity 120 coding strategy. All participants with hearing aids had audiometric thresholds exceeding 90 dB HL at the time of testing, but some of them may have been exposed to acoustic hearing pre-implantation. For example, one of the participants, implanted at age 12, had high thresholds (∼80 dB HL) until he suffered sudden profound hearing loss and subsequently received a CI. Although perception and production measures were all made with only the CI in place, thus excluding any effects of acoustic hearing at the time of testing, the presence of hearing during development may be expected to influence perceptual cue-weighting as well as production patterns. Therefore, we included an analysis based on the presence or absence of residual hearing in our participants.

Production Task

All participants were asked to produce the Chinese disyllabic word "yan-jing," with the 2nd syllable pronounced with Tone 1 (a high level tone) and Tone 4 (a high falling tone) to represent "eyes" (眼睛) and "eyeglasses" (眼鏡), respectively. The first syllable is always pronounced with Tone 3 (a dipping tone). Participants were asked to produce the target words in a natural way, just as how they would say it in everyday life. Three repetitions of each target word were recorded from each participant, in order to increase the number of observations and determine to what extent the extracted acoustical parameters varied from one recording to another. The recordings were performed at two clinical sites, the Chi-Mei Medical Center in Tainan and the Chang Kung Memorial Hospital in Taoyuan. The experimental sessions were set up at both sites in the following, comparable fashion: Signals were recorded at a 44.1 kHz sampling rate with a 16-bit resolution, with a minidisc recorder (Sony MZ-RH1) through a stereo microphone (Audio-Technica AT9440) placed approximately 10 cm from the speaker, in a sound-treated booth. The recordings were transferred from the mini disc to a laptop through the Sonic Stage (Digital Music Manager Version 3.4) software program and saved as .wav files for further editing. With the Adobe Audition 3.0 software program, each of the signals was cut with 400-ms of silence before the onset and after the offset.

Perception Task

The perception task followed the same methods identical to that in Peng et al. (2017). A continuum between Tones 1 and 4 was created in the lab by orthogonally manipulating

(1) the slope of the F0 contour, and (2) the duration, of the second syllable of the word "yan-jing." The range of slopes varied from −1, −0.8, −0.6, −0.4, −0.3, −0.2, −0.1, to 0 octave. The range of durations varied from 40, 60, 80, 100, 120, and 140% of the initial duration. These manipulations were performed at a F0 height of 120 Hz (typical of male voices) or 220 Hz (typical of female voices), resulting in a total of 96 tokens per testing session. Each participant completed three or four sessions, in which all tokens were presented one after another in random order. This task followed a single-interval, two-alternative forced-choice paradigm (2AFC) in which the participant was asked to indicate whether a given stimulus meant "eyes" or "eyeglasses" by clicking on one of two response buttons shown on the computer screen. Sounds were delivered from an external soundcard (Soundmax Integrated HD Audio) connected to a laptop through loudspeakers (Audio Pro) at approximately 65 dB SPL at the child's ears. Although the amplitude contour (which is naturally correlated with the F0 contour, at least within the NH population) conveys some information about lexical tones (Whalen and Xu, 1992), the overall intensity does not. All stimuli were therefore, presented at the same root-mean-square (RMS) level. The amplitude contour was not manipulated in this study, as the study focused on the trade-off between F0 and duration cues.

PRODUCTION DATA ANALYSIS

Extracting Acoustic Parameters

Acoustic analyses were performed on each of the two syllables for all recorded tokens. We extracted the intensity pattern sampled every 5 ms with Praat (Boersma and Weenink, 2018), and ran a peak detection algorithm with a peak prominence of 20 dB. In 8 cases, this algorithm did not permit us to successfully locate the two peaks because the intensity pattern dropped by less than 20 dB between the two syllables. When this occurred, the peak detection algorithm was reiterated with a lower peak prominence until it successfully located two peaks. Each syllable was then trimmed on either side of the intensity peak. The choice of 20-dB cutoff permitted selection of the entire syllable, including the last phoneme /n/ in the first syllable or /η/ in the second syllable. The F0 values were also sampled every 5 ms. All recordings were first concatenated together and F0 points were extracted within a default range 75 to 600 Hz. This resulted in a dominant distribution with a few outliers that were octave jumps. To prevent those errors, the F0 distribution was then fitted with a *normal* probability density function on a logarithmic axis, essentially to reflect the vocal range of a

given child. The mean of the fit was chosen as the center of the vocal range which was subsequently restricted to $+/-6$ semitones around. Each token was then analyzed using Praat with this narrow F0 range. Visual inspections were performed to identify any abnormality. Abnormalities occurred in four occasions for the NH population and 11 occasions for the CI population over all tokens, either because (1) the production was not sufficiently voiced, or because (2) the F0 contour exceeded the $+/-6$ semitones range (e.g., higher range for Tones 1 than 4). In cases (1), the voicing threshold was adjusted manually to 0.1 rather than the default 0.45 (parameter in Praat) as a way to reduce the influence of unvoiced portions (while keeping a 20-dB cutoff window). In cases (2), the F0 range was expanded up to $+12$ semitones and down to -9 semitones relative to the center of the vocal range. The entire F0 contour was recorded, but for the purpose of this study, the analyses focused on

two descriptors: F0 median and F0 movement from the first to the last 30 ms.

An additional analysis was performed with a 10-dB cutoff, revealing qualitatively similar findings as with the 20-dB cutoff (see **Appendix**). Its rationale was that the final phoneme (voiced consonant) contributed in some cases to the F0 contours of Tones 1 and 4 (e.g., right panels **Figure 1**). Since the middle vowel was more intense than the phonemes embedding it, this procedure allowed a closer focus on the voiced part of each syllable, which provided more canonical F0 contours even though it was too conservative.

As an example, **Figure 1** shows the parameters extracted from the recordings of Tones 1 and 4 produced by a male NH participant (top two rows) and by a female CI participant (bottom two rows). In each panel, the black vertical lines delimit the window selected from the 20-dB cutoff relative to the intensity
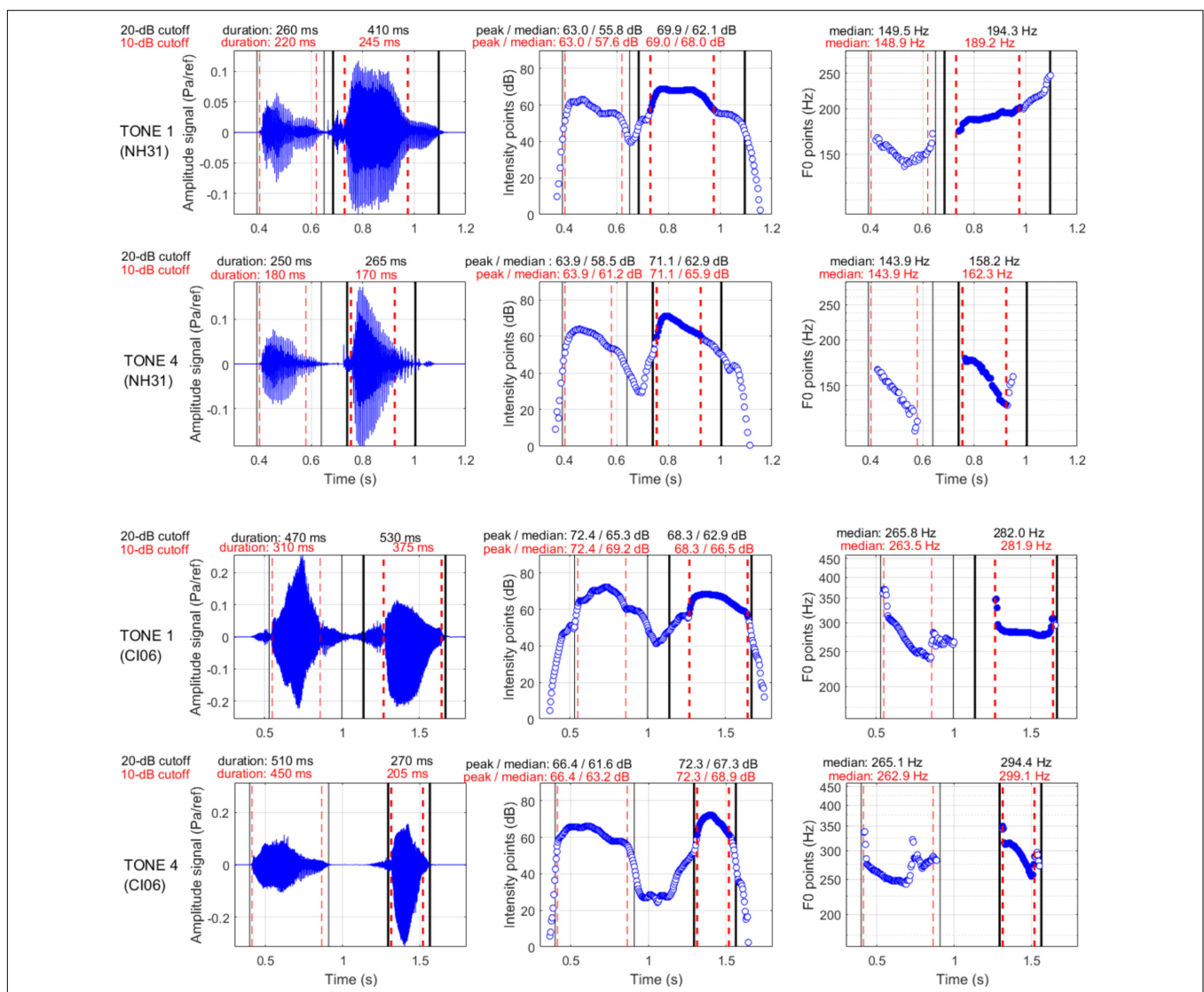
FIGURE 1 | The "yan-jing" production, with the second syllable produced as Tone 1 (top) or Tone 4 (bottom) by a 12.7-year-old participant with NH (top two rows) and by a 10.9 year-old participant with a CI (bottom two rows).

peak, and the red dashed lines delimit the window selected from the 10-dB cutoff. Several traits can be observed. For the NH boy, there was little difference in duration between the two syllables; for both tones, the intensity was greater for the second than for the first syllable. For the CI girl, the syllable produced as Tone 4 was markedly short (possibly due to extended duration of the first syllable); the syllable produced as Tone 1 was markedly soft (possibly due to greater intensity of the first syllable). As shown in the right panels, the F0 pattern of the first syllable was either V-shape or falling. This pattern was quite common, and occurred whether the following syllable was Tone 1 or Tone 4. As anticipated, the boy produced the second syllable either with a higher-level/slowly rising pattern for Tone 1 or a rapidly falling pattern for Tone 4. The girl produced a falling F0 for Tone 4 but produced a largely monotonous pattern for Tone 1 that was similar to the F0 range of the first syllable.

Statistical Analyses

The statistical analysis was performed for one acoustical parameter at a time. We used a linear mixed effects (LME) approach, where the initial model had two fixed effects: hearing status (NH vs. CI) and lexical tone contrast (Tone 1 vs. Tone 4), including an interaction term. We also considered random intercepts for each participant as well as a random slope for the effect of contrast because both of these additions significantly improved the initial model. Any other addition (random intercepts for "repetitions," random slopes for the effect of hearing status, random slopes for the effect of sex, random slopes for the effect of chronological age, or even sex as a third fixed effect) did not improve the model and were therefore, excluded. Thus, the final model was of the form "parameter $\sim$ 1 + Contrast*Hearing + (1+Contrast | Participant).

PERCEPTION DATA ANALYSIS

Data from all testing sessions were pooled together and the proportion of Tone 1 responses served as the dependent variable in a logistic mixed-effect analysis. There were three fixed factors: (1) population, (2) slope of F0 variation, and (3) duration, including interaction terms. Note that the duration scale was log-transformed for centering purposes. We also included a random intercept per subject, and random slopes for the effect of F0-slope, duration, as well as F0-height. Thus, the final model was of the form "responses $\sim$ Population*F0variation*Duration + (1+F0variation+Duration+F0height | Participant)." This enabled extraction of coefficients for each subject that reflected the reliance on pitch or duration cues, which could then be correlated against the production outcomes.

PRODUCTION RESULTS

Duration

As displayed in the top panels of **Figure 2**, children with NH prolonged the duration of the second syllable by about 10–20% relative to the first syllable. In contrast, participants with CIs did so for Tone 1 (by about 30%) but not for Tone 4. In other words, participants with CIs tended to contrast the duration patterns to distinguish between Tones 1 and 4. The LME was further performed on the duration ratio between the two syllables (top-right panel). This ratio permitted to discard variances associated
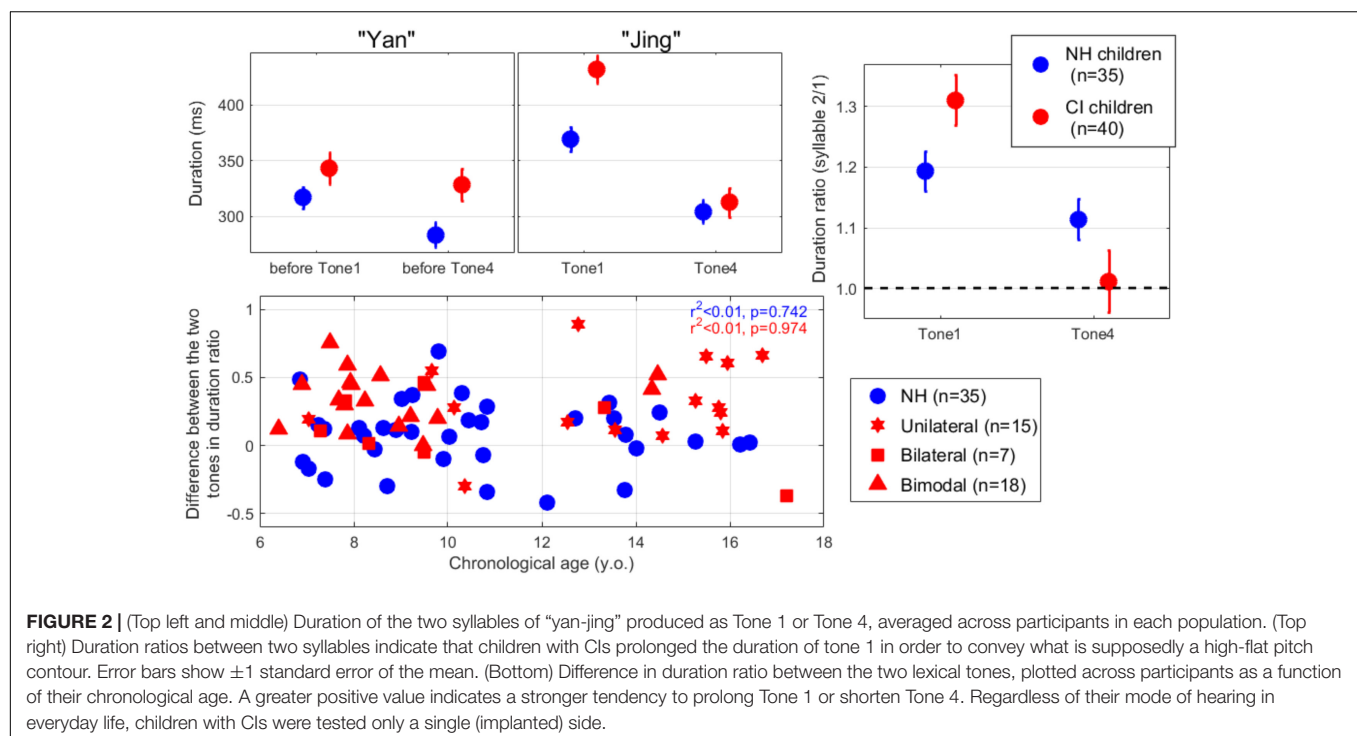
FIGURE 2 | (Top left and middle) Duration of the two syllables of "yan-jing" produced as Tone 1 or Tone 4, averaged across participants in each population. (Top right) Duration ratios between two syllables indicate that children with CIs prolonged the duration of tone 1 in order to convey what is supposedly a high-flat pitch contour. Error bars show ±1 standard error of the mean. (Bottom) Difference in duration ratio between the two lexical tones, plotted across participants as a function of their chronological age. A greater positive value indicates a stronger tendency to prolong Tone 1 or shorten Tone 4. Regardless of their mode of hearing in everyday life, children with CIs were tested only a single (implanted) side.

with individual speaking characteristics, i.e., different speaking rates among participants. There was an effect of hearing status [$t(446) = 2.2$, $p = 0.029$], no effect of contrast [$t(446) = -1.9$, $p = 0.062$], and an interaction between the two [$t(446) = -3.8$, $p < 0.001$]. This interaction was the key evidence that participants with CIs utilized duration to contrast Tones 1 and 4, whereas participants with NH did not.

A question of interest was whether, there was a particular profile of pediatric CI recipients who displayed this "alternative behavior," i.e., shortening Tone 4 or prolonging Tone 1 as a substitute for their respective F0 contours. The bottom panel shows the difference between the duration ratios of Tones 1 and 4. Here, a positive value indicates that the participant produced a longer duration for Tone 1 than for Tone 4 (still with the 2nd syllable duration being relative to that of the first syllable). This alternative behavior was shared by most of the children with CIs (with two exceptions), and was not found to be related to chronological age ($p = 0.974$). There was also no evidence that this alternative behavior was driven by age at implantation or duration of CI experience (respectively, $p = 0.136$ and $p = 0.450$, not shown).

Intensity

As the intensity and F0 contours are correlated (Whalen and Xu, 1992), and because the intensity contour might be more salient for children with CIs, it might be that pediatric CI users adjust intensity during production to emphasize or deemphasize specific parts of tones. Accordingly, we examined the dynamics of the produced intensity contours. The two left panels of **Figure 3** (referring to the non-contrastive syllable) bore a striking similarity with (1) a peak arising about one third of the total duration of the first syllable, and (2) a peak of similar magnitude whether this syllable preceded Tone 1 or Tone 4. For the second syllable, the intensity pattern of Tone 4 closely resembled an inverse V-shape, whereas a high-level intensity was maintained over a longer portion of Tone 1, dropping much closer to the edge of the time window (like an inverse U-shape). It was also apparent that, on average, NH children strengthened the intensity of the second syllable relative to the first, for both tones. In contrast, CI children did so for Tone 4 only.

Visual inspection of intensity peak of each syllable (top-left panels of **Figure 4**) indicates that NH children produced the second syllable at a higher intensity than the first, in both target words (i.e., *eyes* and *eyeglasses*). Children with CIs, on the other hand, did not when producing the target word with Tone 1 (i.e., *eyes*). The LME analysis was performed on the intensity peak of the second syllable, relative to the peak of the first syllable (top-right panel of **Figure 4**). There was no effect of hearing status [$t(446) = -1.9$, $p = 0.061$], an

FIGURE 3 | Mean (thick line) and one standard error (areas) for the intensity patterns extracted for each syllable when producing "yan-jing" with Tone 1 (top panels) vs. Tone 4 (bottom panels). The patterns were delimited in time with a 20-dB cutoff from the intensity peak, and the resulting duration was normalized from 0 to 100% of the total duration to enable averaging across repetitions and participants.

FIGURE 4 | (Top-left) Intensity peak of the first and second syllable of the word "yan-jing" spoken as Tone 1 or Tone 4, averaged across all participants in each population. (Top-right) Intensity peak of each tone relative to the syllable preceding it. All children tended to soften the intensity of Tone 1 relative to Tone 4. (Bottom) Difference in relative intensity peak between the two lexical tones, plotted across participants as a function of their chronological age. A lower negative value is indicative of a stronger tendency to soften Tone 1 relative to Tone 4.

effect of contrast [$t(446) = 3.1$, $p = 0.002$], and no interaction [$t(446) = 0.8$, $p = 0.411$].

We further examined the individual differences in contrasting the two lexical tones, in relative intensity peak across participants (bottom panel). Here, a negative value indicates that the participant produced a softer intensity for Tone 1 than for Tone 4 (intensity being normalized by what occurred in the first syllable). This was the case for 77% of the participants with NH and 73% of the participants with CIs. The difference was found to be weakly related to chronological age only among NH children ($p = 0.045$, although this would not survive Bonferroni correction). Among children with CI, this behavior was not predicted either by age at implantation ($p = 0.638$) or length of device experience ($p = 0.833$). In summary, all children utilized intensity to some degree to contrast the two tones.

F0 Pattern

Figure 5 shows the mean F0 pattern for each syllable in each lexical tone, normalized in duration (by resampling 100 points over the length of the pattern) and normalized in its scale (by expressing F0 in semitones relative to the mean over the first syllable). The F0 contour exhibited in the first syllable (left panels) was supposedly a falling-rising contour, but this pattern was washed away to some degree in the averaging process, since the timing of the local minimum varied considerably across repetitions and across participants. More importantly, this pattern was similar whether it preceded Tone 1 or Tone 4, decreasing within a 2–3 semitones scale,

and similar for both subject groups, allowing for a consistent reference with which to compare F0 patterns in the second syllable. The top-right panel of **Figure 5** shows that participants with NH expressed Tone 1 by starting about 3 semitones above the preceding syllable and slowly raised their voice pitch to another 2–3 semitones higher. Participants with CIs also started about three semitones above the preceding syllable but did not raise their voice over the course of the tone. Both participant groups expressed Tone 4 by dropping their voice pitch by 4–5 semitones (bottom-right panel). Two specific analyses were performed, one based on the F0 median relative to the precedent syllable, and the other based on the F0 movement calculated as the difference between the first and last 30-ms portion.

F0 Median

The LME analysis revealed an effect of hearing status [$t(446) = -2.7$, $p = 0.007$], an effect of lexical tone contrast [$t(446) = -4.1$, $p < 0.001$], and an interaction [$t(446) = 2.6$, $p = 0.009$]. As displayed in the top-left panel of **Figure 6**, participants with NH raised their voice pitch relative to the first syllable by over 4 semitones to express Tone 1, whereas participants with CIs did it to a smaller degree. Seen across participants (bottom-left panel), twelve children with CI changed their voice pitch between the syllables by fewer than two semitones, whereas practically all NH children did it by more than two semitones. This is evidence that at least a fraction of the CI population exhibited a relatively monotonous production

FIGURE 5 | F0 pattern extracted for each syllable in each tone, with normalized duration (extracted with a 20-dB cutoff from the intensity peak) and normalized in its scale by expressing F0 in semitones relative to the mean over the first syllable. Lines represent the mean over all participants in a given group and areas represent one standard error of the mean.
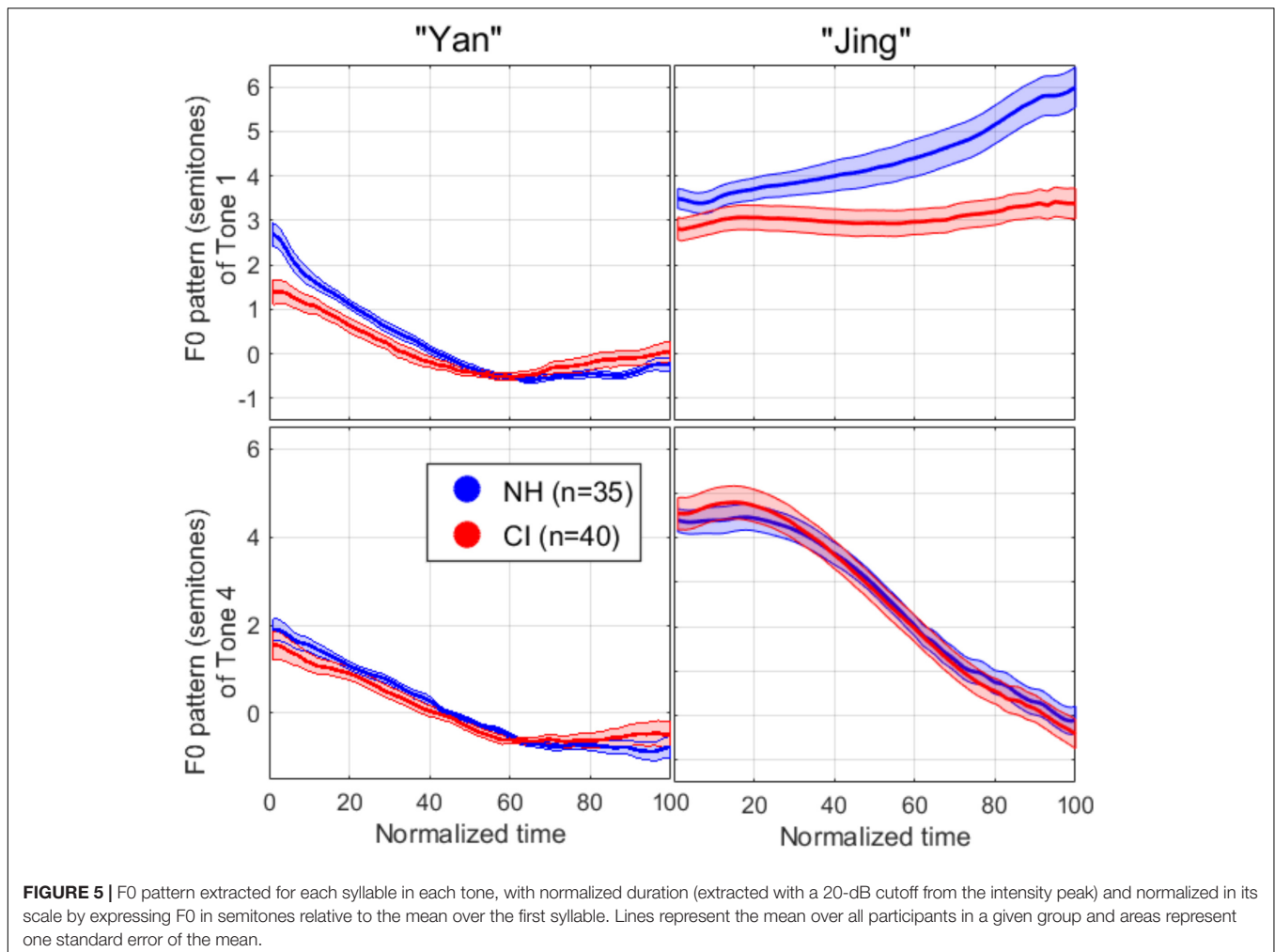
since they were not able to indicate Tone 1 as "high" (although they were able to indicate it as "flat" – see next section).

F0 Movement

The LME analysis revealed an effect of hearing status [$t(446) = -4.7$, $p < 0.001$], an effect of contrast [$t(446) = -13.8$, $p < 0.001$], and a significant interaction [$t(446) = 2.2$, $p = 0.027$]. As displayed in the top-right panel of **Figure 6**, NH and CI groups differed primarily in the rising versus flat contour of Tone 1. To produce Tone 1 (not shown), 72% of CI participants exhibited a slightly rising F0 contour while the rest exhibited a downward movement. NH participants produced a more accentuated rising contour which contributed to the difference in F0 median aforementioned. To produce Tone 4 (bottom-right panel of **Figure 6**), all but two participants exhibited a downward movement (−4.5 semitones on average). Interestingly, younger children were more likely to produce a steep downward movement than were older children.

We also examined the extent to which these F0 parameters could depend on years of CI use (**Figure 7**). This experience-related factor was a stronger predictor than chronological age in explaining how much participants with CIs dropped

their voice pitch within Tone 4. As displayed on the right panel, participants with the longest experience with their CIs produced Tone 4 with the shallowest falling slope ($p = 0.016$) accounting for 14% of the variance. Note that excluding one subject with 15 years of experience (16.7 years old, the second oldest of our sample) whose productions were very good, this relationship strengthened considerably ($r^2 = 0.24$, $p = 0.002$). In addition, there was a non-significant trend, where the long-term users produced smaller differences in F0 median between the two syllables when producing Tone 1 (left panel), and this relationship was considerably strengthened by ignoring the same 16.7 years old subject ($r^2 = 0.19$, $p = 0.006$). Despite a large inter-subject variability that is often observed among CI users, there is some evidence that long-term CI experience was associated with a more monotonous F0 production.

Role of Acoustic Hearing

Although all participants were tested with their earlier-implanted CI only, they varied in their everyday device configurations. A between-subjects analysis of variance in the production outcomes discussed above (1: difference between the two tones in duration ratio, 2: difference between the two tones in intensity

FIGURE 6 | (Top-left) F0 median over the second syllable, expressed relatively to that in the first syllable. A higher value implies the use of a higher pitch range relative to syllable 1, and is particularly relevant for Tone 1's examination. (Top-right) F0 movement calculated as the difference between the last and first 30-ms of the F0 pattern over the second syllable only. A negative value means a falling inflection, and is particularly relevant for Tone 4's examination. (Bottom-left) F0 median for Tone 1 and (bottom-right) F0 movement for Tone 4, plotted across participants as a function of their chronological age.

FIGURE 7 | F0 median for Tone 1 (left) and F0 movement for Tone 4 (right) plotted across participants as a function of their years of CI use. The CI participants who had used their CI for the longest time exhibited less modulation of their vocal chords either to differentiate the pitch range between syllables (as in the case of Tone 1) or to indicate the direction of a pitch sweep (as in the case of Tone 4).

ratio, 3: F0 median over Tone 1 relative to the preceding syllable, and 4: F0 movement over Tone 4) was conducted one by one, with Bonferroni corrections, based on whether the listeners were Bimodal ($N = 18$, 45%), Unilateral-CI ($N = 15$, 37.5%) or Bilateral-CI ($N = 7$, 17.5%) users. The results did not show consistent patterns. No significant differences were observed between the groups in duration characteristics [$F_{(2,37)} = 2.4$, $p = 0.108$], and only a marginal difference in intensity characteristics [$F_{(2,37)} = 3.3$, $p = 0.046$] driven by a significant difference between Unilateral-CI and Bilateral-CI users ($p = 0.036$). Another marginal effect of group was observed for F0 median [$F_{(2,37)} = 3.1$, $p = 0.059$], driven by

a difference between Unilateral-CI and Bimodal listeners, with Bimodal listeners producing a higher F0 median than Unimodal listeners for Tone 1 ($p = 0.047$). However, this effect was unlikely to be due to residual hearing, because there was no difference between Bilateral-CI and Bimodal users ($p = 0.780$). Finally, differences between the groups in F0 movement also failed to reach significance [$F_{(2,37)} = 3.0$, $p = 0.064$], and did not point either toward a benefit of residual hearing (i.e., Bimodal users producing F0 drops of about −5 semitones, while Bilateral-CI and Unilateral-CI users produced drops of −6 and −3.5 semitones, respectively, and no pairwise comparison reached significance). Also, the groups were different in chronological age,

with Unilateral-CI users being significantly older than bimodal users (mean ages 13.4 vs. 8.9 years of age, $p < 0.001$) and marginally older than bilateral users (13.4 vs. 10.4, $p = 0.055$). As duration of device experience co-varied with chronological age, this may have also contributed to the differences between groups.

PERCEPTION RESULTS

The data for the perception task are shown in **Figure 8**. A logistic mixed-effect analysis revealed a significant interaction between population and the slope of F0 variation [$t(20451) = 11.4$, $p < 0.001$]. The proportion of Tone 1 responses for NH children rose dramatically from about 20 to 90% (on average over the two F0 heights) as the F0 drop changed in a subtle manner between -0.3 to -0.1 octave. For participants with CIs, the

proportion of Tone 1 responses varied more gradually between 20 and 75% over the entire scale of F0 variation. As a consequence, the estimated coefficient for F0 variation differed between the groups: -20.7 and -4.3 for NH and CI participants, respectively (**Figure 9**). There was a main effect of duration [$t(20451) = 10.4$, $p < 0.001$], favoring Tone 1 responses the longer the syllable. Its estimated coefficient was 6.6, and it did not differ between NH and CI participants [$t(20451) = -0.5$, $p = 0.604$]. Finally, there was an interaction between the two experimental manipulations, F0 variation and duration [$t(20451) = -4.9$, $p < 0.001$], which itself interacted in a 3-way with population [$t(20451) = 2.8$, $p = 0.005$]. This can be appreciated when considering that NH children made use of duration only when the pitch contours were ambiguous (F0-slopes of -0.3 to -0.1 octave), whereas CI children made use of duration cues throughout all manipulations. Notably, at the extremes (for CI children): extending the duration

FIGURE 8 | Proportion of stimuli perceived as Tone 1 among a continuum of stimuli varying orthogonally in F0-variation, duration, and F0-height. A steeper slope of the psychometric function along a given dimension (e.g., F0 variation) reflects a stronger reliance on this cue.

FIGURE 9 | Coefficients extracted from the logistic mixed-effect model reflecting the reliance on F0-variation (left) and duration (right) in the lexical tone recognition task, as a function of the children's age.

from 40 to 140% still caused a +10% increase in Tone 1 responses when the F0 contour dropped by a full octave, and caused a +45% increase in Tone 1 responses when the F0 contour was flat.

Note that F0-height was not included as a fixed factor, as it did not represent an experimental manipulation but was included to represent the natural variability in lexical tones (male or female voices). It made overall little difference to the responses (top versus bottom panels, **Figure 8**), and estimates of the per-subject random slopes allocated to F0-height did not differ between participants with NH and with CIs [$t(73) = -1.3$, $p = 0.207$]. Also, F0-height did not correlate with any of the production parameters, and is not discussed any further.

Estimates of the per-subject slopes for F0-variation and duration were plotted across participants (**Figure 9**). To homogenize the variability between the two populations, the estimates for F0-variation were expressed in log_{10} of the absolute value. There was an age effect among participants with NH ($p = 0.049$, although it would not survive Bonferroni correction), whereby the older children placed slightly more weight (than the younger children) on the slope of F0 variation. In contrast, there was an age effect among participants with CIs, whereby the older children placed more weight on duration, explaining 20% of the variance ($p = 0.004$). Note that the present participants in the CI group were implanted before age 3 (median = 2.5 years of age); their chronological ages correlated strongly with the length of CI experience [$r^2 = 0.72$, $p < 0.001$], and a similar correlation could therefore, be obtained when the variable *chronological age* was substituted with the *length of CI experience* [$r^2 = 0.19$, $p = 0.005$]. Additionally, the participants with NH did not exhibit the same trend when compared to that of the participants with a CI [$r^2 = 0.01$, $p = 0.572$]. Taken together, these findings suggest that chronological development itself does not contribute to the observed trend that older participants with a CI placed more weight on duration cues

compared to those younger ones. In other words, rather than a developmental factor, this effect could well be driven by the opportunity to have learned cue-trading through the experience with CIs.

Finally, we addressed the question of whether the perceptual weights that a given child placed upon F0 contours and duration cues could be related to the production outcomes discussed earlier. For the NH group, we did not observe any relationship. For the CI group, however, two interesting correlations were observed. First, the participants who placed greater weight to duration cues perceptually were the individuals who exhibited little downward movement when producing Tone 4 ($p = 0.016$, right panel of **Figure 10**). This is exemplified by the two participants who relied the most on duration (coefficient of 13–14), and despite being quite different in age (9.5 vs. 16), they both expressed Tone 4 with less than a two semitones drop. Second, there was a marginal correlation ($p = 0.058$, left panel), where the users who relied more on F0 perceptually tended to raise their voice pitch more between Tone 1 and the syllable preceding it. An account based on the sensitivity to F0 would easily explain such a link: the users who are lucky enough to discriminate a static F0 difference of 1–2 semitones (Deroche et al., 2014) or track a F0 glide down to 8 semitones/s (Deroche et al., 2019) in the voice of other speakers could afford to rely on F0 contours perceptually even though the trajectory of the F0 inflection is coarse, and similarly this sensitivity may be just enough for their auditory feedback to exhibit this coarse inflection in their own production. Therefore, it could prevent the "monotonizing" impact of hearing through a CI over many years (**Figure 7**). However, it must be acknowledged that the perceptual weights on duration did not correlate with the production outcome respective to duration ratio ($p = 0.734$); and the perceptual weights on F0 variation did not correlate with the F0 movement of Tone 4 ($p = 0.266$). Therefore, on the whole, our hypothesis that "reliance on specific acoustic dimensions in the identification of tones

FIGURE 10 | Coefficients extracted from the logistic mixed-effect model reflecting the reliance on F0-variation (left) and duration (right) in the lexical tone recognition task, as a function of the F0 parameters extracted from production. Coefficients for F0 variation are expressed in log_{10} of their absolute value to improve homogeneity of variance between the two participant groups.

would be reflected in their production by the same individuals" received mixed support.

GENERAL DISCUSSION

Peng et al. (2017) reported that pediatric CI recipients used both F0 and duration cues to discriminate between Tones 1 and 4, while NH peers relied exclusively on F0 cues. This result seemed consistent with cue-trading, in which CI listeners use alternative acoustic dimensions that co-vary with F0 contour to compensate for the limited functional spectral resolution. However, when the same children were asked to identify lexical tones in a set of 40 naturally spoken words, their performance was predicted by their reliance on F0 rather than on duration cues. This makes sense considering that, in connected speech, the four lexical tones actually show little difference in duration. With minimal semantic or linguistic context, it is hard to see how those children could indeed make use of duration cues. In other words, while pediatric CI users may rely on amplitude and/or duration cues as additional sources of information to perceive lexical tones, it is their sensitivity to F0 contours that predicts lexical tone recognition in everyday listening. Duration cues may not be very helpful at the sentence level, and as such, the degraded F0 contour may still be the more reliable information in ecological situations.

This begged the question of whether CI children would still attempt to modulate their voice pitch despite ignoring how well they succeeded in doing so, or whether they would attempt to convey those tones via other dimensions that they have adequate representation of, even though these co-varying cues may not be relied upon by NH listeners. The present results provided several key points, as follows.

First, pediatric CI recipients produced Tone 4 shorter than Tone 1 (**Figure 2**). This behavior simply exaggerated that of NH participants (more easily observable with voicing duration), which is why CI users were able to produce meaningful tone distinctions using duration cues. This finding highlights that the patterns produced by all participants reflect to some degree the biological or mechanical constraints of human vocal production. Thus, CI users cannot produce tones in an arbitrary way; they can only refine their productions based on what is acceptable and meaningful in the natural lexicon. On a side note, it is notable that the participants with CIs exhibited longer vowels than their NH peers, and this may not be coincidental. By slowing down their speech overall, these children increase their capacity to shorten specific syllables when they need it, without reaching a narrow window, where this might conflict with audibility. Additionally, speaking slowly tempers fast fluctuations in intensity, i.e., it gives them a better control over loudness changes. These results are consistent with findings of Chuang et al. (2012), where CI children are reported to exhibit longer vowels as well as longer pauses between words, resulting in a slower speaking rate than their NH peers matched in age, sex, and educational level.

Second, while individuals with NH stressed the second syllable relative to the first in both Tones (relying on pitch to convey the tone identity), pediatric CI recipients tended to soften Tone

1 relative to the syllable preceding it (**Figure 4**). However, we did not capture a trait of the intensity pattern that would highlight a significant interaction between population and tone. Rather, a marginal effect of group ($p = 0.061$) showed that participants with CIs simply did not emphasize the second syllable as much as their NH peers did. Perhaps, they are less aware of which syllable contains to the critical information that distinguishes the two tones and consequently do not feel the need to emphasize it. Arguably, compression of dynamic range in the auditory feedback that CI children received from their own voice should hinder their ability to detect small increments in loudness. However, this explanation suffers from the clear difference between the two tones (>2 dB) that CI children successfully exhibited (as NH children did). Another factor that could have played a role here is the fact that NH children listened binaurally to their voice's feedback while CI users listened monaurally (regardless of whether they used another CI or a HA in their everyday life). This means that several of the CI children did not experience the binaural summation they are normally used to experience, and this could have led them to speak louder (for both tones).

Third, there were signs (although subtle) of atypical F0 productions among some pediatric CI recipients. This was reflected by a lower tendency to (1) mark the F0 median of Tone 1 as *higher* than the syllable preceding it, and (2) mark the F0 movement of Tone 4 as *falling*. However, several notes of caution must be acknowledged. The first observation was partly accounted for by a difference in F0 movement between the two groups. Children with NH actually raised the F0 over the course of Tone 1 by 2 semitones, while children with CI produced it as flat (as it is supposed to be, in isolation). This raises a doubt about NH children's production quality which must be answered. Much of the literature on Mandarin's phonetics is based on monosyllabic words. Xu (1997) demonstrated that with bi-tonal sequences, there are anticipatory and carry-over effects to consider, and most relevant here, "*a considerable portion of the F0 curve for Tone 1 has a rising contour when the preceding tone is Tone 3 or Tone 4, both of which have a low offset*" (p. 69). Thus, the present behavior of NH children is perfectly normal and expected, given that Tone 3 was used in the first syllable. In contrast, the fact that CI children stuck to the canonical form of Tone 1 indicates that they did not take the tonal context into consideration. The second observation is also debatable since, at a population level, there was no difference between the two groups in F0 movement for Tone 4; only the older CI users reduced the degree of their drop in F0. Therefore, in both measures (F0 median for Tone 1 or F0 movement for Tone 4), we think that the interesting finding is about the effect of CI experience (**Figure 7**), rather than a deficit of the entire population.

The effect of CI experience raises discrepancies among earlier studies. In earlier studies, the quality of lexical tone productions was rated (by a NH adult or a machine), and those ratings were dominated by the quality of the F0 contour. Hence, such ratings should in principle be consistent with acoustic parameters such as those presented here. Han et al. (2007) reported that CI experience was *beneficial* to production ratings (with $N = 14$). However, Zhou et al. (2013) failed to replicate this finding with

a large sample size (N = 110). Earlier, Peng et al. (2004) did not find such a benefit of experience (with N = 30) and this was not possible for Xu et al. (2004) to investigate, as their sample included only 4 CI children. Here, we found CI experience to be rather *detrimental* to the quality of F0 productions. Arguably, the linear trends were modest (e.g., accounting for only 14% of the variance in the right panel of **Figure 7**), but we note that those CI users who produced the shallower falling tone were also the subjects who placed greater perceptual weights on duration (**Figure 10**, right panel, also accounting for 14% of the variance). We do not believe this relation to be coincidental, and it suggests a "monotonizing" process (and a shift in perceptual strategies) that takes place over years of hearing through the device, perhaps as many as 20 years given the current slopes. The most trivial interpretation is that the poor feedback of voice pitch provided by current CIs reinforces the percept of a monotonous voice and over many years, some CI users adapt and no longer modulate their vocal chords (as this seems to have no impact on their auditory feedback). This being said, children with CIs have ample opportunities to receive direct or indirect feedback from caregivers, teachers, clinicians, and other NH children, on how to produce better F0 contours to enhance their intelligibility. These interactions should mitigate the saliency of the monotonous voice percept, but perhaps it is difficult to learn F0 control from outsiders' advice.

Rather than a "monotonizing" impact of CI experience, an alternative explanation is that pediatric CI recipients exhibited a stronger developmental trajectory in their F0 control than did their NH peers. Adults and older children generally speak with narrower fluctuations in F0 than do young children (e.g., review by Kent, 1976). Older CI users could have spoken on a narrow range of F0 fluctuations, not because their voice was monotonous per se, but because they had already refined the control of their vocal chords to operate within a range that is just enough to convey the tonal information. This interpretation suffers from two weaknesses: (1) the correlations with CI experience were stronger than those with chronological age of CI children (**Figure 7** vs. **Figure 6**), and (2) there was no effect of chronological age among NH children (**Figure 6**). However, the developmental trajectories of CI children are known to differ from those of NH children, and it is easy to imagine that the refinement process in F0 variability requires hearing. So, this interpretation should certainly not be discarded until one can test precisely whether these F0 fluctuations would eventually (with very long-term exposure) flatten or show a similarly narrow distribution as for NH adults.

One of the most efficient ways to disambiguate such interpretations is to examine production with and without auditory feedback, i.e., by turning the CI on and off. Such studies differ in their outcome, with some showing differences between the two conditions (Poissant et al., 2006; Bharadwaj et al., 2007) and others finding no difference in the acoustics of their speech (Tye-Murray et al., 1996; Turgeon et al., 2017). Applied to lexical tones, similar designs would be greatly informative. Also critical is the fact that the mechanics of speech production may actually differ (e.g., when the feedback is on or off) even when no acoustic difference is observed in the recordings, which is why articulatory

measures may eventually be necessary to fully understand the abnormal vocal production by CI users and their relation to experience-related plasticity (Turgeon et al., 2015, 2017).

Fourth, a number of results in the lexical tone recognition task were found to be consistent with previous findings (Peng et al., 2017): (1) the tonal boundary along a continuum of F0 slopes was very sharp for NH children but more gradual for CI children; (2) the tonal boundary along a continuum of compressed to stretched syllables was generally shallower (than for the F0 slope) and CI children relied on duration across the entire range of F0 slopes, whereas NH children used it only in very few cases, where the F0 slope was ambiguous; (3) as they got older, NH children relied even more on F0 cues while CI children relied even more on duration cues. This latter finding is particularly important because it is potentially the reason why prelingually deafened CI users improve over time in this task, i.e., not because they somehow get better at processing F0 contours but because they have learnt to detect other cues. This interpretation would seem consistent with a study by Tao et al. (2015) who observed considerable deficits in melodic contour identification by Mandarin-speaking CI users (aged 6–26 years) while performing well in a lexical tone recognition task. Also, performance in the two tasks was not correlated among the 33 users in their study (children and adults, pre- or post-lingual). The authors concluded that CI users must compensate their deficits in F0 processing by using the amplitude and duration cues in lexical tones. Note that this learning to trade among cues must take place while hearing, but among prelingually deafened children, it is always difficult to disentangle developmental effect from that of CI experience itself. Zhou et al. (2013) reported CI experience to account for 18% of the variance in lexical tone identification; in very good agreement, we found it to account for 19%.

Finally, the present study focused on the contrast between Tone 1 vs. Tone 4, as this pair permitted us to examine the changes in perceptual weighting between two acoustic dimensions (F0 and duration) known to contribute to lexical tone recognition for Mandarin Chinese. This Tone 1 vs. Tone 4 contrast is also suitable for our targeted patient populations and listeners who are relatively young in age, given the relatively simple linguistic meanings of the chosen bi-syllabic words with these two lexical tones (i.e., *eyes* vs. *eyeglasses*), in addition to the fact that they do not involve complex contour changes as with Tones 2 and 3 (Peng et al., 2017). Ideally, it would be necessary to replicate the present findings with other pairs of lexical tones and considering different tonal context environments.

CONCLUSION

This study analyzed acoustic recordings of Mandarin Chinese, pediatric CI recipients, and their age-matched peers with NH. All participants were asked to produce disyllabic words with contrastive lexical tones (i.e., Tones 1 and 4). Pediatric CI recipients, at least the older and more experienced ones, exhibited narrower modulations of their voice pitch (both within and across syllables). However, it remains unclear whether this

represents a "monotonizing" impact of CI experience or rather a refinement in the control of vocal chords to convey the tonal information more like adults. Perhaps as a compensatory mechanism, CI children contrasted the duration properties of the second syllable that distinguish Tones 1 and 4. To explore this interplay further and link it to perception, the same children took part in a lexical tone recognition task, discriminating among parametric variations of many tokens in a Tone 1–Tone 4 continuum. The perceptual weights extracted from this task confirmed that CI children relied less on F0 cues than did NH children. CI children used duration cues all the time, whereas NH children used them only when F0 cues were ambiguous. CI children who placed greater weight on duration cues also tended to have the most monotonous tone production. This result supports the idea that perception and production are reasonably coupled, even with this clinical population having an auditory feedback of relatively poor quality.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Institutional Review Board at the Chi-Mei Medical Center which approved the protocol. All subjects gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

MD analyzed the data and wrote the manuscript. H-PL collected all the data. Y-SL was responsible for the project supervision in Taiwan. MC obtained funding for this research program, worked closely with S-CP on the rationale, with MD on the analyses, and edited the manuscript. S-CP developed the rationale, designed the experimental tasks, contributed to analyses, and edited the manuscript.

REFERENCES

Barry, J. G., and Blamey, P. J. (2004). The acoustic analysis of tone differentiation as a means for assessing tone production in speakers of Cantonese. *J. Acoust. Soc. Am.* 116, 1739–1748. doi: 10.1121/1.1779272

Bharadwaj, S. V., Graves, A. G., Bauer, D. D., and Assmann, P. F. (2007). Effects of auditory feedback deprivation length on the vowel /epsilon/ produced by pediatric cochlear implant users. *J. Acoust. Soc. Am. EL* 121, 196–202.

Boersma, P., and Weenink, D. (2018). *Praat: Doing Phonetics by Computer" [computer program]. Version 6.0.37.* Available at: http://www.praat.org/ (accessed January 1, 2018).

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., and Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *J. Acoust. Soc. Am.* 101, 2299–2310. doi: 10.1121/1.418276

Chatterjee, M., and Peng, S. C. (2008). Processing F0 with cochlear implants: modulation frequency discrimination and speech intonation recognition. *Hear. Res.* 235, 143–156. doi: 10.1016/j.heares.2007.11.004

Chatterjee, M., Zion, D. J., Deroche, M. L. D., Burianek, B. A., Limb, C. J., Goren, A. P., et al. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear. Res.* 322, 151–162. doi: 10.1016/j.heares.2014.10.003

Chen, Y., Wong, L. L., Chen, F., and Xi, X. (2014). Tone and sentence perception in young Mandarin-speaking children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 78, 1923–1930. doi: 10.1016/j.ijporl.2014.08.025

Chuang, H. F., Yang, C. C., Chi, L. Y., Weismer, G., and Wang, Y. T. (2012). Speech intelligibility, speaking rate, and vowel formant characteristics in Mandarin-speaking children with cochlear implant. *Int. J. Speech Lang. Pathol.* 14, 119–129. doi: 10.3109/17549507.2011.639391

Ciocca, V., Francis, A. L., Aisha, R., and Wong, L. (2002). The perception of Cantonese lexical tones by early-deafened cochlear implantees. *J. Acoust. Soc. Am.* 111, 2250–2256.

Cullington, H., and Zeng, F. G. (2008). Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects. *J. Acoust. Soc. Am.* 123, 450–461. doi: 10.1121/1.2805617

Deroche, M. L. D., Lu, H.-P., Kulkarni, A. M., Caldwell, M., Barrett, K. C., Peng, S.-C., et al. (2019). A tonal-language benefit for pitch in normally-hearing and cochlear-implanted children. *Sci. Rep.* 9:109. doi: 10.1038/s41598-018-36393-1

Deroche, M. L. D., Lu, H.-P., Limb, C. J., Lin, Y.-S., and Chatterjee, M. (2014). Deficits in the pitch sensitivity of cochlear-implanted children speaking English or Mandarin. *Front. Neurosci.* 8:282. doi: 10.3389/fnins.2014.00282

Elman, J. L. (1981). Effects of frequency-shifted feedback on the pitch of vocal productions. *J. Acoust. Soc. Am.* 70, 45–50. doi: 10.1121/1.386580

Fu, Q. J., Chinchilla, S., and Galvin, J. J. III (2004). The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users. *J. Assoc. Res. Otolaryngol.* 5, 253–260. doi: 10.1007/s10162-004-4046-1

Galvin, J. J. III, Fu, Q. J., and Nogaki, G. (2007). Melodic contour identification by cochlear implant listeners. *Ear Hear.* 28, 302–319. doi: 10.1097/01.aud.0000261689.35445.20

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear Hear.* 28, 412–423. doi: 10.1097/aud.0b013e3180479318

Han, D., Zhou, N., Li, Y., Chen, X., Zhao, X., and Xu, L. (2007). Tone production of Mandarin Chinese speaking children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 71, 875–880. doi: 10.1016/j.ijporl.2007.02.008

Hazan, V., and Barrett, S. (2000). The development of phonetic categorization in children aged 6-12. *J. Phonet.* 28, 377–396. doi: 10.1006/jpho.2000.0121

Houde, J. F., and Jordan, M. I. (1998). Sensorimotor adaptation in speech production. *Science* 279, 1213–1216. doi: 10.1126/science.279.5354.1213

Kent, R. D. (1976). Anatomical and neuromuscular maturation of the speech mechanism: evidence from acoustical studies. *J. Speech Hear. Res.* 19, 421–447. doi: 10.1044/jshr.1903.421

Kong, Y. Y., Stickney, G. S., and Zeng, F. G. (2005). Speech and melody recognition in binaurally combined acoustic and electric hearing. *J. Acoust. Soc. Am.* 117, 1351–1361. doi: 10.1121/1.1857526

Larson, C. R., Burnett, T. A., Kiran, S., and Hain, T. C. (2000). Effects of pitch-shift velocity on voice Fo responses. *J. Acoust. Soc. Am.* 107, 559–564. doi: 10.1121/1.428323

Li, C. N., and Thompson, S. (1977). The acquisition of tone in Mandarin-speaking children. *J. Child Lang.* 4, 185–199. doi: 10.1017/s0305000900001598

Liu, S., and Samuel, A. G. (2004). Perception of Mandarin lexical tones when F0 information is neutralized. *Lang. Speech* 47, 109–138. doi: 10.1177/00238309040470020101

Liu, Y.-W., Tao, D.-D., Chen, B., Cheng, X., Shu, Y., Galvin, J. J. III, et al. (2019). Factors affecting bimodal benefit in pediatric Mandarin-speaking Chinese cochlear implant users. *Ear Hear.* doi: 10.1097/AUD.0000000000000712 [Epub ahead of print].

Luo, X., Fu, Q. J., Wei, C. G., and Cao, K. L. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear Hear.* 29, 957–970. doi: 10.1097/AUD.0b013e3181888f61

Nittrouer, S. (1996). Discriminability and perceptual weighting of some acoustic cues to speech perception by 3-year-olds. *J. S. Lang. Hear. Res.* 2, 278–297. doi: 10.1044/jshr.3902.278

Peng, S.-C., Chatterjee, M., and Lu, N. (2012). Acoustic cue integration in speech intonation recognition with cochlear implants. *Trends Amplif.* 16, 67–82. doi: 10.1177/1084713812451159

Peng, S.-C., Lu, H.-P., Lu, N., Lin, Y.-S., Deroche, M. L. D., and Chatterjee, M. (2017). Processing of acoustic cues in lexical-tone identification by pediatric cochlear-implant recipients. *J. Speech Lang. Hear. Res.* 60, 1223–1235. doi: 10.1044/2016_JSLHR-S-16-0048

Peng, S.-C., Lu, N., and Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiol. Neuro Otol.* 14, 327–337. doi: 10.1159/000212112

Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y. S., and Wang, L. S. (2004). Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear.* 25, 251–264. doi: 10.1097/01.aud.0000130797.73809.40

Peng, S. C., Tomblin, J. B., and Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear Hear.* 29, 336–351. doi: 10.1097/AUD.0b013e318168d94d

Poissant, S. F., Peters, K. A., and Robb, M. P. (2006). Acoustic and perceptual appraisal of speech production in pediatric cochlear implant users. *Int. J. Pediatr. Otorhinolaryngol.* 70, 1195–1203. doi: 10.1016/j.ijporl.2005.12.008

Qin, M. K., and Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *J. Acoust. Soc. Am.* 114, 446–454. doi: 10.1121/1.1579009

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear. Res.* 11, 157–189. doi: 10.1016/0378-5955(83)90077-1

Tao, D., Deng, R., Jiang, Y., Galvin, J. J. I. I. I., Fu, Q.-J., and Chen, B. (2015). Melodic pitch perception and lexical tone perception in Mandarin-speaking cochlear implant users. *Ear Hear.* 36, 102–110. doi: 10.1097/AUD.0000000000000086

Turgeon, C., Premont, A., Trudeau-Fisette, P., and Ménard, L. (2015). Exploring consequences of short- and long-term deafness on speech production: a lip-tube perturbation study. *Clin. Lingu. Phon.* 29, 378–400. doi: 10.3109/02699206.2015.1007527

Turgeon, C., Trudeau-Fisette, P., Fitzpatrick, E., and Ménard, L. (2017). Vowel intelligibility in children with cochlear implants: an acoustic and articulatory study. *Int. J. Pediatr. Otorhinolaryngol.* 101, 87–96. doi: 10.1016/j.ijporl.2017.07.022

Tye-Murray, N., Spencer, L., Bedia, E. G., and Woodworth, G. (1996). Differences in children's sound production when speaking with a cochlear implant turned on and turned off. *J. Speech Lang. Hear. Res.* 39, 604–610. doi: 10.1044/jshr.3903.604

Wang, S., Liu, B., Dong, R., Zhou, Y., Li, J., Qi, B., et al. (2012). Music and lexical tone perception in Chinese adult cochlear implant users. *Laryngoscope* 122, 1353–1360. doi: 10.1002/lary.23271

Wang, S., Liu, B., Zhang, H., Dong, R., Mannell, R., Newall, P., et al. (2013). Mandarin lexical tone recognition in sensorineural hearing-impaired listeners and cochlear implant users. *Acta Otolaryngol.* 133, 47–54. doi: 10.3109/00016489.2012.705438

Wang, W., Zhou, N., and Xu, L. (2011). Musical pitch and lexical tone perception with cochlear implants. *Int. J. Audiol.* 50, 270–278. doi: 10.3109/14992027.2010.542490

Whalen, D. H., and Xu, Y. (1992). Information for Mandarin tones in the amplitude contour and in brief segments. *Phonetica* 49, 25–47. doi: 10.1159/000261901

Winn, M. B., Chatterjee, M., and Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing. *J. Acoust. Soc. Am.* 131, 1465–1479. doi: 10.1121/1.3672705

Winn, M. B., Chatterjee, M., and Idsardi, W. J. (2013). Roles of voice onset time and F0 in stop consonant voicing perception: effects of masking noise and low-pass filtering. *J. Speech Lang. Hear. Res.* 56, 1097–1107. doi: 10.1044/1092-4388(2012/12-0086)

Wu, J. L., and Yang, H. M. (2003). Speech perception of Mandarin Chinese speaking young children after cochlear implant use: effect of age at implantation. *Int. J. Pediatr. Otorhinolaryngol.* 67, 247–253. doi: 10.1016/s0165-5876(02)00378-6

Xu, L., Chen, X., Lu, H., Zhou, N., Wang, S., Liu, Q., et al. (2011). Tone perception and production in pediatric cochlear implant users. *Acta-Oto-Laryngol.* 131, 395–398. doi: 10.3109/00016489.2010.536993

Xu, L., Li, Y., Hao, J., Chen, X., Xue, S. A., and Han, D. (2004). Tone production in Mandarin-speaking children with cochlear implants: a preliminary study. *Acta Otolaryngol.* 124, 363–367. doi: 10.1080/00016480410016351

Xu, Y. (1997). Contextual tonal variations in Mandarin. *J. Phonet.* 25, 61–83. doi: 10.1006/jpho.1996.0034

Zeng, F. G. (2002). Temporal pitch in electric hearing. *Hear. Res.* 174, 101–106. doi: 10.1016/s0378-5955(02)00644-5

Zhou, N., Huang, J., Chen, X., and Xu, L. (2013). Relationship between tone perception and production in prelingually deafened children with cochlear implants. *Otol. Neurotol.* 34, 499–506. doi: 10.1097/MAO.0b013e318287ca86

Zhou, N., and Xu, L. (2008). Development and evaluation of methods for assessing tone production skills in Mandarin-speaking children with cochlear implants. *J. Acoust. Soc. Am.* 123, 1653–1664. doi: 10.1121/1.2832623

Zhu, M., Fu, Q. J., Galvin, J. J. III, Jiang, Y., Xu, J., Xu, C., et al. (2011). Mandarin Chinese speech recognition by pediatric cochlear implant users. *Int. J. Pediatr. Otorhinolaryngol.* 75, 793–800. doi: 10.1016/j.ijporl.2011.03.009

APPENDIX

Throughout visual screening of each production (see examples in **Figure 1**), we noted that the second syllable often exhibited an elevation of F0 from the last phoneme, especially for Tone 4 (e.g., right panels in **Figure 1**). This is the kind of observation that led us to reiterate the analysis with a 10-dB cutoff to selectively capture the stereotypical shape of the tones, even though it necessarily restricted the amplitude of F0 movements. Here, we report on the statistical results derived for duration, F0 median, and F0 movement with the narrower window.

The duration of syllables extracted with a 10-dB cutoff provided a closer estimation of the voicing duration. On average, this voicing duration was reduced by about 100 ms compared to the syllable duration (which was described in **Figure 2**, top-left panels), but the overall pattern was largely similar. When expressed as a ratio between the two syllables, the LME analysis revealed no effect of hearing status $[t(446) = 1.1, p = 0.259]$, an effect of contrast $[t(446) = -5.0, p < 0.001]$, and an interaction between the two $[t(446) = -2.2, p = 0.028]$. Those results were therefore, qualitatively similar as those revealed with the 20-dB window, but with a stronger role for contrast (tone 4 being overall less voiced than tone 1).

When expressing F0 median in semitones relative to the first syllable, the LME analysis revealed a marginal effect of hearing status $[t(446) = -2.0, p = 0.045]$, a marginal effect of contrast $[t(446) = -1.8, p = 0.069]$, and a critical interaction between the two $[t(446) = 2.4, p = 0.015]$. When examining F0 movement, the LME analysis revealed an effect of hearing status $[t(446) = -4.2, p < 0.001]$, a strong effect of contrast $[t(446) = -14.9, p < 0.001]$, and a critical interaction between the two $[t(446) = 2.3, p = 0.022]$. Again, these outcomes were similar to the ones aforementioned. Overall, therefore, the size of the window considered (20 or 10 dB drop from the intensity peak) made qualitatively no difference to the results described in the rest of the manuscript.

Assessing Cerebral White Matter Microstructure in Children With Congenital Sensorineural Hearing Loss: A Tract-Based Spatial Statistics Study

Muliang Jiang[1,2†], Zuguang Wen[1†], Liling Long[1*], Chi Wah Wong[3], Ningrong Ye[2], Chishing Zee[4] and Bihong T. Chen[2]

[1] Department of Radiology, First Affiliated Hospital of Guangxi Medical University, Nanning, China, [2] Department of Diagnostic Radiology, City of Hope National Medical Center, Duarte, CA, United States, [3] Center for Informatics, City of Hope National Medical Center, Duarte, CA, United States, [4] Department of Radiology, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

Objectives: To assess the microstructural properties of cerebral white matter in children with congenital sensorineural hearing loss (CSNHL).

Methods: Children (>4 years of age) with profound CSNHL and healthy controls with normal hearing (the control group) were enrolled and underwent brain magnetic resonance imaging (MRI) scans with diffusion tensor imaging (DTI). DTI parameters including fractional anisotropy, mean diffusivity, axial diffusivity, and radial diffusivity were obtained from a whole-brain tract-based spatial statistics analysis and were compared between the two groups. In addition, a region of interest (ROI) approach focusing on auditory cortex, i.e., Heschl's gyrus, using visual cortex, i.e., forceps major as an internal control, was performed. Correlations between mean DTI values and age were obtained with the ROI method.

Results: The study cohort consisted of 23 children with CSHNL (11 boys and 12 girls; mean age $\pm$ SD: 7.21 ± 2.67 years; range: 4.1–13.5 years) and 18 children in the control group (11 boys and 7 girls; mean age $\pm$ SD: 10.86 ± 3.56 years; range: 4.5–15.3 years). We found the axial diffusivity values being significantly greater in the left anterior thalamic radiation, right corticospinal tract, and corpus callosum in the CSHNL group than in the control group ($p < 0.05$). Significantly higher radial diffusivity values in the white matter tracts were noted in the CSHNL group as compared to the control group ($p < 0.05$). The fractional anisotropy values in the Heschl's gyrus in the CSNHL group were lower compared to the control group ($p = 0.0015$). There was significant negative correlation between the mean fractional anisotropy values in Heschl's gyrus and age in the CSNHL group < 7 years of age ($r = -0.59$, $p = 0.004$).

Conclusion: Our study showed higher axial and radial diffusivities in the children affected by CNHNL as compared to the hearing children. We also found lower fractional anisotropy values in the Heschl's gyrus in the CSNHL group. Furthermore, we identified

negative correlation between the fractional anisotropy values and age up to 7 years in the children born deaf. Our study findings suggest that myelination and axonal structure may be affected due to acoustic deprivation. This information may help to monitor hearing rehabilitation in the deaf children.

INTRODUCTION

Congenital sensorineural hearing loss (CSNHL) is a type of deafness that occurs prior to language development. Permanent childhood hearing loss due to CSNHL is estimated to be 1.2 to 1.7 cases per 1,000 live births, and up to 30% of these affected children have profound hearing loss (Kral and O'donoghue, 2010; Paludetti et al., 2012). Congenital deprivation of sound stimuli in children with CSNHL may delay language acquisition and alter the development of auditory neural pathways, resulting in brain structural changes (Chari and Chan, 2017). Delayed diagnosis of hearing loss in infants and young children with CSNHL results in severe deficits in learning and development. In addition, severe-to-profound hearing loss causes a severe societal burden due to reduced work productivity and the need for expensive special education resources for children with prelingual deafness (Mohr et al., 2000). Cochlear implant being the most effective treatment for CSNHL should be placed within the first 3.5 to 4 years of life when the developing brain has the maximal plasticity for reorganization in order to optimize cognitive functioning (Kral and Sharma, 2012). Nevertheless, cochlear implant placed late or even in adult age may enable the patients to hear and could help to compensate for the deficits in the brain development. However, prior studies have shown that individuals with the cochlear implant placed late or in adult age may have persistent auditory deficits, and they may not gain effective speech understanding even with the cochlear implant for a long period of time (Kral and Sharma, 2012). Therefore, it is prudent to diagnose and treat children with CSNHL early in life to improve their learning and development outcomes.

Neuroimaging studies have contributed to our understanding of brain alterations and neuroplasticity in deafness. Prior studies of macrostructural differences in deaf people, compared to control participants without hearing loss, showed lower white matter (WM) volume but preserved gray matter volume in the auditory region, including the Heschl's gyrus (HG) and the adjacent temporal lobe (Hribar et al., 2014). In addition to the reported WM macrostructural changes, there are also WM microstructural alterations in people with hearing loss (Kim et al., 2009; Miao et al., 2013; Park et al., 2018). However, less is known about the WM microstructural properties specifically in children with CSNHL.

Diffusion tensor imaging (DTI) has been commonly used to study WM microstructure. DTI is an established magnetic resonance imaging (MRI) method that measures the directionality of water diffusion in the brain and is more sensitive for detecting WM microstructural alterations than volumetric methods (Douaud et al., 2011). A prior DTI study found abnormal WM microstructure in adolescents with prelingual deafness, compared to healthy control adolescents (Park et al., 2018). Their study found lower fractional anisotropy (FA) values in the WM tract of the HG and inferior fronto-occipital fasciculus (IFOF), among other WM areas, in deaf children younger than 4 years but not in older children. However, another DTI study showed lower FA values and higher radial diffusivity (RD) in bilateral superior temporal gyri, HG, and splenium of the corpus callosum in older children (10–18 years of age) with prelingual deafness (Miao et al., 2013). These results indicate that WM microstructural properties may be different in children with hearing loss, but there is limited information about how brain alters in children affected by CSNHL. Furthermore, the existing study results are inconsistent, leaving a gap in knowledge.

To address the knowledge gap, we designed the current study to test the hypothesis that children with CSHNL have different WM microstructural properties compared to the healthy controls with normal hearing. We evaluated the cerebral WM microstructural properties using two different methods, including a whole-brain voxel-wise tract-based spatial statistics (TBSS) method and a region of interest (ROI) approach focusing on the auditory cortex, i.e., the HG, using the visual cortex, i.e., the forceps major (FM) as an internal control.

MATERIALS AND METHODS

Participants

This was a study of children with CSNHL and healthy controls with normal hearing who underwent brain MRI scans with acquisition of DTI data. The eligibility criteria included the following: children between 3.5 and 18 years of age with profound CSNHL, a mean brain stem response threshold of >91 dB, auditory steady-state evoked potential >80 dB, hearing loss within the range of 250–4,000 Hz, and grossly normal MRI results for the brain and inner ear with no evidence of major structural abnormalities. None of the children with CSNHL had a history of using hearing aids upon enrollment. The exclusion criteria were the following: severe neurological disorders such as epilepsy, congenital leukodystrophy, or severe cognitive impairment such as autism or hyperkinetic syndrome, and poor-quality MRI scans that were rendered suboptimal for data analysis.

This study was carried out in accordance with all institutional, local, and national guidelines. Our study was approved by the institutional review board of our hospital. The parents or legal guardians of the children gave written informed consent in accordance with the Declaration of Helsinki. Statistical analysis was performed on the demographic information of

the participants including age and gender. For continuous variable such as age, a two-sample *t*-test was used. For categorical variables such as gender, Fisher's exact tests were used. A two-sided *p*-value less than 0.05 was considered as statistically significant.

MRI Image Acquisition

The MRI images were acquired with a Siemens Magnetom Verio 3T MRI scanner (Siemens Healthcare; Erlangen, Germany) using a 12-channel phased-array head coil (Siemens). The head position was stabilized with sponge support. All study participants were sedated for the MRI scan with oral chloral hydrate under the care of medical professionals. There were no complications with regard to the oral sedatives.

A three-dimensional magnetization-prepared rapid gradient-echo (MP-RAGE) sequence was obtained with the following protocol: slice thickness, 0.9 mm; slice interval, 0 mm; repetition time, 1,700 ms; echo time, 2.9 ms; inversion time, 900 ms; matrix, 256×256; field of view, 220×220 mm^2; and slice number, 140. The DTI scanning protocol was as follows: slice thickness, 2 mm; slice interval, 0 mm; slice number, 60; repetition time, 9,000 ms; echo time, 93 ms; and imaging matrix, 128×128. Motion-probing gradients were applied along 30 non-collinear directions with a *b* factor of 1,000 s/mm^2 after an acquisition without diffusion weighting ($b = 0$ s/mm^2). Additional parameters included the following: diffusion-sensitive gradient direction, 30; field of view, 220×220 mm^2; number of excitations, 2.

Diffusion Tensor Imaging Preprocessing and Tract-Based Spatial Statistics Analysis

The FMRIB Software Library (FSL) was used for DTI data preprocessing (version 5.0; FMRIB, Oxford, United Kingdom)[1]. TBSS[2] was used to perform voxel-wise statistical analysis (Smith et al., 2006). First, the raw diffusion datasets were corrected for eddy current effects (Andersson and Sotiropoulos, 2016). Then, we used the Brain Extraction Tool for brain extraction (Smith, 2002). Subsequently, we fitted the preprocessed diffusion data into a tensor model to create the FA images (Basser et al., 1994). A non-linear registration was used to align the FA data from all the subjects into a common space, the Montreal Neurological Institute (MNI) MNI152 space (Fonov et al., 2009, 2011). A study-specific FA template was built as the target FA image for our study cohort, which was younger than 18 years. A mean FA image of all subjects was created according to a previously described method (Andersson et al., 2007). This mean FA image was then thinned to create a mean FA skeleton, which represented the centers of all the tracts common to all subjects. Each subject's aligned data were then projected onto the MNI template image, and voxel-wise cross-subject statistics were applied to the data. The maps generated and analyzed were FA, mean diffusivity (MD), RD, and axial

diffusivity (AD) according to a previously described method (Wheeler-Kingshott and Cercignani, 2009). λ1, λ2, and λ3 were the three eigenvalues of diffusion, while λ1 was also called AD (**Figure 1**), and averaging the maps of λ2 and λ3 represented RD (**Figure 2**).

Voxel-wise statistical analyses were performed using the FSL tool, Randomise[3]. This method implemented permutation-based inference (Wang et al., 2007). Age was treated as a covariant in the whole-brain TBSS analysis but not in the ROI analysis. Multiple comparison correction was performed using both the threshold-free cluster enhancement and cluster-based thresholding options (Smith and Nichols, 2009).

Region of Interest Analysis Focusing on Auditory Cortex and Correlation Between Diffusion Tensor Imaging and Age

Additional DTI analysis was performed with an ROI approach focusing on the auditory cortex, i.e., the WM tracts leading to HG, using the visual cortex, i.e., the FM of corpus callosum connecting the bilateral visual cortex as internal control as seen in **Figure 3**. We extracted the mean DTI values from HG and FM. Two-sample *t*-test was used to compare between-group differences in the mean FA, AD, and RD values for each tract of interest. The $p < 0.05$ was considered statistically significant. Correlations between mean FA, AD, and RD values and age in each group were obtained with the ROI methods focusing on HG while using FM as an internal control by computing pairwise Pearson correlation coefficients and the associated *p*-values before and after stratification of each group with a cutoff age of 7 years. The CSNHL group was then divided into two subgroups, i.e., the CSNHL group < 7 years and the CSNHL group ≥ 7 years, and the control group was divided into two subgroups in a similar manner with a cutoff age of 7 years. A value of $p < 0.05$ was considered statistically significant after Bonferroni correction for multiple comparisons.

RESULTS

Demographic Information

A total of 34 children over 3.5 years old with profound CSNHL were enrolled in the study between February 1, 2017, and October 1, 2018, and each participant underwent brain MRI scanning for acquisition of DTI data and 3D T1-weighted sequences. Eleven children were subsequently excluded from the study: six children had major neurological diseases or major cognitive impairment, and five children with CSNHL had poor-quality MRI brain scans that prevented optimal imaging analysis. The final study cohort consisted of 23 children with profound CSNHL (11 boys and 12 girls; mean age ± SD: 7.21 ± 2.67 years; range: 4.1–13.5 years). The control group included 18 age-matched children (11 boys and 7 girls; mean age ± SD: 10.86 ± 3.56 years; range: 4.5–15.3 years). All study participants in both the CSNHL group and the control group were right-handed (**Table 1**).

---

[1] http://www.fmrib.ox.ac.uk/fsl
[2] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/TBSS/

[3] https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Randomise
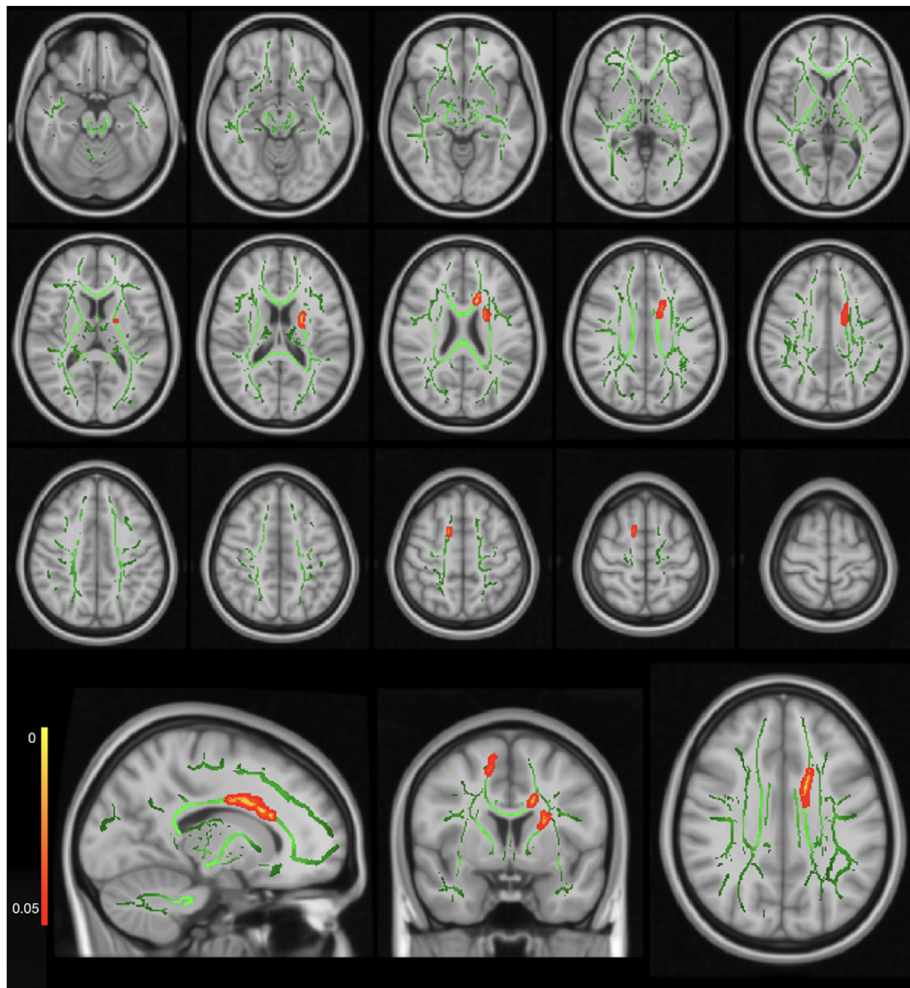
FIGURE 1 | Axial diffusivity (AD) map showing significantly higher AD values in the CSHNL group as compared to the control group. The red–yellow highlighted areas indicate higher AD values in the right corticospinal tract, left anterior thalamic radiation, and left corpus callosum. The background image consists of the standard MNI T1-weighted template at 1-mm thickness and the FA skeleton (*green*). The images at the bottom row from the left side to the right side indicate sagittal, coronal, and axial images for one representative axial level.

No significant differences between the CSNHL group and the control group were noted for gender ($p = 0.60$). Regarding age, the CSNHL group was significantly younger than the control group ($p = 0.0006$).

Whole-Brain Voxel-Wise Tract-Based Spatial Statistics Data

The whole-brain TBSS analysis showed no significant differences between the CSNHL and control groups for either the FA or the MD map (p-corrected > 0.05). Significantly higher AD values were observed in the left anterior thalamic radiation, right corticospinal tract, and left corpus callosum in the CSNHL group, as compared to the control group (p-corrected < 0.05) (**Table 2**). The coordinates of the local maxima and cluster sizes are presented in **Figure 1** and **Table 2**. There were significantly higher RD values for the left anterior thalamic radiation and

left IFOF (p-corrected < 0.05 for both threshold-free cluster correction and cluster-based thresholding correction) in the CSNHL group, relative to the control group, as shown in **Figure 2** and **Table 2**.

Region of Interest Data Focusing on Auditory Cortex and Correlation Between Diffusion Tensor Imaging and Age

There were statistically significant lower FA values in the CSNHL group compared to that in the control group in the HG ($p = 0.0015$) and FM ($p = 0.0033$) as presented in **Table 3** and **Figure 4**.

There was statistically significant negative correlation between the mean FA values in the HG relative to the FM (as an internal control) and age in the CSNHL group < 7 years ($n = 14$, $r = -0.59$, $p = 0.004$), while no similar correlation was noted
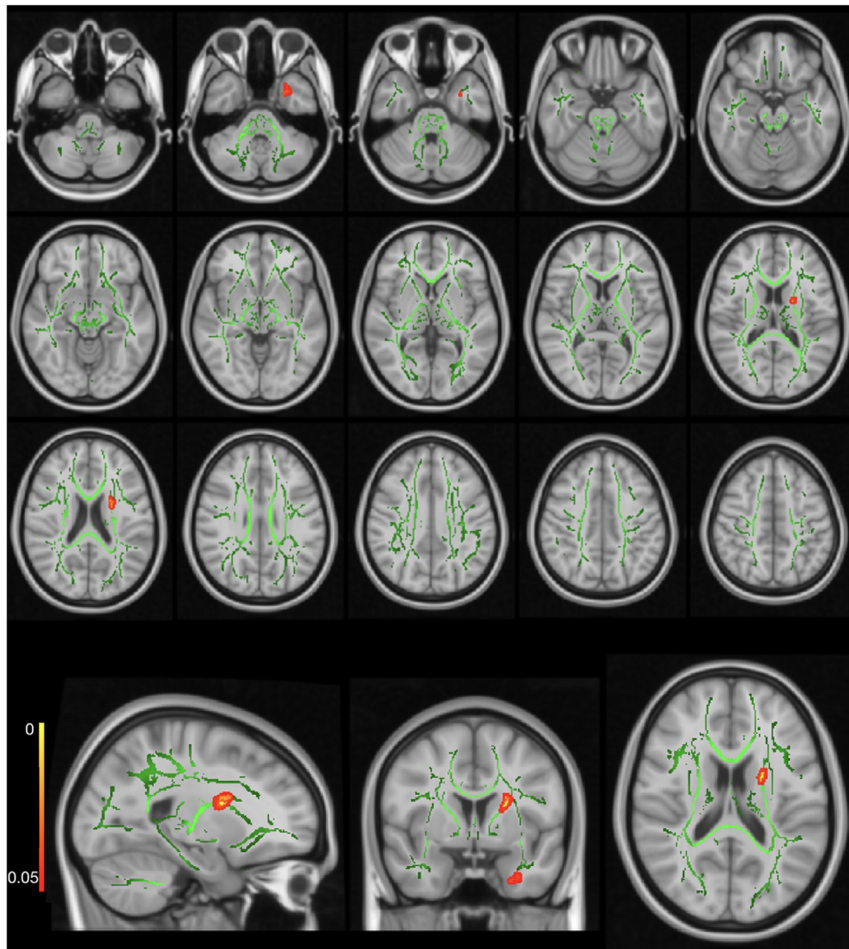
FIGURE 2 | Radial diffusivity (RD) map showing significantly higher RD values in the CSHNL group as compared to the control group. The red–yellow highlighted areas indicate higher RD values in left inferior fronto-occipital fasciculus (IFOF) and left anterior thalamic radiation. The background image consists of the standard MNI T1-weighted template at 1-mm thickness and the fractional anisotropy (FA) skeleton (*green*). The images at the bottom row from the left side to the right side indicate sagittal, coronal, and axial images for one representative axial level.

Heschl's gyrus (HG)   Forceps Major (FM)
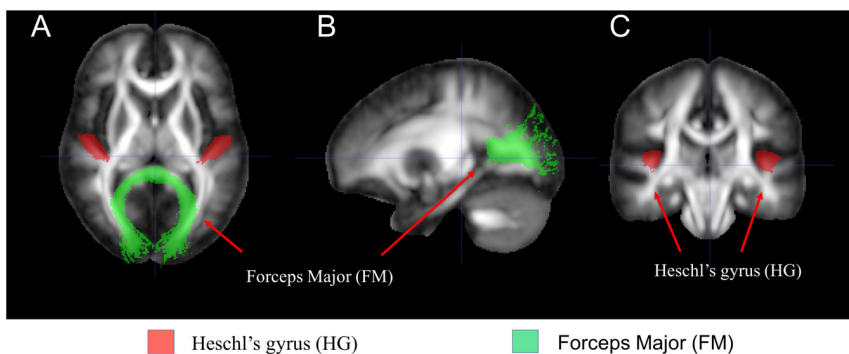
FIGURE 3 | The white matter (WM) tracts of interest selected for region of interest (ROI) analysis. **(A)** Axial, **(B)** sagittal, and **(C)** coronal images of the selected brain regions.

in the CSNHL group ≥ 7 years ($n = 9$, $r = 0.20$, $p = 0.60$). While in the control group stratified by the age of 7 years, there were no statistically significant correlation with age either in the control group < 7 years ($n = 3$, $r = 0.99$, $p = 0.10$) or in the control group ≥ 7 years ($n = 15$, $r = 0.45$, $p = 0.09$) as presented in **Figure 5**.

	CSNHL Group (*n* = 23)	Control Group (*n* = 18)
Age DTI obtained (years)	7.21 ± 2.67 (4.1–13.5)	10.86 ± 3.56 (4.5–15.3)
Age CSNHL diagnosed (years)	0.77 ± 0.18 (0.5–1.17)	\
Gender	11 Male/12 Female	11 Male/7 Female
Handedness	all righthanded	All righthanded
dB hearing loss: Left ear	108.45 ± 7.59 (94–120dB)	\
dB hearing loss: Right ear	106.62 ± 8.06 (91–120dB)	\

Age (years) are presented as mean age ± SD; hearing loss in left ear and right ear are presented as mean dB ± SD. dB: decibels.

TABLE 2 | Summary of DTI parameters showing significant differences between the CSHNL group and the Control Group.

	Voxels	P value	MNI Coordinates			Atlases (Tract from JHU)
			X	Y	Z	
AD	225	< 0.001	102	144	93	Corticospinal tract R
	133	0.007	108	123	84	Anterior thalamic radiation L
	105	0.022	75	131	122	Corpus callosum L
RD	103	0.016	115	134	31	IFOF L
	81	0.037	109	126	86	Anterior thalamic radiation L

JHU, Johns Hopkins University; L, Left; R, Right.

DISCUSSION

In this study, we found higher AD and RD values in the children with CSNHL, compared to the healthy children with normal hearing. These differences were present in several WM tracts, including the left anterior thalamic radiation, right corticospinal tract, corpus callosum, and left IFOF. The higher AD values indicated the presence of WM microstructural alterations in axonal formation, and the higher RD values indicated altered myelination in the children with CSNHL. To the best of our knowledge, our study was the first to report differences in the AD and RD values in children with CSNHL between 4.1 and 13.5 years of age. Moreover, our ROI analysis focusing on the auditory cortex showed lower FA values in the HG and negative correlation between the FA values and age up to 7 years in the children born deaf.

Tract-based spatial statistics is a commonly used program for comparing groups on a voxel-by-voxel basis to identify differences in FA and other DTI parameters (Maller et al., 2014). Using TBSS, we performed a whole-brain analysis of normalized brain to evaluate the WM microstructure. Our TBSS approach allowed for a broad characterization of between-group differences with stringent statistical corrections (Smith and Nichols, 2009). FA and MD are the two common diffusion indicators for DTI data, and they have been used to evaluate WM properties in various diseases, such as prelingual deafness and traumatic brain injury (Le Bihan et al., 2001; Hribar et al., 2014; Park et al., 2018). FA values are related to the WM microstructural integrity and directionality of axonal fibers within a voxel (Assaf and Pasternak, 2008). MD is the mathematical combination of

TABLE 3 | DTI parameters for the region of interest (ROI) analysis focusing on the Heschl's gyrus (HG) while using the forceps major (FM) as an internal control.

		CSNHL Group	Control Group	p
		n = 23	*n* = 18	
		Mean (SD)	Mean (SD)	
FA	HG*	0.24 (0.01)	0.27 (0.03)	<0.01
	FM**	0.27 (0.02)	0.29 (0.02)	<0.01
RD	HG	0.87 (0.02)	0.85 (0.03)	0.08
	FM	0.98 (0.03)	0.99 (0.04)	0.66
AD	HG	0.98 (0.03)	0.97 (0.03)	0.15
	FM	1.13 (0.04)	1.15 (0.04)	0.32

*p = 0.0015: when comparing the mean FA value of HG between the CSNHL group and the Control Group. **p = 0.0033: when comparing the mean FA value in the FM between the CSNHL group and the Control Group. Values indicate the mean values in each parameter. FA: fractional anisotropy, RD: radial diffusivity, AD: axial diffusivity.
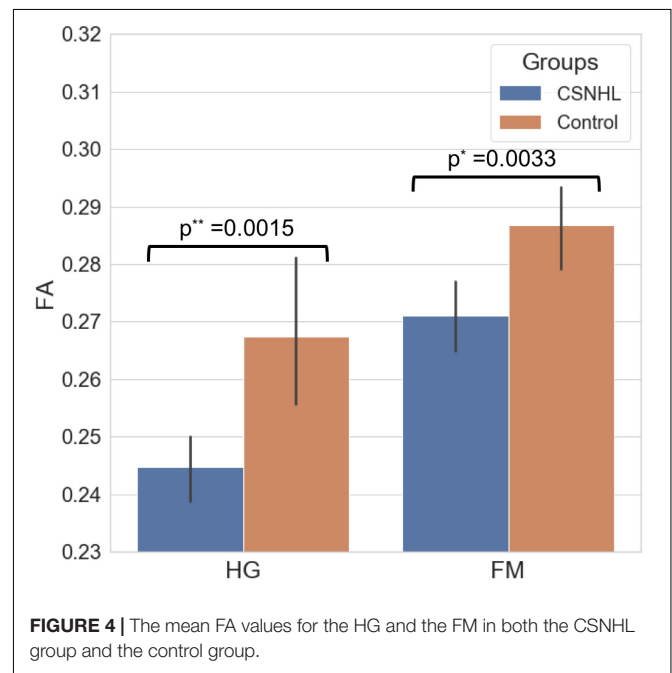
FIGURE 4 | The mean FA values for the HG and the FM in both the CSNHL group and the control group.

RD and AD. Nevertheless, assessing the diffusion properties such as AD and RD in regions parallel to and perpendicular to WM tracts may provide additional information about brain microstructure such as axonal formation and the presence of myelin (Song et al., 2002, 2005; Counsell et al., 2006; Sun et al., 2006), reflecting distinct aspects of WM microstructure. Directional diffusivity shows great potential to be specific biomarkers for injury to the axons and myelin (Xu et al., 2008), and DTI makes it possible to evaluate microstructural diffusivity (Hribar et al., 2014). In our study, we evaluated WM microstructural properties by analyzing RD and AD along with FA and MD. Although we found no significant differences in FA or MD values between the CSNHL group and the control group in the whole-brain TBSS analysis, our finding of higher RD and AD values in specific WM regions indicated the presence

FIGURE 5 | Correlation analysis between the mean FA values and age. The mean FA values on the Y-axis indicate the differences between the Heschl's gyrus (HG) and the forceps major (FM). **(A)** Correlation between the mean FA values and age in the CSNHL group < 7 years of age. **(B)** Correlation between the mean FA values and age in the CSNHL group ≥ 7 years of age.
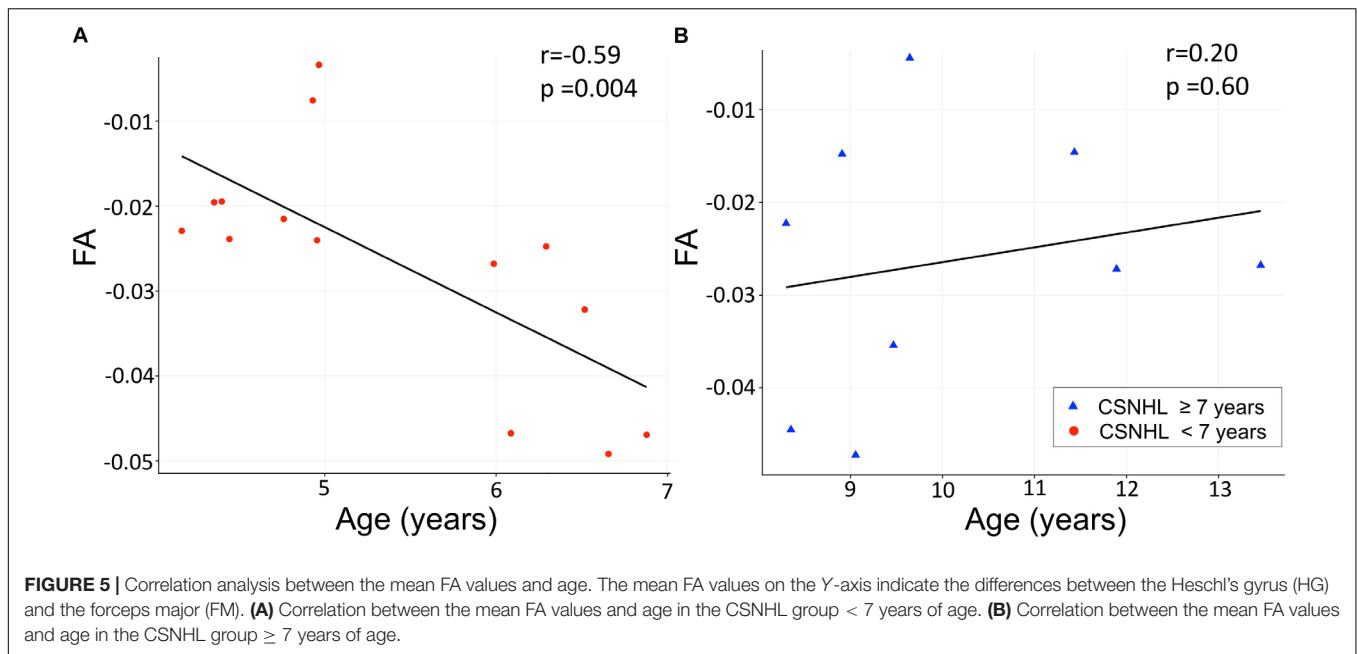
of WM microstructural differences in the children with CSNHL compared to the children with normal hearing.

The published literature is conflicting regarding the WM microstructural properties in the congenital deaf people. Various inconsistent structural differences in both auditory and non-auditory areas have been reported in the literature. This might be due to different analytical approaches with some studies using a whole-brain voxel-wise TBSS approach while others focusing on the specific regions of interests such as the auditory cortex. In addition, some studies adopted a combined approach using both the whole-brain TBSS and ROI methods, and some used a volumetric method to analyze the volumes of the auditory cortex (Emmorey et al., 2003; Park et al., 2018). For instance, a human study using a volumetric approach showed less WM in the HG leading to larger gray matter–WM ratio in the congenitally deaf adults compared to the hearing adults (Emmorey et al., 2003). Their findings suggest that lack of auditory stimuli from birth may lead to less myelination and possibly fewer fibers connected to the auditory cortices. A study with cortical cytoarchitectural analysis showed morphological changes in the auditory cortex with thinner auditory fields in a deaf animal research (Berger et al., 2017).

Non-auditory effects of congenital deafness have also been reported including the visual cortex, somatosensory projections, and motor tracts (Hribar et al., 2014; Kral et al., 2019). We speculate the non-auditory changes found in our whole-brain TBSS voxel-wise analysis might be due to the following reasons. First, neuroplasticity may occur in the children with CSNHL. The developing brain may exert compensatory cross-modal plasticity with reorganization of other sensory modalities such as visual ability, thus demonstrating a phenomena described in the literature as a visual "take over" of auditory areas in deaf children (Kral et al., 2019). In addition, sensory and motor inputs are linked to each other during development, and motor

cortices may mediate the increased auditory-evoked activity to sounds in humans (Reznik et al., 2015). Therefore, it is not surprising to find motor tracts such as the corticospinal tracts showing higher AD and RD values in our cohort with CSNHL compared to the hearing children. Second, our relatively small sample size consisted of the children ranging from 4.1 to 13.5 years of age, which was different from some studies in the literature (Miao et al., 2013; Park et al., 2018). In addition, our study was limited by the fact that our two study groups were not matched by age with the CSNHL group being younger than the control group. All these factors may decrease the sensitivity of detecting subtle differences in the DTI measures. Third, heterogenous FA images are commonly seen in children's developing brains. Therefore, there is an increased risk for false-positive results, which have been noted in the published literature (Smith et al., 2006). Nevertheless, our additional analysis with the ROI approach focusing on the auditory cortex showed the HG was affected in the CSNHL group with lower FA values compared to the hearing children. As for our future study, we will perform volumetric analysis of the auditory cortex and resting-state functional connectivity to further understand the brain structural and functional differences between the children with CSHNL and the children with normal hearing.

Some of our study results were consistent with the published literature, while some were divergent from prior reports. We found no significant differences for FA or MD values between the CSNHL group and the control group in the whole-brain TBSS analysis. One prior study showed reduced FA values but unchanged MD values along the auditory pathways of patients with sensorineural hearing loss when compared to the controls. Another study, which used TBSS to study prelingually deaf children, found no significant differences between the FA maps of patients and controls, but positive correlations between FA values and age for the deaf children in almost all WM tracts.

Differences in age may partially explain the differences between our study results and those of the studies by Chang et al. (2004) and Park et al. (2018). The study by Park et al. (2018) included patients from 1 to 7 years old, while our study participants with CSNHL were between 4.1 and 13.5 years old. In their additional analysis focusing on children over 4 years old, they found no significant differences between the deaf and control children for the DTI parameters, which was generally in line with our finding. Our study enrolled the children with CSNHL over 3.5 years old for two reasons. First, prior studies have shown that the brain changes the fastest from 1 to 3 years of age (Hermoye et al., 2006; Tierney and Nelson, 2009). Studying the developing brain during this period is challenging because the brain is small but rapidly growing. The regional topology and myelination also change during this period. Furthermore, brain function might not match the anatomy precisely at this stage of brain development (Fan et al., 2011). In addition, although 3.5 years of age is an optimal time for language function and brain development (Sharma et al., 2002, 2015), it is challenging to enroll study participants at this age. Some families did not seek medical treatment such as cochlear implant until the children were older enough for kindergarten or elementary school. In our experience, this delay in treatment was more commonly seen in rural areas with limited social and financial support. We therefore designed our study targeting the children with CSNHL between 3.5 and 18 years because this age group was particularly vulnerable and was at a disadvantage for learning in school due to hearing impairment. Because of the children's ages in our study cohort, it was expected to see our study results different from others targeting different ages of children with hearing loss.

Our additional analysis of age relationships in the WM microstructure showed significant correlation between the mean FA values and the age in the deaf children younger than 7 years. These findings implicate the WM microstructural alterations occur in younger children but may be compensated in older children with CSNHL. WM maturation occurs mostly in the first 12 months of life and plateaus after 24 months of age (Muftuler et al., 2012). Our study participants were all older than 2 years, and therefore, most of the major developmental changes should have already occurred in our cohort. Nevertheless, we assumed that some minor WM alterations may still be ongoing in our cohort. This assumption was supported by prior reports indicating gradual increases of FA values in various WM tracts as the brain matures from childhood to adulthood (Muftuler et al., 2012). However, we did not detect the expected age effect in our study with the whole-brain TBSS analysis. We believe this study result might be due to the wide age range in our study cohort, the two groups not age matched, and the relatively small sample size. Nevertheless, our additional ROI analysis focusing on the auditory cortex showed an age effect in the CSNHL group with significant negative correlation between the mean FA values in the HG and age in the subgroup of deaf children younger than 7 years.

We found that the WM microstructural differences were located predominantly in the frontal brain, including the anterior thalamic radiation connecting the anterior thalamus with prefrontal areas, and IFOF connecting the occipital lobe to the temporal lobe and frontal lobe. Frontal brain is associated with cognitive function such as executive function, speech, and language development. There are positive correlations between cognitive function and refined fiber organization of WM tracts (Muftuler et al., 2012). In addition, studies have shown an association between trajectories of brain WM maturation and intellectual performance (Tamnes et al., 2010). Therefore, our study findings of WM differences in the frontal brain implicate potential association between WM microstructural properties and cognitive function in children with CSNHL. We hypothesize that the detrimental effects on the WM microstructures as reflected by lower FA in young deaf children may represent potential neural correlates of cognitive development in the children with CSNHL and may contribute to their developmental delay. We recognize that this hypothesis was based on our speculation since we did not collect data to assess developmental delay or cognitive development for the current study. Nevertheless, our study has clinical relevance indicating preferential frontal WM involvement in the vulnerable children with hearing loss, which may serve as preliminary data for hypothesis generating for future large-scale studies of CSNHL, cognitive functioning, and developmental delay.

Our study showed higher values for AD and RD in the left anterior thalamic radiation in the children with CSNHL compared to the hearing children. The acoustic radiation is a component of the anterior thalamic radiation, and it is critical for hearing. This fiber tract originates in the medial geniculate nucleus and terminates in the primary auditory cortex, i.e., the HG (Wakana et al., 2004). Our study finding of significantly higher values in both AD and RD indicates loss of WM microstructural integrity of the acoustic radiation, as reported in another study (Husain et al., 2011). The higher value in AD implicates axonal abnormalities, intrinsic neuronal degradation, and morphologic differences in neuronal fibers (Lin et al., 2008; Profant et al., 2014). The higher value in RD may reflect abnormal myelination such as demyelination or dysmyelination (Song et al., 2002, 2005; Hasan et al., 2008), or it could reflect a higher number of crossing fibers, as shown in a study of congenital deafness in adults (Karns et al., 2017). For the children with CSNHL in our study, we speculate that the higher AD values may be due to abnormal axonal maturity such as less axonal density and caliber, while the higher RD might be due to abnormal myelination such as loss of myelin integrity.

Our study showed higher AD and RD in the corticospinal tract, corpus callosum, and IFOF in children with CSHNL as compared to the hearing children. These brain structures regulate motor, language, and visual function (Vanderah, 2016; Wu et al., 2016). The corticospinal tract is a collection of axons that carry motor-related information from the cerebral cortex to the spinal cord (Vanderah, 2016). Prior studies have found that hearing-impaired children with sensory organizational deficits have poor balance and have presented with motor problems (Crowe and Horak, 1988; Hartman et al., 2011). Our study findings of higher AD in the corpus callosum of the children with CSNHL is in agreement with prior studies of deaf individuals (Kim et al., 2009). The corpus callosum is responsible for

transmitting neural messages between the right and left hemispheres and plays an important role in transmitting auditory, visual, and linguistic signals. Therefore, it is not surprising to see different diffusivity values in the corpus callosum in the children with CSNHL compared to the hearing children in our cohort, who may undergo brain reorganization from the lack of sound stimuli.

There are several limitations to our study. First, our study cohort of patients with CSNHL had a wide age range, from 4.1 to 13.5 years of age, and were not age matched with the control group, which may confound the findings in our study. Moreover, we did not evaluate the CSNHL patients younger than 4 years, which may limit the generalizability of our study results to all children with CSHNL. Second, our sample size was not large enough to stratify the patients based on their clinical and lifestyle factors, such as knowledge of sign language and the extent of special education for the hearing impaired, which may alter brain structure and function. Third, our TBSS method involved whole-brain analysis with stringent statistical comparisons, which may limit the power to detect subtle differences in specific brain regions between the two groups. Other methods, such as the ROI approach, may help us to evaluate differences in the DTI parameters in the particular regions of WM that are relevant to the auditory pathway (Karns et al., 2017). Therefore, we performed additional ROI analysis focusing on the auditory cortex. Fourth, although DTI is commonly used to study the WM microstructure, it has its own inherent limitations as an imaging method. For instance, the tensor model only represents one major fiber direction in a voxel, so DTI analysis could potentially be confused by cross-fiber regions (Behrens et al., 2007). Last, our study was limited due to the lack of analysis on the sedation effect. We used chloral hydrate to induce a light-sedated state because it was easy to use as a liquid for oral intake for children, and it had a low adverse respiratory effect. However, all children in our study were sedated during the MRI scans. We therefore could not perform an analysis of sedation effect by comparing the group with sedation with the group without sedation. Nevertheless, we recognize that sedation may affect brain structure and function. A brain connectivity analysis of resting-state functional MRI study examined the effect of chloral hydrate-induced light sedation in school-aged children (Wei et al., 2013). Their study showed a global detrimental effect of sedation on the functional interactions of the brain, especially on the information-processing properties. We will incorporate the sedation effect in our future studies of brain structure and function in children with CSNHL.

In summary, we assessed the cerebral WM microstructural properties in children with CSNHL. We found higher AD and RD values in the left anterior thalamic radiation, right corticospinal tract, corpus callosum, and left IFOF in the children with CSNHL compared to the hearing children. We also found lower FA values in the auditory cortex in the deaf children compared to the hearing children. Furthermore, we identified the FA values being negatively correlated with age up to 7 years in the children born deaf. We speculate that the WM microstructural alterations in the children with CSNHL could be due to injury of the auditory pathway or brain functional neuroplasticity with reorganization. This information may help to design and track hearing rehabilitation in the children with CSNHL.

DATA AVAILABILITY

The raw data supporting the conclusion of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This study was carried out in accordance with all institutional, local, and national guidelines. Our study was approved by the institutional review board of the First Affiliated Hospital of Guangxi Medical University. The parents or legal guardians of the children gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

MJ and ZW contributed equally to this work. MJ and LL designed the study. MJ, BC, CW, NY, and CZ contributed to manuscript preparation and revision. ZW performed data acquisition and analysis. MJ, ZW, CW, NY, and BC performed data interpretation. LL, BC, and CZ were responsible for the study supervision. All authors approved the final manuscript.

ACKNOWLEDGMENTS

REFERENCES

Andersson, J. L., Jenkinson, M., and Smith, S. (2007). *Non-Linear Registration, Aka Spatial Normalisation*. FMRIB technical report TR07JA2. Oxford: FMRIB Centre.

Andersson, J. L. R., and Sotiropoulos, S. N. (2016). An integrated approach to correction for off-resonance effects and subject movement in diffusion MR imaging. *Neuroimage* 125, 1063–1078. doi: 10.1016/j.neuroimage.2015.10.019

Assaf, Y., and Pasternak, O. (2008). Diffusion tensor imaging (DTI)-based white matter mapping in brain research: a review. *J. Mol. Neurosci.* 34, 51–61. doi: 10.1007/s12031-007-0029-0

Basser, P. J., Mattiello, J., and LeBihan, D. (1994). Estimation of the effective self-diffusion tensor from the NMR spin echo. *J. Magn. Reson. B* 103, 247–254.

Behrens, T. E., Berg, H. J., Jbabdi, S., Rushworth, M. F., and Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34, 144–155. doi: 10.1016/j.neuroimage.2006.09.018

Berger, C., Kuhne, D., Scheper, V., and Kral, A. (2017). Congenital deafness affects deep layers in primary and secondary auditory cortex. *J. Comp. Neurol.* 525, 3110–3125. doi: 10.1002/cne.24267

Chang, Y., Lee, S. H., Lee, Y. J., Hwang, M. J., Bae, S. J., Kim, M. N., et al. (2004). Auditory neural pathway evaluation on sensorineural hearing loss using diffusion tensor imaging. *Neuroreport* 15, 1699–1703.

Chari, D. A., and Chan, D. K. (2017). Diagnosis and treatment of congenital sensorineural hearing loss. *Curr. Otorhinolaryngol. Rep.* 5, 251–258. doi: 10.1007/s40136-017-0163-3

Counsell, S. J., Shen, Y., Boardman, J. P., Larkman, D. J., Kapellou, O., Ward, P., et al. (2006). Axial and radial diffusivity in preterm infants who have diffuse white matter changes on magnetic resonance imaging at term-equivalent age. *Pediatrics* 117, 376–386. doi: 10.1542/peds.2005-0820

Crowe, T. K., and Horak, F. B. (1988). Motor proficiency associated with vestibular deficits in children with hearing impairments. *Phys. Ther.* 68, 1493–1499.

Douaud, G., Jbabdi, S., Behrens, T. E., Menke, R. A., Gass, A., Monsch, A. U., et al. (2011). DTI measures in crossing-fibre areas: increased diffusion anisotropy reveals early white matter alteration in MCI and mild Alzheimer's disease. *Neuroimage* 55, 880–890.

Emmorey, K., Allen, J. S., Bruss, J., Schenker, N., and Damasio, H. (2003). A morphometric analysis of auditory brain regions in congenitally deaf adults. *Proc. Natl. Acad. Sci. U.S.A.* 100, 10049–10054. doi: 10.1073/pnas.1730169100

Fan, Y., Shi, F., Smith, J. K., Lin, W., Gilmore, J. H., and Shen, D. (2011). Brain anatomical networks in early human brain development. *Neuroimage* 54, 1862–1871. doi: 10.1016/j.neuroimage.2010.07.025

Fonov, V., Evans, A. C., Botteron, K., Almli, C. R., McKinstry, R. C., Collins, D. L., et al. (2011). Unbiased average age-appropriate atlases for pediatric studies. *Neuroimage* 54, 313–327. doi: 10.1016/j.neuroimage.2010.07.033

Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C., and Collins, D. (2009). Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* 47:S102.

Hartman, E., Houwen, S., and Visscher, C. (2011). Motor skill performance and sports participation in deaf elementary school children. *Adapt. Phys. Activ. Q.* 28, 132–145.

Hasan, K. M., Eluvathingal, T. J., Kramer, L. A., Ewing-Cobbs, L., Dennis, M., and Fletcher, J. M. (2008). White matter microstructural abnormalities in children with spina bifida myelomeningocele and hydrocephalus: a diffusion tensor tractography study of the association pathways. *J. Magn. Reson. Imaging* 27, 700–709. doi: 10.1002/jmri.21297

Hermoye, L., Saint-Martin, C., Cosnard, G., Lee, S. K., Kim, J., Nassogne, M. C., et al. (2006). Pediatric diffusion tensor imaging: normal database and observation of the white matter maturation in early childhood. *Neuroimage* 29, 493–504. doi: 10.1016/j.neuroimage.2005.08.017

Hribar, M., Suput, D., Carvalho, A. A., Battelino, S., and Vovk, A. (2014). Structural alterations of brain grey and white matter in early deaf adults. *Hear Res.* 318, 1–10. doi: 10.1016/j.heares.2014.09.008

Husain, F. T., Medina, R. E., Davis, C. W., Szymko-Bennett, Y., Simonyan, K., Pajor, N. M., et al. (2011). Neuroanatomical changes due to hearing loss and chronic tinnitus: a combined VBM and DTI study. *Brain Res.* 1369, 74–88. doi: 10.1016/j.brainres.2010.10.095

Karns, C. M., Stevens, C., Dow, M. W., Schorr, E. M., and Neville, H. J. (2017). Atypical white-matter microstructure in congenitally deaf adults: a region of interest and tractography study using diffusion-tensor imaging. *Hear Res.* 343, 72–82. doi: 10.1016/j.heares.2016.07.008

Kim, D.-J., Park, S.-Y., Kim, J., Lee, D. H., and Park, H.-J. (2009). Alterations of white matter diffusion anisotropy in early deafness. *Neuroreport* 20, 1032–1036. doi: 10.1097/WNR.0b013e32832e0cdd

Kral, A., Dorman, M. F., and Wilson, B. S. (2019). Neuronal development of hearing and language: cochlear implants and critical periods. *Annu. Rev. Neurosci.* doi: 10.1146/annurev-neuro-080317-061513 [Epub ahead of print]

Kral, A., and O'donoghue, G. M. (2010). Profound deafness in childhood. *N. Engl. J. Med.* 363, 1438–1450.

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Le Bihan, D., Mangin, J. F., Poupon, C., Clark, C. A., Pappata, S., Molko, N., et al. (2001). Diffusion tensor imaging: concepts and applications. *J. Magn. Reson. Imaging* 13, 534–546.

Lin, Y., Wang, J., Wu, C., Wai, Y., Yu, J., and Ng, S. (2008). Diffusion tensor imaging of the auditory pathway in sensorineural hearing loss: changes in radial diffusivity and diffusion anisotropy. *J. Magn. Reson. Imaging* 28, 598–603. doi: 10.1002/jmri.21464

Maller, J. J., Thomson, R. H., Pannek, K., Rose, S. E., Bailey, N., Lewis, P. M., et al. (2014). The (Eigen)value of diffusion tensor imaging to investigate depression after traumatic brain injury. *Hum. Brain. Mapp.* 35, 227–237. doi: 10.1002/hbm.22171

Miao, W., Li, J., Tang, M., Xian, J., Li, W., Liu, Z., et al. (2013). Altered white matter integrity in adolescents with prelingual deafness: a high-resolution tract-based spatial statistics imaging study. *AJNR Am. J. Neuroradiol.* 34, 1264–1270. doi: 10.3174/ajnr.A3370

Mohr, P. E., Feldman, J. J., Dunbar, J. L., McConkey-Robbins, A., Niparko, J. K., Rittenhouse, R. K., et al. (2000). The societal costs of severe to profound hearing loss in the United States. *Int. J. Technol. Assess. Health Care* 16, 1120–1135.

Muftuler, L. T., Davis, E. P., Buss, C., Solodkin, A., Su, M. Y., Head, K. M., et al. (2012). Development of white matter pathways in typically developing preadolescent children. *Brain Res.* 1466, 33–43. doi: 10.1016/j.brainres.2012.05.035

Paludetti, G., Conti, G., Di Nardo, W., De Corso, E., Rolesi, R., Picciotti, P., et al. (2012). Infant hearing loss: from diagnosis to therapy official report of XXI conference of italian society of pediatric otorhinolaryngology. *Acta. Otorhinolaryngol. Ital.* 32:347.

Park, K. H., Chung, W.-H., Kwon, H., and Lee, J.-M. (2018). Evaluation of cerebral white matter in prelingually deaf children using diffusion tensor imaging. *BioMed. Res. Int.* 2018:6795397. doi: 10.1155/2018/6795397

Profant, O., Skoch, A., Balogova, Z., Tintera, J., Hlinka, J., and Syka, J. (2014). Diffusion tensor imaging and MR morphometry of the central auditory pathway and auditory cortex in aging. *Neuroscience* 260, 87–97. doi: 10.1016/j.neuroscience.2013.12.010

Reznik, D., Ossmy, O., and Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *J. Neurosci.* 35, 2173–2180. doi: 10.1523/jneurosci.3723-14.2015

Sharma, A., Campbell, J., and Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. *Int. J. Psychophysiol.* 95, 135–144. doi: 10.1016/j.ijpsycho.2014.04.007

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear* 23, 532–539.

Smith, S. M. (2002). Fast robust automated brain extraction. *Hum. Brain. Mapp.* 17, 143–155. doi: 10.1002/hbm.10062

Smith, S. M., Jenkinson, M., Johansen-Berg, H., Rueckert, D., Nichols, T. E., Mackay, C. E., et al. (2006). Tract-based spatial statistics: voxelwise analysis of multi-subject diffusion data. *Neuroimage* 31, 1487–1505. doi: 10.1016/j.neuroimage.2006.02.024

Smith, S. M., and Nichols, T. E. (2009). Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference. *Neuroimage* 44, 83–98. doi: 10.1016/j.neuroimage.2008.03.061

Song, S.-K., Sun, S.-W., Ramsbottom, M. J., Chang, C., Russell, J., and Cross, A. H. (2002). Dysmyelination revealed through MRI as increased radial (but unchanged axial) diffusion of water. *Neuroimage* 17, 1429–1436.

Song, S.-K., Yoshino, J., Le, T. Q., Lin, S.-J., Sun, S.-W., Cross, A. H., et al. (2005). Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage* 26, 132–140.

Sun, S. W., Liang, H. F., Trinkaus, K., Cross, A. H., Armstrong, R. C., and Song, S. K. (2006). Noninvasive detection of cuprizone induced axonal damage and demyelination in the mouse corpus callosum. *Magn. Reson. Med.* 55, 302–308. doi: 10.1002/mrm.20774

Tamnes, C. K., Ostby, Y., Fjell, A. M., Westlye, L. T., Due-Tonnessen, P., and Walhovd, K. B. (2010). Brain maturation in adolescence and young adulthood: regional age-related changes in cortical thickness and white matter volume and microstructure. *Cereb. Cortex* 20, 534–548. doi: 10.1093/cercor/bhp118

Tierney, A. L., and Nelson, C. A. III (2009). Brain development and the role of experience in the early years. *Zero Three* 30, 9–13.

Vanderah, T. (2016). *Nolte's the Human Brain*. Amsterdam: Elsevier.

Wakana, S., Jiang, H., Nagae-Poetscher, L. M., van Zijl, P. C., and Mori, S. (2004). Fiber tract-based atlas of human white matter anatomy. *Radiology* 230, 77–87. doi: 10.1148/radiol.2301021640

Wang, Z., Childress, A. R., Wang, J., and Detre, J. A. (2007). Support vector machine learning-based fMRI data group analysis. *Neuroimage* 36, 1139–1151. doi: 10.1016/j.neuroimage.2007.03.072

Wei, Z., Alcauter, S., Jin, K., Peng, Z. W., and Gao, W. (2013). Graph theoretical analysis of sedation's effect on whole brain functional system in school-aged children. *Brain Connect.* 3, 177–189. doi: 10.1089/brain.2012.0125

Wheeler-Kingshott, C. A., and Cercignani, M. (2009). About "axial" and "radial" diffusivities. *Magn. Reson. Med.* 61, 1255–1260. doi: 10.1002/mrm.21965

Wu, Y., Sun, D., Wang, Y., and Wang, Y. (2016). Subcomponents and connectivity of the inferior fronto-occipital fasciculus revealed by diffusion spectrum imaging fiber tracking. *Front. Neuroanat.* 10:88. doi: 10.3389/fnana.2016.00088

Xu, J., Sun, S. W., Naismith, R. T., Snyder, A. Z., Cross, A. H., and Song, S. K. (2008). Assessing optic nerve pathology with diffusion MRI: from mouse to human. *NMR Biomed.* 21, 928–940. doi: 10.1002/nbm.1307

Somatosensory Cross-Modal Reorganization in Children With Cochlear Implants

Garrett Cardon[1] and Anu Sharma[2]*

[1] Department of Psychology, Colorado State University, Fort Collins, CO, United States, [2] Department of Speech, Language, and Hearing Sciences, University of Colorado Boulder, Boulder, CO, United States

Deprived of sensory input, as in deafness, the brain tends to reorganize. Cross-modal reorganization occurs when cortices associated with deficient sensory modalities are recruited by other, intact senses for processing of the latter's sensory input. Studies have shown that this type of reorganization may affect outcomes when sensory stimulation is later introduced via intervention devices. One such device is the cochlear implant (CI). Hundreds of thousands of CIs have been fitted on people with hearing impairment worldwide, many of them children. Factors such as age of implantation have proven useful in predicting speech perception outcome with these devices in children. However, a portion of the variance in speech understanding ability remains unexplained. It is possible that the degree of cross-modal reorganization may explain additional variability in listening outcomes. Thus, the current study aimed to examine possible somatosensory cross-modal reorganization of the auditory cortices. To this end we used high density EEG to record cortical responses to vibrotactile stimuli in children with normal hearing (NH) and those with CIs. We first investigated cortical somatosensory evoked potentials (CSEP) in NH children, in order to establish normal patterns of CSEP waveform morphology and sources of cortical activity. We then compared CSEP waveforms and estimations of cortical sources between NH children and those with CIs to assess the degree of somatosensory cross-modal reorganization. Results showed that NH children showed expected patterns of CSEP and current density reconstructions, such that postcentral cortices were activated contralaterally to the side of stimulation. Participants with CIs also showed this pattern of activity. However, in addition, they showed activation of auditory cortical areas in response to somatosensory stimulation. Additionally, certain CSEP waveform components were significantly earlier in the CI group than the children with NH. These results are taken as evidence of cross-modal reorganization by the somatosensory modality in children with CIs. Speech perception in noise scores were negatively associated with CSEP waveform components latencies in the CI group, suggesting that the degree of cross-modal reorganization is related to speech perception outcomes. These findings may have implications for clinical rehabilitation in children with cochlear implants.

Keywords: cochlear implants, cross-modal reorganization, somatosensory, vibrotactile, cortical somatosensory evoked potential, high density EEG, independent components analysis, sLORETA

INTRODUCTION

Permanent hearing loss in children is a common condition that is found in 2–3 of every 1,000 live births (Centers for Disease Control and Prevention [CDC], 2010). Children who are identified with more severe cases of hearing loss (i.e., ≥70 dBHL), are often candidates for treatment with a cochlear implant (CI). CIs are devices that restore hearing to deaf individuals via direct electrical stimulation of the auditory (VIII) nerve. As of 2013, approximately 324,200 CIs had been fitted worldwide (Food and Drug Administration [FDA], 2012). These devices have proven extremely useful in restoring auditory function to many children born with hearing loss. However, many other implant recipients have had relatively little success in behavioral speech understanding (Harrison et al., 2005; Nicholas and Geers, 2007; Holt and Svirsky, 2008). Despite ongoing improvement in CIs and implantation procedures, there remains a high degree of variability in the behavioral outcomes (e.g., speech and language development) of children with CIs (Svirsky et al., 2000; Sarant et al., 2001, 2014; Tobey et al., 2003; Harrison et al., 2005; Geers, 2006; Nicholas and Geers, 2007; Holt and Svirsky, 2008; Geers et al., 2009; Lund, 2016; Szagun and Schramm, 2016). Given this variability, it is difficult to predict the level of benefit an implant will provide a given patient. Recent investigation has been aimed at discovering the underlying factors associated with this variability (Svirsky et al., 2000; Sarant et al., 2001; Tobey et al., 2003; Geers, 2006; Geers et al., 2009; Szagun and Schramm, 2016). However, despite these efforts, only these factors explain on a portion of the variability (i.e., approximately 35–62%; Fink et al., 2007), though one that seems to stand out is age at implantation – earlier implantation appears to lead to greater chances for favorable outcome (Sharma et al., 2002a,b; Svirsky et al., 2004; Geers, 2006; Geers et al., 2009; Niparko et al., 2010).

Sensory loss (i.e., blindness and deafness) can lead to reorganization of the cerebral cortex. In deafness, this reorganization manifests itself when sensory modalities that are intact recruit auditory cortices for their own processing – termed *cross-modal reorganization* (see Bavelier and Neville, 2002; Merabet and Pascual-Leone, 2010 for reviews). In conjunction with age of implantation, it is likely that cortical development and neuroplastic processes, such as cross-modal reorganization, play a role in outcomes for children with CIs. Though most of the work that characterizes this type of plastic change has examined cross-modal reorganization of auditory cortex by vision (e.g., Rebillard et al., 1977; Neville et al., 1983; Finney et al., 2001, 2003; Fine et al., 2005; Sadato, 2005; Doucet et al., 2006; Bavelier and Hirshorn, 2010; Lomber et al., 2010; Meredith and Lomber, 2011; Campbell and Sharma, 2014, 2016; Clemo et al., 2014; Kok et al., 2014; Scott et al., 2014), a number of studies have also shown evidence of cross-modal reorganization between the somatosensory and auditory cortices in both animals and humans (Levänen et al., 1998; Baldwin, 2002; Auer et al., 2007; Sharma et al., 2007; Allman et al., 2009; Meredith and Lomber, 2011; Meredith and Allman, 2012; Karns et al., 2012). However, while such investigations have been carried out in adults, no study has examined somatosensory cross-modal reorganization of auditory cortical areas in pediatric CI recipients. Thus,

the goal of this study was to examine possible cross-modal reorganization between the somatosensory and auditory systems in children with CIs (relative to age-matched NH controls) and its relationship to behavioral speech perception.

MATERIALS AND METHODS

Participants

Participants for the current study consisted of two groups of individuals: those with NH and children with CI. Children with CIs were limited to those with sensorineural hearing loss (SNHL). All participants were recruited and tested in accordance with the Institutional Review Board of the University of Colorado at Boulder. As such, signed informed consent was obtained from parents or guardians of all subjects in the current study. We recruited 35 NH children between the ages of 5 and 17 years of age (17 female). The overall group was divided into three age groups for recruiting and analysis. These groups were: (1) 5–7-year-old children ($n = 9$; mean age = 6.95 years; SD = ±0.53 years); (2) 8–10-year-old children ($n = 11$; mean age = 9.81 years; SD = ±0.97 years); (3) 11-year-old and older children and adolescents ($n = 15$; mean age = 12.9 years; ±SD = 1.45 years). All of these individuals had NH, which was defined as auditory thresholds at or below 20 dBHL at 250, 500, 1,000, 2,000, 4,000, and 8,000 Hz. These thresholds were obtained in each participant by a certified clinical audiologist. Additionally, none of the participants had any history of neurological disorder.

We recruited children with CIs (CI group; $n = 12$; mean age at test = 12.42 years; S.D. = ±4.16 years). A subset of the above NH group was formed for comparison with CI children (NH group; $n = 17$; mean age at test = 12.29 years; S.D. = ±2.46 years). These 17 children were selected, rather than 12, to increase statistical power, while still maintaining similarity in age between the CI and NH groups. Statistical comparison of the ages of the NH and CI groups confirmed that they were not significantly different ($p > 0.05$). Ten out of 12 CI participants were bilaterally implanted, while the remaining two subjects had unilateral CIs. All bilateral CI recipients were implanted sequentially – seven of ten received their first implant in the right ear. The mean age of first implantation for the CI group as a whole was 3.90 years (S.D. = ±4.03 years), while the average age of second implantation for bilaterally implanted children was 7.33 years (S.D. = ±4.47 years). The average duration of first CI use at the time of testing (i.e., time between first CI fitting and testing) was 8.51 years (S.D. = ±4.09 years), while the duration of 2nd CI use was 5.65 years (S.D. = ±2.31 years). Make and model of CI and speech processing strategy was not accounted for given the limited sample size of the CI group.

Stimuli

250 Hz tones, each 90 ms in duration, with 10 ms linear ramps at onset and offset, were used to elicit cortical somatosensory evoked potentials (CSEP). These stimuli were presented to each participant via a standard clinical bone oscillator (Sensory Systems d.b.a. Radioear Inc., B71 Bone Transducer), which was electrically shielded with copper mesh so that any electrical

noise produced by the device would not be registered by the EEG electrodes. During testing, this transducer was temporarily affixed to the participant's right or left index finger using medical tape. For consistency, all participants underwent testing with right finger stimulation. Additional testing with left finger stimulation was achieved in a subset of the study participants ($n = 6$), though, due to time constraints and subject cooperation, this was not done in all children. Stimulus presentation timing was controlled by E-Prime® 2.0 software (Psychology Software Tools, Inc.). All stimuli were presented at a level of 55 dBHL on the audiometer, which resulted in vibrotactile sensation in all participants (approximately 0.122 g or 1.2 m/s^2 of acceleration output) that was sufficient to elicit CSEPs, but never uncomfortable (Weinstein, 1968). For all CI participants, CIs were turned off during CSEP recording to ensure that the vibrotactile stimuli were only felt and not heard. Continuous white noise was played via a loudspeaker at a level of 50 dBHL on the side of stimulation in order to mask any auditory artifact of vibrotactile stimulation for all participants. Procedures were similar to those described previously in studies from our laboratory and others (Yamaguchi and Knight, 1991; Sharma et al., 2015, 2016; Cardon and Sharma, 2018). All participants reported that they could *feel*, but not *hear*, the stimulus.

EEG Recording and Analysis

During testing, each participant was seated in a comfortable chair situated in a sound treated room. They were fitted with a 128-channel EEG recording net (Electrical Geodesics, Inc.) that had been soaked in a solution of water, baby shampoo, and sodium chloride. EEG recordings were sampled at 1 kHz and band-pass filtered online between 0.1 and 200 Hz. Following recording, EEG data were initially high-pass filtered offline at 1 Hz. These data were then segmented into epochs that consisted of 100 ms pre- and 495 ms post-stimulus intervals. Then, data were exported for further analysis in the EEGLAB toolbox (Delorme and Makeig, 2004) running within the Matlab® software package (The MathWorks®2014). Once imported, channels containing excessive amounts of noise were rejected. Then, epochs that presented with data exceeding $\pm 100\ \mu V$ in amplitude were also eliminated. The sampling rate of the data was then changed to 250 Hz to allow for subsequent processing efficiency. Data then underwent re-referencing to a common average reference. Finally, rejected channels' data were replaced via spherical interpolation, which was necessary to appropriately address highly noisy channels and remove their effects on subsequent analyses.

The region of interest (ROI) employed for initial CSEP analysis in the large NH group consisted of 24 electrodes that covered the parietal and temporal areas of the left hemisphere of the scalp (Hämäläinen et al., 1990). Waveforms from the designated electrodes from this ROI were averaged together to form a composite waveform. Peak latencies and absolute and peak-to-peak amplitudes for the P50, N70, P100, N140a, and N140b CSEP waveform components were then extracted from waveforms from the ROI for each participant. These were later used for statistical comparison.

For CI children (and smaller age-matched group of NH children) electrodes were divided into more specific ROIs in the temporal and parietal regions of both hemispheres in order to evaluate possible effects of cross-modal reorganization on CSEP responses. ROI selection was based on a combination of visual inspection of the 128-channel data and optimal recording locations of CSEPs reported in Hämäläinen et al. (1990) and Cardon and Sharma (2018). ROIs included: (1) Left Temporal ROI (LTemp ROI; electrodes: TP7, T9, P9, TP9, T5-P7); (2) Left Parietal ROI (LPar ROI; electrodes: P3, P5, CP1, P1, PO7, PO3); (3) Right Parietal ROI (RPar ROI; electrodes: P4, P6, CP2, P2, PO8, PO4); (4) Right Temporal ROI (RTemp; electrodes: TP8, T10, P10, TP10, T6-P8). These electrode positions represent approximate 10–20 system electrode locations, as reported in Luu and Ferree (2000), since EGI uses a geodesic electrode organization system.

Average CSEPs were calculated for each participant for all ROIs. Then, each participant's ROI CSEP waveform component latencies and absolute and peak-to-peak amplitudes were noted (i.e., the difference between the amplitude of the CSEP peak of interest and the preceding peak were calculated). We measured peak-to-peak amplitudes due to the larger inherent variability in measurement of absolute amplitudes (e.g., Johnson et al., 1975). Statistical comparisons were then performed with these CSEP peak latency and amplitude values between the CI and NH groups using non-parametric Mann–Whitney U Tests for each ROI. CSEP latencies and amplitudes for the CI group were also correlated (Spearman's Rank correlations) with behavioral speech perception in noise scores (see "Speech Perception in Noise" section) to evaluate potential relationships between neural activity and behavioral speech perception in noise. Correction for multiple comparisons was performed on both between group statistics and correlations using the False Discovery Rate method presented by Benjamini and Hochberg (1995; $q \leq 0.1$).

Current Density Reconstruction

In preparation for current density reconstruction (CDR), each subject's data epochs were concatenated and subjected to independent components analysis (ICA). One application of ICA is artifact rejection. Thus, independent components (ICs) containing eye blinks or movement, electrical noise, or muscle artifact were removed from each participant's dataset. After ICA artifact rejection, ICs that accounted for the highest portion of the variance around each peak of the CSEP were saved for inclusion in CDR. Data were then transferred to the Curry® Scan 7 Neuroimaging suite (Compumedics Neuroscan™) for cortical source estimation. Initial processing steps toward CDR included baseline correction, noise estimation using the pre-stimulus interval, averaging of participants individual CSEP waveforms to for grand average waveforms, and additional ICA.

Modeling of the head was accomplished using the Boundary Element Method (BEM; e.g., Fuchs et al., 2002; Hallez et al., 2007). Within this head model, white matter volumes were adjusted to match age-related values (Wilke et al., 2006; Gilley et al., 2008). CDR were then performed for each CSEP waveform component (P50, N70, P100,

N140a, N140b) using the sLORETA algorithm (Pascual-Marqui, 2002; see Grech et al., 2008 for a review). The results of this method appear as color gradients that represent the *F*-distribution of the data, which were overlaid using the Montreal Neurologic Institute (MNI) average brain (Evans et al., 1993).

Speech Perception in Noise

Speech perception ability was assessed in each participant in the CI group using the BKB-SIN™ test (Bench et al., 1979; Etymotic Research, 2005). During this testing, participants sat facing a loudspeaker at 0° azimuth with his or her CI on and functioning as it normally would. Sentences – two lists of six sentences each – were then presented to the participant via the loudspeaker at 65 dBHL. As the sentences progressed, background noise (multi-talker babble) level was increased with each sentence. This noise increase occurred in five steps, each of 5 dB increments, or from 25 dB SNR (least challenging) to 0 dB SNR (most challenging). The participant was asked to repeat the words of the sentence he or she heard. The tester marked key words from each sentence as correct or incorrect and computed a score for each list, based on the number of words repeated correctly. Participants received an SNR score, representing the level at which they could perceive and repeat 50% of key words – lower scores indicated better performance. Scores from the two presented lists were then averaged together to obtain a composite BKB-SIN score for each participant. In addition, age corrections were applied to participants' composite scores to normalize results for comparison across subjects (Etymotic Research, 2005). Finally, BKB-SIN scores were correlated with CSEP component peak latencies from each ROI to assess the relationship between speech perception in noise and cross-modal reorganization.

RESULTS

Normal Hearing Children (*n* = 35)
Cortical Somatosensory Evoked Potentials

Plots of the grand average CSEP waveforms for each of the age groups (i.e., 5–7-, 8–1-, and 11–17-year-olds) from the temporo-parietal ROI are shown in **Figure 1**. Across all ages, all of the components of the CSEP (i.e., P50, N70, P100, N140) can be reliably identified. In the majority of subjects, regardless of their age, the N140 appeared as a bifid negative going peak. Given this pattern, we classified the first of the N140 peaks as the N140a, while the second was called the N140b. Thus, CSEP waveform morphology appears to be stable (with respect to presence of peak components) across the age range examined in this study.

In order to determine more detailed differences between the age groups' CSEP waveforms, both peak latency and peak-to-peak amplitude results from the aforementioned ROI were subjected to statistical comparison. One latency difference was found following multiple comparisons correction. That is, there was a main effect of age for the N140a CSEP latency ($p = 0.00$; $F = 8.05$). *Post hoc* analysis revealed
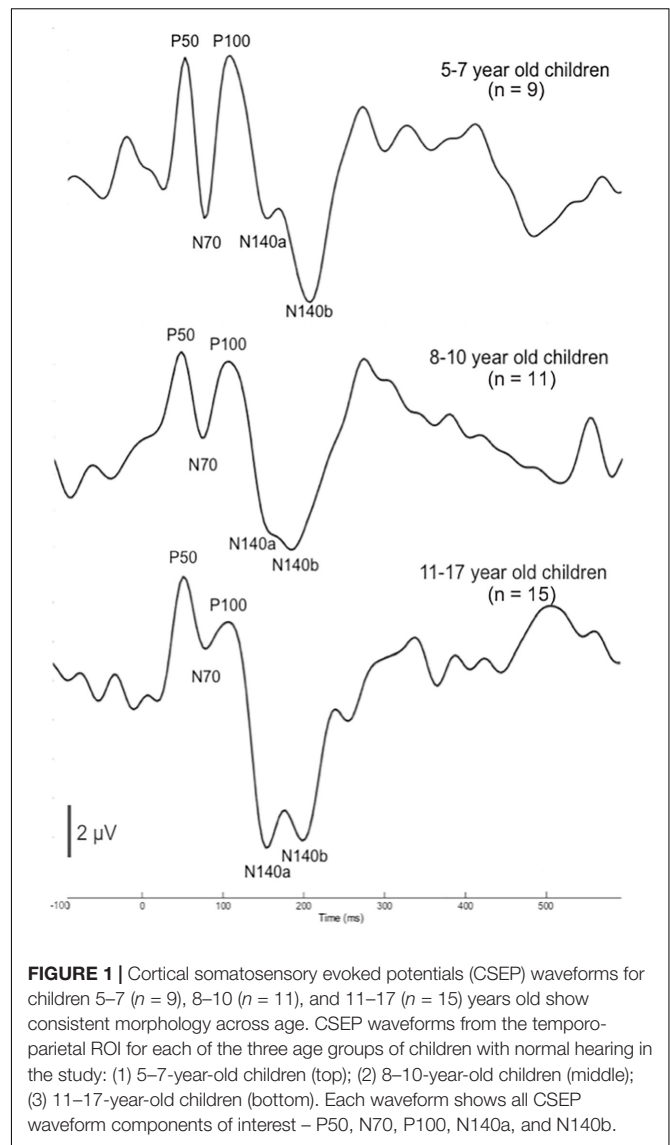
FIGURE 1 | Cortical somatosensory evoked potentials (CSEP) waveforms for children 5–7 (*n* = 9), 8–10 (*n* = 11), and 11–17 (*n* = 15) years old show consistent morphology across age. CSEP waveforms from the temporo-parietal ROI for each of the three age groups of children with normal hearing in the study: (1) 5–7-year-old children (top); (2) 8–10-year-old children (middle); (3) 11–17-year-old children (bottom). Each waveform shows all CSEP waveform components of interest – P50, N70, P100, N140a, and N140b.

that the youngest group (5–7-year-old) showed significantly shorter latencies compared with the 8–10-year-old group for the latency of the N140a CSEP peak ($p = 0.00$). The 5–7-year-old children also exhibited significantly larger CSEP peak-to-peak amplitudes for the N70 ($p = 0.003$; $F = 7.26$), P100 ($p = 0.004$; $F = 6.66$), and N140b ($p = 0.002$; $F = 7.483$) CSEP components relative to the two older groups. The latency finding is reflective of expected developmental patterns and consistent with previous studies (e.g., Allison et al., 1984; Sitzoglou and Fotiou, 1985; Pihko et al., 2009). However, no previous studies have reported on the maturation of amplitude of CSEPs recorded to vibrotactile stimuli in the literature possibly reflecting the inherent variability in absolute amplitude measurements.

Current Density Reconstructions

Results from cortical source localization analysis for NH children (*n* = 35) are shown in **Figure 2**. Initially, sources were
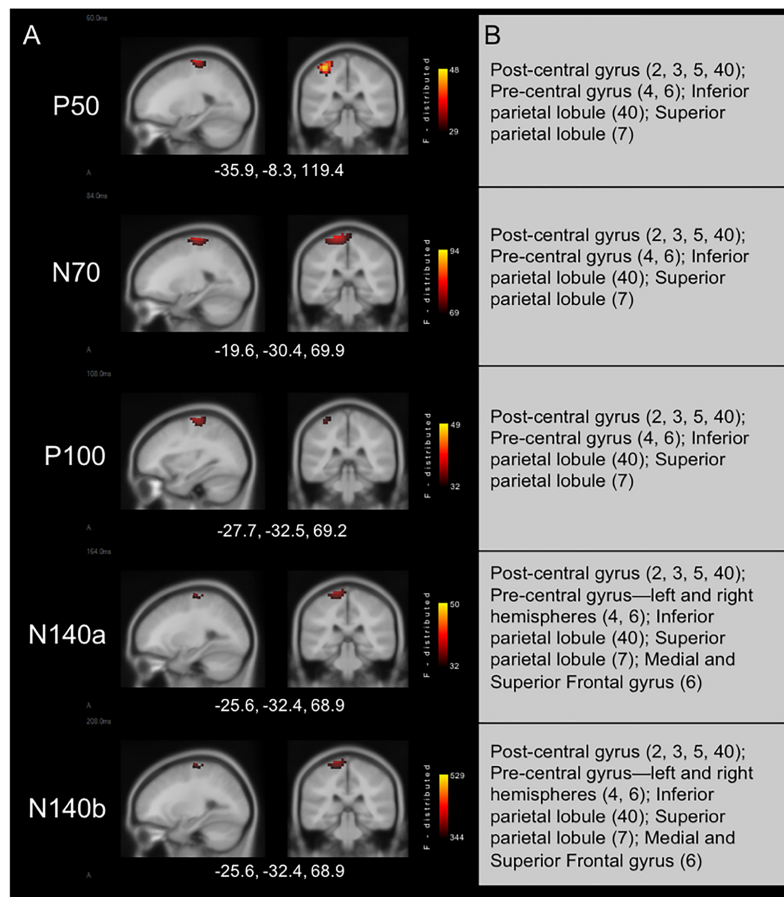
FIGURE 2 | Current density reconstructions for cortical somatosensory evoked potentials (CSEP) waveform components in children with normal hearing (NH) show expected activation of parietal cortices contralateral to the side of stimulation. **(A)** Cortical activations in response to vibrotactile stimulation of the right index finger in children with NH. Activations are organized in rows corresponding to each CSEP waveform component. Sagittal (left) and coronal (right) slices are presented for each of these components. Three-dimensional Montreal Neurological Institute coordinates for each activation are listed below each row of slices. The *F*-distribution scale presents the color gradient associated with the maximum (yellow) through the minimum (black) likelihood for activation as calculated by sLORETA. **(B)** A table listing all areas of significant activation for each CSEP waveform component. Brodmann areas are indicated in parentheses.

computed for each age group separately. However, it was found that all groups' source estimations were comparable. Thus, all participants were combined for final cortical source analysis. Visual inspection and computer-aided determination of the areas of significant activation yielded by sLORETA analysis revealed the following: (1) the P50, N70, and P100 CSEP waveform components presented with virtually the same areas of activation of the left hemisphere. These included, post-central gyrus (BA 2, 3, 5, 40), pre-central gyrus (BA 4, 6), inferior parietal lobule (BA 40), and superior parietal lobule (BA 7); 2) the N140a and N140b generators were also very similar. In addition to all of the previously mentioned activated areas (i.e., for the P50-P100 CSEP components), medial and superior frontal gyri were also activated for the N140a and N140b.

Due to the constancy in peak CSEP components and CDR across the 5–17-year-old age range found in the NH group (*n* = 35) and to increase power for this study, we grouped all CI participants' CSEP data for analysis

and comparison against a subset of age-matched NH children (*n* = 17).

Cochlear Implanted Children (*n* = 12) and Age-Matched NH Children (*n* = 17)
Cortical Somatosensory Evoked Potentials

Both CI and NH children presented with CSEP waveform morphology that was typical of somatosensory evoked responses elicited via vibrotactile stimulation of the finger (Hämäläinen et al., 1990), especially in the parietal ROI contralateral to the side of stimulation. **Figure 3** (left panel) shows grand average results for both groups of children from the LPar ROI during right index finger stimulation. Both groups' results show the characteristic CSEP waveform peaks – P50, N70, P100, N140a, and N140b. While there were no significant differences found in the latencies and amplitudes of the CSEP peaks from this ROI, it is shown here for reference. In contrast, the RTemp ROI waveforms showed a significant
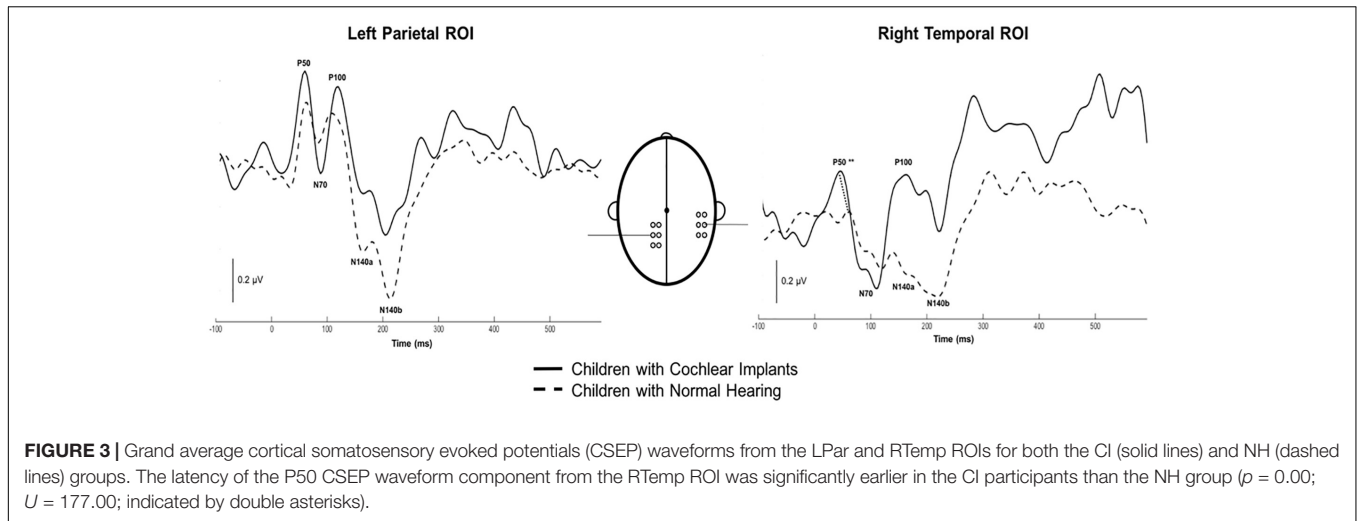
FIGURE 3 | Grand average cortical somatosensory evoked potentials (CSEP) waveforms from the LPar and RTemp ROIs for both the CI (solid lines) and NH (dashed lines) groups. The latency of the P50 CSEP waveform component from the RTemp ROI was significantly earlier in the CI participants than the NH group ($p = 0.00$; $U = 177.00$; indicated by double asterisks).

difference between the latency of the P50 CSEP component between groups (**Figure 3**, right panel; $p = 0.00$; $U = 177.00$) (see also **Table 1**), such that the CI group's latencies were significantly earlier (mean = 49.00 ms; S.D. = 7.5 ms) than the NH group (mean = 65.18 ms; S.D. = 13.17 ms). Additionally, the morphology of the CSEP waveforms differed somewhat between groups and ROIs. For instance, the CI group's grand average waveform shows more robust peaks than those of the NH group. Additionally, the morphology of the CI group's waveform includes a small positivity at approximately 50 ms, followed by a large negativity

around 100 ms, and then another positivity between 150 and 200 ms. This waveform morphology pattern may be more characteristic of the cortical auditory evoked potential, as observed in older children, than the CSEP (e.g., Sharma et al., 1997; Gilley et al., 2005). A shorter P50 latency in the RTemp ROI of the CI children, in response to vibrotactile stimulation, may be a marker of cross-modal reorganization of temporal cortices by the somatosensory system. In addition, somatosensory evoked potentials originating from the auditory cortices may maintain some aspects of auditory evoked potential morphology.

TABLE 1 | Mean latency, standard deviation, 95% confidence intervals, and significance statistics for all CSEP latencies from the LPar and RTemp ROIs.

ROI and Waveform Component	Group	Mean Latency (ms)	Std. Deviation	95% Confidence Interval (Upper – Lower Bound)	p-value; Mann–Whitney U statistic
LPar P50	CI	59.4	7.3	54.8 −64.1	0.39; 121.5
	NH	61.7	6.9	58.1 −65.3	
LPar N70	CI	87.3	9.6	81.2 −93.4	0.19; 72.0
	NH	82.8	8.3	78.5 −87.1	
LPar P100	CI	117.3	14.3	108.2 −126.4	0.37; 81.0
	NH	112.9	14.1	105.7 −120.2	
LPar N140a	CI	152.0	25.3	135.9 −168.0	0.21; 131.0
	NH	161.9	18.2	152.5 −171.3	
LPar N140b	CI	204.7	28.7	186.4 −222.9	0.59; 114.5
	NH	215.3	22.5	203.7 −226.8	
**RTemp P50L	**CI**	**49.0**	**7.5**	**44.3 – 53.7**	**0.00; 177.0**
	NH	**65.2**	**13.2**	**58.4 – 71.9**	
RTemp N70L	CI	83.3	17.5	72.2 −94.4	0.15; 135.0
	NH	94.6	15.9	86.4 −102.7	
RTemp P100L	CI	122.7	24.8	106.9 −138.4	0.37; 81.5
	NH	113.6	18.7	104.0 −123.3	
RTemp N140a	CI	148.3	24.9	132.5 −164.2	0.40; 82.0
	NH	141.4	20.6	130.8 −151.9	
RTemp N140b	CI	216.0	26.1	199.4 −232.6	0.36; 81.5
	NH	204.5	25.7	191.2 −217.7	

Bold type indicates a significant difference between the CI and NH groups. **p < 0.00.

A

	NH Right Finger Stimulation (n = 17)	CI Right Finger Stimulation (n = 12)	CI Left Finger Stimulation (n = 6)
P50	−35.9, −8.3, 119.4	−11.9, −27, 76.1	51.3, −36.7, −7.0
N70	−19.6, −30.4, 69.9	−15.9, −38.1, 71.9	51.3, −36.7, −7.0
P100	−27.7, −32.5, 69.2	−16.4, −30.3, 73.0	51.3, −36.7, −7.0
N140a	−25.6, −32.4, 68.9	−46.0, −34.2, 57.8	59.0, −17.0, −1.7
N140a	−25.6, −32.4, 68.9	−46.0, −34.2, 57.8	59.0, −17.0, −1.7

B

	NH	CI Right Finger	CI Left Finger
P50	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7)	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4); Inferior parietal lobule (40); Superior Temporal gyrus (29, 41, 42); Transverse Temporal gyrus (41, 42); Supramarginal gyrus (40); Angular gyrus (39); Superior Frontal gyrus (6); Paracentral Lobule (6); Insula (13)	Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40); Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Cuneus (18); Superior occipital gyrus (19); Posterior cingulate (30)
N70	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7)	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4); Inferior parietal lobule (40); Superior Temporal gyrus (29, 41, 42); Transverse Temporal gyrus (41, 42); Supramarginal gyrus (40); Angular gyrus (39); Superior Frontal gyrus (6); Paracentral Lobule (6); Insula (13)	Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40); Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Cuneus (18); Superior occipital gyrus (19); Posterior cingulate (30)
P100	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7)	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4); Inferior parietal lobule (40); Superior Temporal gyrus (29, 41, 42); Transverse Temporal gyrus (41, 42); Supramarginal gyrus (40); Angular gyrus (39); Superior Frontal gyrus (6); Paracentral Lobule (6); Insula (13)	Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40); Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Cuneus (18); Superior occipital gyrus (19); Posterior cingulate (30)
N140a	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7)	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4); Inferior parietal lobule (40); Superior Temporal gyrus (29, 41, 42); Transverse Temporal gyrus (41, 42); Supramarginal gyrus (40); Angular gyrus (39); Superior Frontal gyrus (6); Paracentral Lobule (6); Insula (13); Superior Frontal gyrus (10)	Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40); Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Cuneus (18); Superior occipital gyrus (19); Posterior cingulate (30)
N140a	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7)	Post-central gyrus (2, 3, 5, 40); Pre-central gyrus (4); Inferior parietal lobule (40); Superior Temporal gyrus (29, 41, 42); Transverse Temporal gyrus (41, 42); Supramarginal gyrus (40); Angular gyrus (39); Superior Frontal gyrus (6); Paracentral Lobule (6); Insula (13); Superior Frontal gyrus (10)	Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40); Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Cuneus (18); Superior occipital gyrus (19); Posterior cingulate (30)

FIGURE 4 | Current density reconstructions (CDR) for cortical somatosensory evoked potentials (CSEP) in normal hearing and cochlear-implanted children. **(A)** Cortical activations in response to vibrotactile stimulation of the right index finger in children with normal hearing (NH – left panel; n = 35) and cochlear implants (CI – middle panel; n = 12). Additionally, cortical activations in response to stimulation of the left finger in a subset of CI children who received their initial CIs in the right ear are shown in the right-most panel (n = 6). Activations are organized in rows corresponding to each CSEP waveform component (P50, N70, P100, N140a, N140b). CDRs are presented on coronal slices for each of these components. Three-dimensional Montreal Neurological Institute coordinates for each activation are listed below each MRI slice. The F-distribution scale (bottom) presents the color gradient associated with the maximum (yellow) through the minimum (black) likelihood for activation as calculated by sLORETA. **(B)** A table listing all areas of significant activation for each CSEP waveform component. Brodmann areas are indicated in parentheses.

Current Density Reconstructions

Current Density Reconstructions were performed for each of the CSEP waveform components. **Figure 4** shows CDR results for vibrotactile stimulation of the right finger in both the CI group and NH group. In addition, cortical activations in response to left finger stimulation in a subgroup of CI children who received their first implant on the right side are shown in **Figure 4**. In response to vibrotactile stimulation of the right finger (**Figure 4** – middle panel), CI children, as a group, show clear activation of the left (i.e., contralateral to the side of stimulation) somatosensory cortices (i.e., post-central gyrus; BA 3, 2, 5), as well as pre-central gyrus (BA 4, 6), inferior and superior parietal lobules (BA 40 and 7), respectively. Contralateral activations in these areas were expected (i.e., due to the crossover of ascending somatosensory pathways) and consistent with those calculated for the NH group (**Figure 4** – left panel). However, the CI group also showed robust activation of the left temporal cortex – superior temporal gyrus (BA 29, 41, 42); transverse temporal gyrus (BA 41, 42); supramarginal gyrus (BA 40); Angular gyrus (BA 39); superior frontal gyrus (BA 6); paracentral lobule (BA 6); and insula (BA 13). This pattern of activation was consistent for the P50, N70, and P100 CSEP waveform components (**Figure 3**). Both the N140a and N140b presented with CDRs that matched the above CSEP components. However, in addition, frontal cortices contributed to these later components in the CI group. CDR analysis showed that another portion of the superior frontal gyrus contributed to these components (i.e., BA 10; see **Figure 3**).

In a subset of CI participants (n = 6), the left index finger was stimulated in addition to the right (separate conditions). All of these children received implants in their right ears first (mean age at first implantation = 2.89 years; S.D. = ±2.67 years). Five out of six of these participants were also implanted in the left ear at a later date (mean age at second implantation = 7.47 years; S.D. = ±2.91 years). **Figure 4** (right panel) shows the CDR for the right and left index finger stimulation in these children. Interestingly, the cortical activations to the left finger stimulation in the subgroup of children who had received their first CI in the right ear appeared to be centered primarily in auditory cortical areas, with some activity evident in known somatosensory cortical regions. These activated areas included: Superior temporal gyrus (39, 22); Middle temporal gyrus (39, 22); Post-central gyrus (3, 5, 7); Pre-central gyrus (4, 6); Inferior parietal lobule (40); Superior parietal lobule (7); Angular gyrus (39); Supramarginal gyrus (40). These areas of activation were largely found in the right hemisphere, though in the P100 and N140b CSEP components, post-central gyrus (i.e., somatosensory cortex) activations were partially located in the left hemisphere. Activation of auditory processing areas (BA 39, 22) in response to vibrotactile stimulation of the left finger suggests additional cross-modal reorganization of the auditory cortex ipsilateral to the side of first implantation.

CSEP Correlation With Speech Perception in Noise

Cortical somatosensory evoked potentials peak measurements from the LTemp, LPar, RPar, and RTemp ROIs were correlated with results on the BKB-SIN for the CI group to assess the relationship between neurophysiological activity and behavioral speech perception in noise. The latencies of the P50, P100, and N140a from the RTemp ROI all showed significant negative correlations with BKB-SIN score (**Figures 5A–C**; r = −0.679, p = 0.015; r = −0.72, p = 0.008; r = −0.756, p = 0.004, respectively). That is, as latency decreased, BKB-SIN score worsened (see **Figure 5**). That decreased behavioral performance was related to earlier CSEP latencies from the right temporal region of the scalp may suggest that children who have trouble with speech perception in noise, show more evidence of cross-modal reorganization consistent with previous studies (Doucet et al., 2006; Sandmann et al., 2012; Campbell and Sharma, 2016).

DISCUSSION

The objective of the current study was to determine whether cochlear-implanted children would show evidence of cross-modal reorganization of the auditory cortex by the somatosensory system, and if this reorganization would be correlated with behavioral outcomes in these children. Using high-density EEG recorded in response to vibrotactile stimulation of the right and left index finger, we found the following main results: (i) NH children showed expected small age-related variations in CSEP waveform component latencies between 5–7 and 8–10 years of age. However, the generators of cortical somatosensory activation localized to

FIGURE 5 | Scatter plots describing significant relationships between scores on the BKB-SIN and the CSEP P50 **(A)**, P100 **(B)**, and N140a **(C)** waveform components in the CI group.
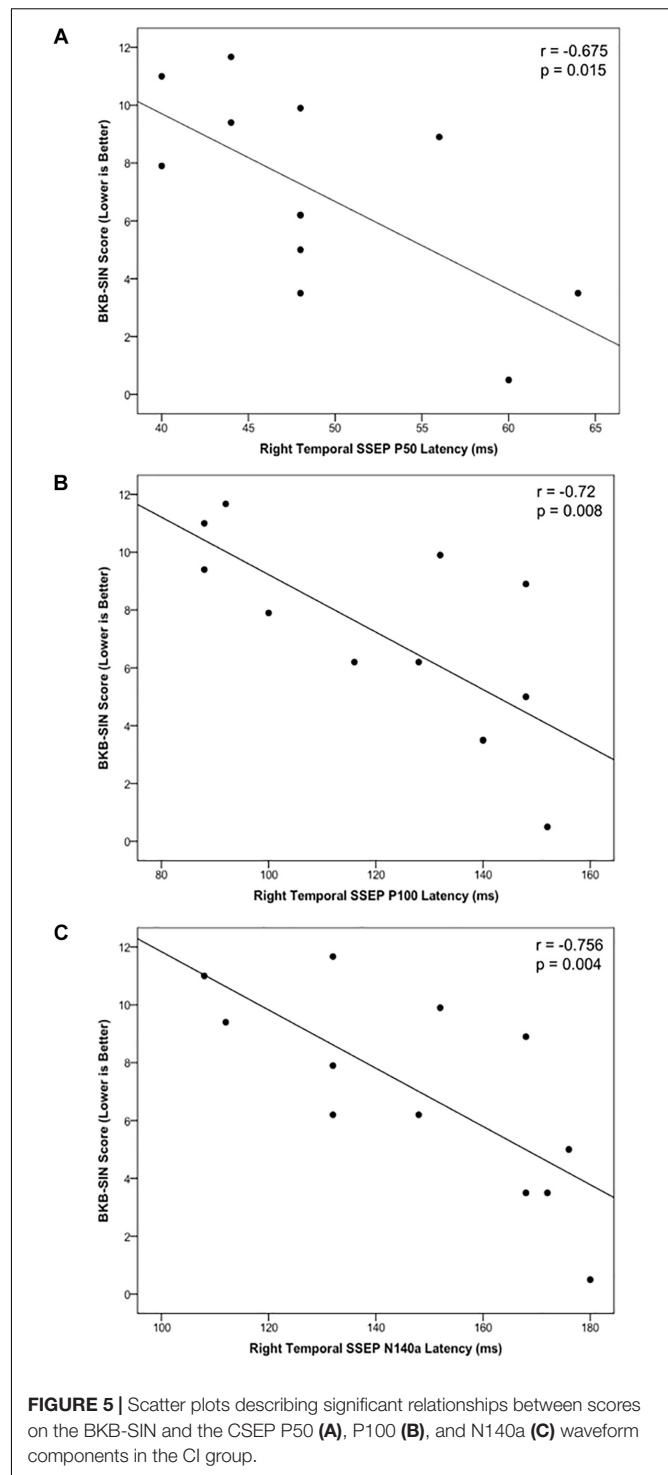
the post-central gyrus, association cortices of the parietal lobe, and pre-central gyrus contralateral to the side of stimulation across the age span; (ii) CSEP morphology and latencies were consistent between CI and NH children in the LPar ROI, but not the RTemp ROI – the latter exhibiting significantly earlier CSEP latencies in the CI group, (iii) CDR of right finger vibrotactile stimulation revealed expected activation of

the left somatosensory cortices in both NH and CI children. However, CI participants showed activation of auditory processing areas in the left temporal and parietal association cortex by vibrotactile stimulation; (iv) In a subset of children who received their first CIs in the right ear, we saw significant cross-modal activation in the right hemisphere, suggesting that the cortex ipsilateral to the first CI (i.e., the cortex less activated by the first implant) is highly susceptible to cross-modal activation; (v) CSEP latencies from the RTemp ROI were significantly correlated with speech perception in noise results, which may be an indication that poorer behavioral outcomes with the CI are associated with greater degrees of cross-modal reorganization.

CSEP Development in Typical Children

The morphological aspects of the participants' CSEP data (**Figure 1**) were consistent with previous reports. For instance, one study (Hämäläinen et al., 1990) used vibrotactile stimuli applied to the middle finger to evoke potentials from the primary and secondary somatosensory cortex and showed P50, N70, P100, and N140 CSEP components. The CSEP waveforms in the current study consistently showed all peak components across the age range. This pattern of stability of peak components across age differs from the developmental progression of cortical evoked potentials recorded to visual and auditory stimuli, which changes significantly throughout the age range studied here (e.g., Placzek et al., 1985; Ellingson, 1986; Ponton et al., 2000; Gilley et al., 2005). In fact, we have observed that the CSEP waveforms of adults (ranging in age from 21 to 71) are morphologically comparable to the CSEP waveforms presented in the current study (Cardon and Sharma, 2018). Thus, it appears that the CSEP may be unique among modalities, in that major peak components are present and remain constant from school age through adulthood. More significant changes in the CSEP waveform likely occur earlier in life (i.e., by age 4) and afterward slow considerably (e.g., Desmedt et al., 1976; Allison et al., 1984; Sitzoglou and Fotiou, 1985; Fagan et al., 1987; Eggermont, 1988; Pihko et al., 2009). Our data seem to support the above notion in that any differences that were seen in amplitude or latency of CSEP components were noted between the youngest and two older groups. These findings may exhibit the ending of the early childhood phase of development of the CSEP (i.e., slowing after age 4 years). Given the age range of the current study, the participants may have been too mature for observation of more robust developmental effects.

Current density reconstructions yielded results that matched both our hypothesis and previously reported findings. Numerous investigations have outlined the generators of the various CSEP components. For instance, previous studies have found that the P50 CSEP component is generated in the post-central gyrus of the cerebral hemisphere contralateral to the side of stimulation in primary somatosensory cortex (SI; e.g., Mauguière et al., 1983). The N70 also appears to be generated in contralateral SI (Michie et al., 1987). Hämäläinen et al. (1990) proposed, based on both animal and human studies (e.g., Hari et al., 1983, 1984; Hämäläinen et al., 1990),

that the P100 originates from a combination of ipsi- and contralateral SII cortex. The N140 CSEP component seems to have a number of generators, which are likely distributed throughout the posterior parietal regions of the cortex, with the strongest contributions coming from cortices contralateral to the stimuli. Specifically, some have proposed that the N140 is influenced by generators in contralateral SII (Hari et al., 1983, 1984, 1993) and also contains activity from Brodmann are 46 and other frontal cortices (Desmedt and Tomberg, 1989; Hämäläinen et al., 1990). The current results mirror these reports' descriptions of the sources of cortical activity that contribute to the CSEP. That is, all CSEP components from the current study were localized to the primary, secondary, and association somatosensory cortices (BA 3, 2, 1, 5, and 7) in the hemisphere contralateral to the side of stimulation. In addition, pre-central gyrus was activated in the CDRs for each of the CSEP components. This activity may be mediated by connections between the pre- and post-central gyrus (e.g., Pandya and Kuypers, 1969). Finally, it may be interesting to note that the N140a and N140b CSEP components show activation of medial and superior frontal cortices (i.e., Brodmann area 6), which is consistent with the characterization of the generators for these components offered by Hämäläinen et al. (1990) that indicate frontal cortex involvement in the generation of the N140 CSEP.

Evidence and Possible Mechanisms of Somatosensory Cross-Modal Reorganization in CI Children

In CI children, we saw at least two types of evidence for somatosensory cross-modal reorganization: earlier CSEP latencies in the RTemp ROI and activation in auditory processing areas in superior and transverse temporal cortices, as well as cortical regions important to language processing (i.e., parts of Wernicke's area), in response to somatosensory stimuli (see **Figures 1**, **2**). A number of previous studies have reported similar findings in both animals and humans (Neville and Lawson, 1987; Levänen et al., 1998; Baldwin, 2002; Auer et al., 2007; Sharma et al., 2007; Shore et al., 2008; Karns et al., 2012; Sandmann et al., 2012; Campbell and Sharma, 2014, 2016; Sharma and Glick, 2016). For instance, a recent study from our lab showed a similar pattern of earlier latencies of cortical visual evoked potentials, as well as activation of auditory processing areas in response to visual stimuli, in CI children (Campbell and Sharma, 2016). In contrast, one study in the literature appears to present conflicting evidence to the present results. That is, Hickok et al. (1997) used MEG to study possible cross-modal reorganization in one deaf young adult. These investigators reported that they found no evidence of somatosensory-to-auditory cross-modal reorganization in this subject. However, these investigators used a tapping stimulus applied to the finger, instead of a vibrotactile stimulus. Because of the similarity between sound and vibration, the auditory cortex may be better suited to process vibrotactile input, while this may not be the case with other types of stimuli (i.e., tapping). Additionally, this study

only assessed these factors in one subject. Thus, the Hickok et al. (1997) study may not be directly comparable to this, and other, studies that do show evidence of somatosensory cross-modal reorganization. Overall, the majority of studies in the literature submit that cross-modal reorganization of the auditory cortex by the somatosensory system can occur in deaf individuals. We add our findings as another piece of converging evidence that supports this notion in CI children. Future studies should endeavor to replicate these initial findings given the limited sample of CI children in the current study.

The current CI participants presented with robust activity in response to vibrotactile stimuli in primary and secondary auditory cortices, as well as supramarginal and angular gyri, which make up part of Wernicke's area, important in receptive language processing. Such findings were not the case in NH participants, despite the presence of continuous auditory masking noise. While some have shown cross-modal reorganization primarily in higher order auditory cortices in deaf individuals (Kral, 2007), there is a precedent for primary auditory cortical reorganization. That is, Auer et al. (2007) presented evidence of activity arising from primary auditory cortices in response to vibrotactile stimulation in six deaf young adults using fMRI. Additionally, MEG source analysis performed by Levänen et al. (1998) showed bilateral activation of superior temporal gyrus (STG) in one adult with congenital deafness. It is possible that normally unisensory areas are taken over by other sensory modalities (Auer et al., 2007). Numerous studies have established a precedent for both intracortical, thalamocortical, and subcortical anatomical (e.g., Foxe et al., 2000; Schroeder et al., 2001; Gobbelé et al., 2003; Kayser et al., 2005; Caetano and Jousmäki, 2006; Hackett et al., 2007), as well as functional (Jousmäki and Hari, 1998; Lakatos et al., 2007; Brett-Green et al., 2008), connections between the somatosensory and auditory systems. Subcortically driven cross-modal reorganization of the primary and secondary auditory cortices appears to be a distinct possibility, especially in congenitally deafened individuals whose deprivation was a factor during the development of subcortical-cortical pathways (Soto-Faraco and Deco, 2009; Zeng et al., 2012). These findings are also in agreement with previous data from MEG recordings performed by our group (Sharma et al., 2007), which showed auditory and multimodal association (i.e., Wernicke's area) activity in response to vibrotactile stimulation of the hands in one deaf adult. In addition to subcortical contributions, given the multimodal nature of these areas, it is possible that unmasking and enhancement of latent multisensory connections when one modality is deprived may contribute to cross-modal reorganization in these cortical regions (Levänen et al., 1998; Auer et al., 2007). Such enhancement could lead to both shorter CSEP latencies – via improved synaptic efficiency – and cross-modal activation.

It may be interesting to note that in all of the previous studies examining cross-modal reorganization in deaf individuals, the duration of deafness was extensive (i.e., into adulthood). For example, the subject recruited for study in Levänen et al.

(1998) was 77 years of age and had been deaf for all or most of his life. Though the duration of deafness in the current participants was lower than many of the previous studies – the average age of implantation of children in the current study was 3.9 years – it was beyond the sensitive period for auditory cortical development (i.e., 3.5 years; Sharma et al., 2002a,b). Given that many more children receive their implants around the FDA approved age of 1 year currently, future studies should investigate cross-modal reorganization in children who were fitted with CIs at early ages in order to determine if cross-modal reorganization takes place when the duration of deafness is very short in childhood (see Meredith and Allman, 2012).

Bilateral Implantation and Somatosensory Cross-Modal Reorganization

In the current results, children who received their CIs in the right ear first and who later received a second CI in the left ear showed differing patterns of cortical activation between the right and left cortical hemispheres in response to somatosensory stimulation of the right and left index fingers. Stimulation of the right index finger lead to activity patterns that, for the most part, were consistent with typical somatosensory responses (post- and pre-central gyri, BA 3, 5; and 4, respectively) and activation of auditory areas (BA 39, 22; consistent with our overall finding of cross-modal recruitment for the CI group as a whole). Results from the stimulation of the left finger were, however, quite distinct. That is, instead of the most robust activations being localized to pre- and post-central gyri, cortical generators were estimated to be in the right temporal areas, especially for the P50 and N70 CSEP components. This finding is suggestive of a higher degree of cross-modal reorganization. Our results agree with the results of a study performed by Kral et al. (2002) in congenitally deaf white cats. These investigators reported that cats who had received their implants late (i.e., >5 months) showed decreased activations in the auditory cortex ipsilateral to the implanted ear, while responses coming from the contralateral auditory cortex did not show the same pattern. Additionally, Gordon and Papsin (2009) reported that longer durations of unilateral CI use in humans (i.e., >2 years) lead to abnormally high lateralization of EEG signals to the auditory cortex contralateral to the CI. In contrast, the auditory cortex ipsilateral to the implant showed very low activation (Gordon et al., 2013). The participants who received their CI in the right ear first were fitted with their first implant around the age of 2.89 years (±2.67 years), which is under the sensitive period for auditory cortical maturation (i.e., 3.5 years) reported by Sharma et al. (2002a,b). Consistent stimulation of the left auditory cortex via a CI placed in the right ear during the sensitive period may have contributed to the results from right finger stimulation that suggest near normal somatosensory activation in children who received their first CI in the right ear and some activation of auditory areas (**Figure 4**, right panel). In contrast to right finger stimulation, left finger

stimulation lead to robust activation of right auditory cortices in these children (**Figure 4**, left panel) suggesting that the "weaker," ipsilateral cortex is highly amenable to cross-modal recruitment by the somatosensory modality. Overall, these children spent years without optimal auditory input to the right auditory cortices, which may have allowed cross-modal reorganization of these cortical areas in the cortex ipsilateral to the CI (e.g., Kral et al., 2013; Jiwani et al., 2016). Unfortunately, the present sample of bilaterally implanted children in which left finger stimulation was performed only amounted to six participants. Thus, the above results should be interpreted with caution. Additional studies should be carried out to further investigate the potential effects of unilateral cochlear implantation and hemispheric differences in cross-modal reorganization.

Connections Between Somatosensory Cross-Modal Reorganization in CI Children and Speech Perception

The current findings suggest a relationship between somatosensory cross-modal reorganization and speech perception in noise in CI children. This relationship was such that children who had poorer speech perception in noise with their implant showed more cross-modal re-organization. This suggests that these individuals may have been activating the somatosensory system to help disambiguate the impoverished signal input from the CI. There are numerous reports in the literature that support the notion of the somatosensory system being involved in speech perception (Liberman et al., 1967; Liberman and Mattingly, 1985; Fadiga et al., 2002; Watkins et al., 2003; Galantucci et al., 2006; Meister et al., 2007; Skipper et al., 2007; Ito et al., 2009; Russo et al., 2012; Ammirante et al., 2013; Hubka et al., 2015). For example, Gick and Derrick (2009) tested NH participants' phoneme perception (e.g., "p" vs. "b") while simultaneously presenting inaudible puffs of air to their skin. Interestingly, these participants more often perceived a phoneme as being aspirated when the air puff was presented, reflecting speech-related auditory-tactile integration. Deaf individuals have also shown evidence that they differentiate same-sex talkers and musical instruments solely by using vibrotactile information (Russo et al., 2012; Ammirante et al., 2013). These abilities suggest that the somatosensory system can decipher information that is highly relevant to speech perception, such as frequency and timbre. Furthermore, Ito et al. (2009) showed evidence that stretching the facial skin affected the perception of an auditory phoneme. They reasoned that, since the somatosensory receptors responsible for stretching and orientation of the skin are constantly and systematically being activated during speech *production*, somatosensory input may also be a vital part of speech *perception*. Animal studies have also presented evidence that the somatosensory system may be involved in vocalization behavior. For instance, Hubka et al. (2015) showed that vocalizations in deaf cats may be (partially) influenced by an auditory feedback loop that is mediated by somatosensory perception. These findings are paralleled by studies that have demonstrated that the motor cortices thought to be related to speech production may be activated during speech perception

(Fadiga et al., 2002; Watkins et al., 2003; Meister et al., 2007). Thus, it is reasonable to believe that CI users may rely on vibrotactile input to improve understanding (Gick and Derrick, 2009; Huang et al., 2017), especially under challenging listening conditions, such as speech presented in background noise. As such, future research efforts should be devoted to exploring the potential benefits of tactile stimulation for aiding CI users (Huang et al., 2017).

CONCLUSION

The current study examined cross-modal reorganization between the somatosensory and auditory systems in children with CIs. CDRs secondary to stimulation of the right index finger revealed cortical activation in somatosensory cortices in both NH and CI groups, while the CI group also presented with cortical activity localized to auditory cortical areas suggestive of cross-moral re-organization. Our results also suggest that the cortex ipsilateral to the first implanted ear (which receives weaker auditory input than the contralateral cortex) is highly susceptible to cross-modal reorganization. Finally, children who have difficulty perceiving speech with the CI are more likely to show cross-modal re-organization, likely as a compensatory adaptation.

DATA AVAILABILITY

All datasets analyzed for this study are included in the manuscript and the supplementary files.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Belmont Report as reviewed by the Institutional Review Board of the University of Colorado Boulder with written informed consent from all subjects or their guardians. Additionally, all children aged seven and above provided written assent prior to participating in the study. All subjects gave written informed consent/assent in accordance with the Declaration of Helsinki. The protocol was approved by the Institutional Review Board of the University of Colorado Boulder.

AUTHOR CONTRIBUTIONS

Both authors contributed equally to the conceptualization, hypothesis development, recruitment, data acquisition, data analysis and interpretation, writing, and editing of the manuscript.

FUNDING

REFERENCES

Allison, T., Hume, A. L., Wood, C. C., and Goff, W. R. (1984). Developmental and aging changes in somatosensory, auditory and visual evoked potentials. *Electroencephalogr. Clin. Neurophysiol.* 58, 14–24. doi: 10.1016/0013-4694(84) 90196-2

Allman, B. L., Keniston, L. P., and Meredith, M. A. (2009). Adult deafness induces somatosensory conversion of ferret auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 106, 5925–5930. doi: 10.1073/pnas.0809483106

Ammirante, P., Russo, F., Good, A., and Fels, D. (2013). Feeling voices. *PLoS One* 8:e53585. doi: 10.1371/journal.pone.0053585

Auer, E. T., Bernstein, L. E., Sungkarat, W., and Singh, M. (2007). Vibrotactile activation of the auditory cortices in deaf versus hearing adults. *Neuroreport* 18, 645–648. doi: 10.1097/wnr.0b013e3280d943b9

Baldwin, R. L. (2002). *Functional Reallocation of the Auditory Cortex in Individuals who are Deaf*. Doctoral dissertation, Gallaudet University, Washington, DC.

Bavelier, D., and Hirshorn, E. (2010). I see where you're hearing: how cross-modal plasticity may exploit homologous brain structures. *Nat. Neurosci.* 13, 1309–1311. doi: 10.1038/nn1110-1309

Bavelier, D., and Neville, H. (2002). Cross-modal plasticity: where and how? *Nat. Rev. Neurosci.* 3, 443–452. doi: 10.1038/nrn848

Bench, J., Kowal, Å., and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br. J. Audiol.* 13, 108–112. doi: 10.3109/03005367909078884

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Brett-Green, B. A., Miller, L. J., Gavin, W. J., and Davies, P. L. (2008). Multisensory integration in children: a preliminary ERP study. *Brain Res.* 1242, 283–290. doi: 10.1016/j.brainres.2008.03.090

Caetano, G., and Jousmäki, V. (2006). Evidence of vibrotactile input to human auditory cortex. *Neuroimage* 29, 15–28. doi: 10.1016/j.neuroimage.2005. 07.023

Campbell, J., and Sharma, A. (2014). Cross-modal re-organization in adults with early stage hearing loss. *PLoS One* 9:e90594. doi: 10.1371/journal.pone.0090594

Campbell, J., and Sharma, A. (2016). Visual cross-modal re-organization in children with cochlear implants. *PLoS One* 11:e0147793. doi: 10.1371/journal. pone.0147793

Cardon, G., and Sharma, A. (2018). Somatosensory cross-modal reorganization in adults with age-related, early-stage hearing loss. *Front. Hum. Neurosci.* 12:172. doi: 10.3389/fnhum.2018.00172

Centers for Disease Control and Prevention [CDC] (2010). Identifying infants with hearing loss - United States, 1999-2007. *MMWR Morb. Mortal. Wkly. Rep.* 59, 220–223.

Clemo, H. R., Lomber, S. G., and Meredith, M. A. (2014). Synaptic basis for cross-modal plasticity: enhanced supragranular dendritic spine density in anterior ectosylvian auditory cortex of the early deaf cat. *Cereb. Cortex* 26, 1365–1376. doi: 10.1093/cercor/bhu225

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j. jneumeth.2003.10.009

Desmedt, J. E., Brunko, E., and Debecker, J. (1976). Maturation of the somatosensory evoked potentials in normal infants and children, with special reference to the early N 1 component. *Electroencephalogr. Clin. Neurophysiol.* 40, 43–58. doi: 10.1016/0013-4694(76)90178-4

Desmedt, J. E., and Tomberg, C. (1989). Mapping early somatosensory evoked potentials in selective attention: critical evaluation of control conditions used for titrating by difference the cognitive P30, P40, P100, and 140. *Electroencephalogr. Clin. Neurophysiol.* 74, 321–346. doi: 10.1016/0168-5597(89)90001-4

Doucet, M. E., Bergeron, F., Lassonde, M., Ferron, P., and Lepore, F. (2006). Cross-modal reorganization and speech perception in cochlear implant users. *Brain* 129(Pt 12), 3376–3383. doi: 10.1093/brain/awl264

Eggermont, J. J. (1988). On the rate of maturation of sensory evoked potentials. *Electroencephalogr. Clin. Neurophysiol.* 70, 293–305. doi: 10.1016/0013-4694(88)90048-x

Ellingson, R. J. (1986). Development of visual evoked potentials and photic driving responses in normal full term, low risk premature, and trisomy-21 infants

during the first year of life. *Electroencephalogr. Clin. Neurophysiol.* 63, 309–316. doi: 10.1016/0013-4694(86)90015-5

Etymotic Research (2005). *BKB-SINTM Speech-in-Noise Test*. Elk Grove Village, IL: Etymotic Research.

Evans, A. C., Collins, D. L., Mills, S. R., Brown, E. D., Kelly, R. L., and Peters, T. M. (1993). "3D statistical neuroanatomical models from 305 MRI volumes," in *Proceedings of the IEEE-Nuclear Science Symposium and Medical Imaging Conference*, San Francisco, CA, 1813–1817.

Fadiga, L., Craighero, L., Buccino, G., and Rizzolatti, G. (2002). Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur. J. Neurosci.* 15, 399–402. doi: 10.1046/j.0953-816x.2001.01874.x

Fagan, E. R., Taylor, M. J., and Logan, W. J. (1987). Somatosensory evoked potentials: part I. A review of neural generators and special considerations in pediatrics. *Pediatr. Neurol.* 3, 189–196. doi: 10.1016/0887-8994(87) 90015-4

Fine, I., Finney, E. M., Boynton, G. M., and Dobkins, K. R. (2005). Comparing the effects of auditory deprivation and sign language within the auditory and visual cortex. *J. Cogn. Neurosci.* 17, 1621–1637. doi: 10.1162/089892905774597173

Fink, N. E., Wang, N. Y., Visaya, J., Niparko, J. K., Quittner, A., Eisenberg, L. S., et al. (2007). Childhood development after cochlear implantation (CDaCI) study: design and baseline characteristics. *Cochlear Implants Int.* 8, 92–116. doi: 10.1002/cii.333

Finney, E. M., Clementz, B. A., Hickok, G., and Dobkins, K. R. (2003). Visual stimuli activate auditory cortex in deaf subjects: evidence from MEG. *Neuroreport* 14, 1425–1427. doi: 10.1097/00001756-200308060-00004

Finney, E. M., Fine, I., and Dobkins, K. R. (2001). Visual stimuli activate auditory cortex in the deaf. *Nat. Neurosci.* 4, 1171–1173. doi: 10.1038/nn763

Food and Drug Administration [FDA] (2012). *Manufacturers' Voluntary Reports of Registered Devices to the U.S. Food and Drug Administration. Retrieved February 22, 2019, from National Institute on Deafness and Other Communication Disorders. (n.d.). Quick Statistics [NIDCD Health Information]*. Available at: http://www.nidcd.nih.gov/health/statistics/pages/ quick.aspx (accessed February 22, 2019).

Foxe, J. J., Morocz, I. A., Murray, M. M., Higgins, B. A., Javitt, D. C., and Schroeder, C. E. (2000). Multisensory auditory-somatosensory interactions in early cortical processing revealed by high-density electrical mapping. *Brain Res. Cogn. Brain Res.* 10, 77–83. doi: 10.1016/s0926-6410(00) 00024-0

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., and Ebersole, J. S. (2002). A standardized boundary element method volume conductor model. *Clin. Neurophysiol.* 113, 702–712. doi: 10.1016/s1388-2457(02)00030-5

Galantucci, B., Fowler, C. A., and Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychon. Bull. Rev.* 13, 361–377. doi: 10.3758/bf03193857

Geers, A. E. (2006). Factors influencing spoken language outcomes in children following early cochlear implantation. *Adv. Otorhinolaryngol.* 64, 50–65. doi: 10.1159/000094644

Geers, A. E., Moog, J. S., Biedenstein, J., Brenner, C., and Hayes, H. (2009). Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. *J. Deaf Stud. Deaf Educ.* 14, 371–385. doi: 10.1093/deafed/ enn046

Gick, B., and Derrick, D. (2009). Aero-tactile integration in speech perception. *Nature* 462, 502–504. doi: 10.1038/nature08572

Gilley, P., Sharma, A., Dorman, M., and Martin, K. (2005). Developmental changes in refractoriness of the cortical auditory evoked potential. *Clin. Neurophysiol.* 116, 648–657. doi: 10.1016/j.clinph.2004.09.009

Gilley, P. M., Sharma, A., and Dorman, M. F. (2008). Cortical reorganization in children with cochlear implants. *Brain Res.* 1239, 56–65. doi: 10.1016/j.brainres.2008.08.026

Gobbelé, R., Schürmann, M., Forss, N., and Juottonen, K. (2003). Activation of the human posterior parietal and temporoparietal cortices during audiotactile interaction. *Neuroimage* 20, 503–511. doi: 10.1016/s1053-8119(03)00312-4

Gordon, K. A., and Papsin, B. C. (2009). Benefits of short interimplant delays in children receiving bilateral cochlear implants. *Otol. Neurotol.* 30, 319–331. doi: 10.1097/MAO.0b013e31819a8f4c

Gordon, K. A., Wong, D. D. E., and Papsin, B. C. (2013). Bilateral input protects the cortex from unilaterally-driven reorganization in children who are deaf. *Brain* 136, 1609–1625. doi: 10.1093/brain/awt052

Grech, R., Cassar, T., and Muscat, J. (2008). Review on solving the inverse problem in EEG source analysis. *J. Neuroeng. Rehabil.* 5:25. doi: 10.1186/1743-0003-5-25

Hackett, T., Smiley, J., Ulbert, I., and Karmos, G. (2007). Sources of somatosensory input to the caudal belt areas of auditory cortex. *Perception* 36, 1419–1430. doi: 10.1068/p5841

Hallez, H., Vanrumste, B., and Grech, R. (2007). Review on solving the forward problem in EEG source analysis. *J. Neuroeng. Rehabil.* 4:46. doi: 10.1186/1743-0003-4-46

Hämäläinen, H., Kekoni, J., Sams, M., Reinikainen, K., and Näätänen, R. (1990). Human somatosensory evoked potentials to mechanical pulses and vibration: contributions of SI and SII somatosensory cortices to P50 and P100 components. *Electroencephalogr. Clin. Neurophysiol.* 75, 13–21. doi: 10.1016/0013-4694(90)90148-d

Hari, R., Hämäläinen, M., and Kaukoranta, E. (1983). Neuromagnetic responses from the second somatosensory cortex in man. *Acta Neurol.* 68, 207–212. doi: 10.1111/j.1600-0404.1983.tb04828.x

Hari, R., Karhu, J., Hämäläinen, M., Knuutila, J., Salonen, O., Sams, M., et al. (1993). Functional organization of the human first and second somatosensory cortices: a neuromagnetic study. *Eur. J. Neurosci.* 5, 724–734. doi: 10.1111/j.1460-9568.1993.tb00536.x

Hari, R., Reinikainen, K., Kaukoranta, E., Hämäläinen, M., Ilmoniemi, R., Penttinen, A., et al. (1984). Somatosensory evoked cerebral magnetic fields from SI and SII in man. *Electroencephalogr. Clin. Neurophysiol.* 57, 254–263. doi: 10.1016/0013-4694(84)90126-3

Harrison, R. V., Gordon, K. A., and Mount, R. J. (2005). Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. *Dev. Psychobiol.* 46, 252–261. doi: 10.1002/dev.20052

Hickok, G., Poeppel, D., Clark, K., and Buxton, R. B. (1997). Sensory mapping in a congenitally deaf subject: MEG and fRMI studies of cross-modal non-plasticity. *Hum. Brain Mapp.* 5, 437–444. doi: 10.1002/(SICI)1097-019319975:6

Holt, R. F., and Svirsky, M. A. (2008). An exploratory look at pediatric cochlear implantation: is earliest always best? *Ear Hear.* 29, 492–511. doi: 10.1097/AUD.0b013e31816c409f

Huang, J., Sheffield, B., Lin, P., and Zeng, F.-G. (2017). Electro-tactile stimulation enhances cochlear implant speech recognition in noise. *Sci. Rep.* 7:2196. doi: 10.1038/s41598-017-02429-1

Hubka, P., Konerding, W., and Kral, A. (2015). Auditory feedback modulates development of kitten vocalizations. *Cell Tissue Res.* 361, 279–294. doi: 10.1007/s00441-014-2059-6

Ito, T., Tiede, M., and Ostry, D. J. (2009). Somatosensory function in speech perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1245–1248. doi: 10.1073/pnas.0810063106

Jiwani, S., Papsin, B. C., and Gordon, K. A. (2016). Early unilateral cochlear implantation promotes mature cortical asymmetries in adolescents who are deaf. *Hum. Brain Mapp.* 37, 135–152. doi: 10.1002/hbm.23019

Johnson, D., Jürgens, R., Kongehl, G., and Kornhuber, H. H. (1975). Somatosensory evoked potentials and magnitude of perception. *Exp. Br. Res.* 22, 331–334.

Jousmäki, V., and Hari, R. (1998). Parchment-skin illusion: sound-biased touch. *Curr. Biol.* 8:R190.

Karns, C., Dow, M., and Neville, H. (2012). Altered cross-modal processing in the primary auditory cortex of congenitally deaf adults: a visual-somatosensory fMRI study with a double-flash illusion. *J. Neurosci.* 32, 9626–9638. doi: 10.1523/JNEUROSCI.6488-11.2012

Kayser, C., Petkov, C. I., Augath, M., and Logothetis, N. K. (2005). Integration of touch and sound in auditory cortex. *Neuron* 48, 373–384. doi: 10.1016/j.neuron.2005.09.018

Kok, M. A., Chabot, N., and Lomber, S. G. (2014). Cross-modal reorganization of cortical afferents to dorsal auditory cortex following early- and late-onset deafness. *J. Comp. Neurol.* 522, 654–675. doi: 10.1002/cne.23439

Kral, A. (2007). Unimodal and cross-modal plasticity in the "deaf" auditory cortex. *Int. J. Audiol.* 46, 479–493. doi: 10.1080/14992020701383027

Kral, A., Hartmann, R., Tillein, J., Heid, S., and Klinke, R. (2002). Hearing after congenital deafness: central auditory plasticity and sensory deprivation. *Cereb. Cortex* 12, 797–807. doi: 10.1093/cercor/12.8.797

Kral, A., Heid, S., Hubka, P., and Tillein, J. (2013). Unilateral hearing during development: hemispheric specificity in plastic reorganizations. *Front. Syst. Neurosci.* 7:93. doi: 10.3389/fnsys.2013.00093

Lakatos, P., Chen, C., O'Connell, M., and Mills, A. (2007). Neuronal oscillations and multisensory interaction in primary auditory cortex. *Neuron* 53, 279–292. doi: 10.1016/j.neuron.2006.12.011

Levänen, S., Jousmäki, V., and Hari, R. (1998). Vibration-induced auditory-cortex activation in a congenitally deaf adult. *Curr. Biol.* 8, 869–872. doi: 10.1016/s0960-9822(07)00348-x

Liberman, A. M., Cooper, F. S., Shankweiler, D. P., and Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychol. Rev.* 74, 431–461.

Liberman, A. M., and Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition* 21, 1–36. doi: 10.1016/0010-0277(85)90021-6

Lomber, S. G., Meredith, M. A., and Kral, A. (2010). Cross-modal plasticity in specific auditory cortices underlies visual compensations in the deaf. *Nat. Neurosci.* 13, 1421–1427. doi: 10.1038/nn.2653

Lund, E. (2016). Vocabulary knowledge of children with cochlear implants: a meta-analysis. *J. Deaf Stud. Deaf Educ.* 21, 107–121. doi: 10.1093/deafed/env060

Luu, P., and Ferree, T. (2000). *Determination of the Geodesic Sensor Nets' Electrode Positions and Their 10–10 International Equivalents.* Eugene, OR: Electrical Geodesics, Inc.

Mauguière, F., Desmedt, J. E., and Courjon, J. (1983). Astereognosis and dissociated loss of frontal or parietal components of somatosensory evoked potentials in hemispheric lesions detailed correlations with clinical signs and computerized tomographic scanning. *Brain* 106, 271–311. doi: 10.1093/brain/106.2.271

Meister, I. G., Wilson, S. M., Deblieck, C., Wu, A. D., and Iacoboni, M. (2007). The essential role of premotor cortex in speech perception. *Curr. Biol.* 17, 1692–1696. doi: 10.1016/j.cub.2007.08.064

Merabet, L. B., and Pascual-Leone, A. (2010). Neural reorganization following sensory loss: the opportunity of change. *Nat. Rev. Neurosci.* 11, 44–52. doi: 10.1038/nrn2758

Meredith, M. A., and Allman, B. L. (2012). Early hearing-impairment results in crossmodal reorganization of ferret core auditory cortex. *Neural Plast.* 2012:601591.

Meredith, M. A., and Lomber, S. G. (2011). Somatosensory and visual crossmodal plasticity in the anterior auditory field of early-deaf cats. *Hear. Res.* 280, 38–47. doi: 10.1016/j.heares.2011.02.004

Michie, P. T., Bearpark, H. M., Crawford, J. M., and Glue, L. C. (1987). The effects of spatial selective attention on the somatosensory event-related potential. *Psychophysiology* 24, 449–463. doi: 10.1111/j.1469-8986.1987.tb00316.x

Neville, H. J., and Lawson, D. (1987). Attention to central and peripheral visual space in a movement detection task: an event-related potential and behavioral study. II. Congenitally deaf adults. *Brain Res.* 405, 268–283. doi: 10.1016/0006-8993(87)90296-4

Neville, H. J., Schmidt, A., and Kutas, M. (1983). Altered visual-evoked potentials in congenitally deaf adults. *Brain Res.* 266, 127–132. doi: 10.1016/0006-8993(83)91314-8

Nicholas, J. G., and Geers, A. E. (2007). Will they catch up? The role of age at cochlear implantation in the spoken language development of children with severe to profound hearing loss. *J. Speech Lang. Hear. Res.* 50, 1048–1062. doi: 10.1044/1092-4388(2007/073)

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N. Y., Quittner, A. L., et al. (2010). Spoken language development in children following cochlear implantation. *JAMA* 303, 1498–1506. doi: 10.1001/jama.2010.451

Pandya, D. N., and Kuypers, H. (1969). Cortico-cortical connections in the rhesus monkey. *Brain Res.* 13, 13–36. doi: 10.1016/0006-8993(69)90141-3

Pascual-Marqui, R. D. (2002). Standardized low resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find. Exp. Clin. Pharmacol.* 24(Suppl. D), 5–12.

Pihko, E., Nevalainen, P., Stephen, J., Okada, Y., and Lauronen, L. (2009). Maturation of somatosensory cortical processing from birth to adulthood revealed by magnetoencephalography. *Clin. Neurophysiol.* 120, 1552–1561. doi: 10.1016/j.clinph.2009.05.028

Placzek, M., Mushin, J., and Dubowitz, L. M. S. (1985). Maturation of the visual evoked response and its correlation with visual acuity in preterm infants. *Dev. Med. Child Neurol.* 27, 448–454. doi: 10.1111/j.1469-8749.1985.tb04567.x

Ponton, C., Eggermont, J., Kwong, B., and Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked

potentials. *Clin. Neurophysiol.* 111, 220–236. doi: 10.1016/s1388-2457(99)00236-9

Rebillard, G., Carlier, E., Rebillard, M., and Pujol, R. (1977). Enhancement of visual responses on the primary auditory cortex of the cat after an early destruction of cochlear receptors. *Brain Res.* 129, 162–164. doi: 10.1016/0006-8993(77)90980-5

Russo, F. A., Ammirante, P., and Fels, D. I. (2012). Vibrotactile discrimination of musical timbre. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 822–826. doi: 10.1037/a0029046

Sadato, N. (2005). How the blind "see" Braille: lessons from functional magnetic resonance imaging. *Neuroscientist* 11, 577–582. doi: 10.1177/1073858405277314

Sandmann, P., Dillier, N., Eichele, T., Meyer, M., Kegel, A., Pascual-Marqui, R. D., et al. (2012). Visual activation of auditory cortex reflects maladaptive plasticity in cochlear implant users. *Brain* 135, 555–568. doi: 10.1093/brain/awr329

Sarant, J., Harris, D., Bennet, L., and Bant, S. (2014). Bilateral versus unilateral cochlear implants in children: a study of spoken language outcomes. *Ear Hear.* 35, 396–409. doi: 10.1097/AUD.0000000000000022

Sarant, J. Z., Blamey, P. J., Dowell, R. C., Clark, G. M., and Gibson, W. P. R. (2001). Variation in speech perception scores among children with cochlear implants. *Ear Hear.* 22, 18–28. doi: 10.1097/00003446-200102000-00003

Schroeder, C., Lindsley, R., Specht, C., Marcovici, A., Smiley, J. F., and Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.* 8, 1322–1327. doi: 10.1152/jn.2001.85.3.1322

Scott, G. D., Karns, C. M., Dow, M. W., Stevens, C., and Neville, H. (2014). Enhanced peripheral visual processing in congenitally deaf humans is supported by multiple brain regions, including primary auditory cortex. *Front. Hum. Neurosci.* 8:177. doi: 10.3389/fnhum.2014.00177

Sharma, A., Campbell, J., and Cardon, G. (2015). Developmental and cross-modal plasticity in deafness: evidence from the P1 and N1 event related potentials in cochlear implanted children. *Int. J. Psychophysiol.* 95, 135–144. doi: 10.1016/j.ijpsycho.2014.04.007

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002a). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539. doi: 10.1097/00003446-200212000-00004

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002b). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *Neuroreport* 13, 1365–1368. doi: 10.1097/00001756-200207190-00030

Sharma, A., Gilley, P. M., Dorman, M. F., and Baldwin, R. (2007). Deprivation-induced cortical reorganization in children with cochlear implants. *Int. J. Audiol.* 46, 494–499. doi: 10.1080/14992020701524836

Sharma, A., and Glick, H. (2016). Cross-modal re-organization in clinical populations with hearing loss. *Brain Sci.* 6:E4. doi: 10.3390/brainsci6010004

Sharma, A., Glick, H., Campbell, J., Torres, J., Dorman, M., and Zeitler, D. M. (2016). Cortical plasticity and reorganization in pediatric single-sided deafness pre- and postcochlear implantation: a case study. *Otol. Neurotol.* 37, e26–e34. doi: 10.1097/MAO.0000000000000904

Sharma, A., Kraus, N., McGee, T. J., and Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalogr. Clin. Neurophysiol.* 104, 540–545. doi: 10.1016/s0168-5597(97)00050-6

Shore, S. E., Koehler, S., Oldakowski, M., Hughes, L. F., and Syed, S. (2008). Dorsal cochlear nucleus responses to somatosensory stimulation are enhanced after noise-induced hearing loss. *Eur. J. Neurosci.* 27, 155–168. doi: 10.1111/j.1460-9568.2007.05983.x

Sitzoglou, C., and Fotiou, F. (1985). A study of the maturation of the somatosensory pathway by evoked potentials. *Neuropediatrics* 43, 205–208. doi: 10.1055/s-2008-1059538

Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., and Small, S. L. (2007). Hearing lips and seeing voices: how cortical areas supporting speech production mediate audiovisual speech perception. *Cereb. Coexrt* 17, 2387–2399. doi: 10.1093/cercor/bhl147

Soto-Faraco, S., and Deco, G. (2009). Multisensory contributions to the perception of vibrotactile events. *Behav. Brain Res.* 196, 145–154. doi: 10.1016/j.bbr.2008.09.018

Svirsky, M. A., Robbins, A. M., Kirk, K. I., Pisoni, D. B., and Miyamoto, R. T. (2000). Language development in profoundly deaf children with cochlear implants. *Psychol. Sci.* 11, 153–158. doi: 10.1111/1467-9280.00231

Svirsky, M. A., Teoh, S.-W., and Neuburger, H. (2004). Development of language and speech perception in congenitally, profoundly deaf children as a function of age at cochlear implantation. *Audiol. Neurootol.* 9, 224–233. doi: 10.1159/000078392

Szagun, G., and Schramm, S. A. (2016). Sources of variability in language development of children with cochlear implants: age at implantation, parental language, and early features of children's language construction. *J. Child Lang.* 43, 505–536. doi: 10.1017/S0305000915000641

The MathWorks (2014). *MatLab (Version R2014b) [Software]*. Natick, MA: The MathWorks, Inc.

Tobey, E. A., Geers, A. E., Brenner, C., Altuna, D., and Gabbert, G. (2003). Factors associated with development of speech production skills in children implanted by age five. *Ear Hear.* 24(Suppl. 1), 36S–45S. doi: 10.1097/01.aud.0000051688.48224.a6

Watkins, K. E., Strafella, A. P., and Paus, T. (2003). Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41, 989–994. doi: 10.1016/s0028-3932(02)00316-0

Weinstein, S. (1968). "Intensive and extensive aspects of tactile sensitivity as a function of body part, sex, and laterality," in *The Skin Senses*, ed. D. R. Kenshalo (Springfield, IL: Thomas), 195–222.

Wilke, M., Krägeloh-Mann, I., and Holland, S. K. (2006). Global and local development of gray and white matter volume in normal children and adolescents. *Exp. Brain Res.* 178, 296–307. doi: 10.1007/s00221-006-0732-z

Yamaguchi, S., and Knight, R. T. (1991). Age effects on the P300 to novel somatosensory stimuli. *Electroencephalogr. Clin. Neurophysiol.* 78, 297–301. doi: 10.1016/0013-4694(91)90184-6

Zeng, C., Yang, Z., Shreve, L., Bledsoe, S., and Shore, S. (2012). Somatosensory projections to cochlear nucleus are upregulated after unilateral deafness. *J. Neurosci.* 32, 15791–15801. doi: 10.1523/JNEUROSCI.2598-12.2012

Check for
updates

The Contribution of Bilingualism, Parental Education, and School Characteristics to Performance on the Clinical Evaluation of Language Fundamentals: Fourth Edition, Swedish

Ketty Andersson[1], Kristina Hansson[1], Ida Rosqvist[1], Viveka Lyberg Åhlander[1], Birgitta Sahlén[1] and Olof Sandgren[1,2]*

[1]Department of Clinical Sciences Lund, Logopedics, Phoniatrics, and Audiology, Faculty of Medicine, Lund University, Lund, Sweden, [2]Department of School Development and Leadership, Faculty of Education and Society, Malmö University, Malmö, Sweden

Assessment of bilingual children in only one language fails to acknowledge their distributed linguistic competence and has been shown to overidentify language disorder in bilingual populations. However, other factors, sometimes associated with bilingualism, may also contribute to low results in language assessments. Our aim was to examine the impact of these factors on language abilities. We used the Clinical Evaluation of Language Fundamentals – Fourth Edition, Swedish (CELF-4) to investigate core language abilities of 224 7- to 8-year-old children. Results showed 30 and 80% of monolinguals and bilinguals, respectively, performing more than 1 SD below the normative sample mean, calling into question the clinical utility of the test. However, participant and school characteristics provided a deeper understanding of the skewed results. In isolation, bilingualism predicted 38% of the variance in the CELF-4 Core scores. With level of parental education entered the variance explained by the model increased to 52%, but the unique contribution of bilingualism was reduced to 20%. Finally, with information added on school characteristics and enrollment in the school's recreation center the model explained an additional two percent, with the unique contribution of bilingualism further reduced to 9%. The results indicate an increased risk for low results on the CELF-4 Core when children present with multiple risk factors. This highlights the need to look beyond bilingualism in language assessment of bilingual children and adolescents and to consider other explanations to academic struggle. Available interventions must be considered and applied proportionately to their respective impact on the individual's development.

Keywords: language assessment, bilingualism, academic achievement, language exposure, language disorder

INTRODUCTION

Large-scale international comparisons have reported a decline in Swedish primary and elementary school students' academic attainment the last 20 years, as compared to peers in other countries. Swedish 15-year olds' skills and knowledge in reading, mathematics, and science, as assessed in the Program for International Student Assessment (PISA), have steadily declined from 2000.

The scores reached an all-time low in 2012 (OECD, 2014) and returned to the OECD mean in 2015 (OECD, 2016). Similar developmental trends have been shown for fourth grade reading comprehension [Progress in International Reading Literacy Study (PIRLS 2016)] and fourth and eighth grade knowledge in mathematics and science [Trends in International Mathematics and Science Study (TIMSS 2015)].

The most recent report from the OECD indicates that the negative trend among Swedish students may be reversed, or at least halted (OECD, 2016). However, the results have caused great concern in the general public. Some policy makers have used the ensuing discussion to make ideologically based claims on necessary changes to the school curriculum.

Research-based analyses have offered several explanations to the declining results. Some question the validity (Brunner et al., 2007) and reliability (Goldstein, 2004) of the large-scale assessments of student performance used in international comparisons. Declining results have been linked to factors such as less able teachers (Meroni et al., 2015), low teacher expectations on student progress (Wang et al., 2018), increased use of computers and handheld devices in classrooms (OECD, 2015), and lack of large-scale funding and coordination of systematic evaluations of educational practices (Pontoppidan et al., 2018). In a Swedish context, high levels of autonomy for school districts have led to greater differences between schools (Swedish National Agency for Education, 2006). Globally, rapid changes in the student cohort demographics, from largely monolingual to bilingual, have been presented as a main (Agirdag and Vanlaar, 2018) or a contributing (OECD, 2018b) factor.

Indeed, many Western countries have seen an increased number of bilinguals (OECD, 2018a), and language is of crucial importance for school success (Pace et al., 2019). Mainstream teaching requires students to be fluent in the majority language not only to follow the teacher's instructions and to participate in the teaching activities but also to access the hidden curriculum, which guides school culture (Baker, 2011). Language, in particular vocabulary and listening comprehension, has repeatedly been shown to be at the core of these competencies. The vocabulary used in classrooms differs greatly from that used in everyday conversation, with more abstract words (Cummins, 1979), and bilingual children need support to keep pace with monolingual peers linguistically and academically. Although bilingual children tend to develop their vocabulary knowledge at the same pace as monolingual children, the gap between the groups remains because bilingual children have a lower starting point and monolinguals gradually improve and, thereby, constitute a continuously moving target (Thordardottir and Juliusdottir, 2013). In order to make similar academic and language achievements as monolinguals, a high oral language proficiency is required (Babayiğit, 2015). Indeed, O'Connor et al. (2018) found no difference between mono- and bilingual children regarding literacy and numeracy when the bilingual children had a high proficiency in receptive English vocabulary. However, bilingual children with less developed receptive vocabulary skills had difficulties meeting the school demands.

However, several factors have impact on language development, only one of which is bilingualism. Children, whether Monolingual and bilingual children alike, living in less affluent areas risk facing "a double dose of disadvantage," experiencing impoverished language input at both home and school (Neuman et al., 2018, see also Hoff, 2013). Neuman et al. (2018) found parents in socioeconomically disadvantaged areas to use fewer words and shorter, less complex sentences when interacting with their children, than parents in a more diverse working-class comparison neighborhood. In addition, Neuman et al. (2018) found similar differences in conversation in school; teachers in the less affluent neighborhood used less varied vocabulary and less complex syntax than the teachers in the working-class neighborhood, thereby failing to provide compensation for the limited language input in these children's home environments. With a reduced experience of interaction with adults who use complex syntax and vocabulary, known to enhance children's expressive language development, these children start school with less robust language experiences, which, in turn, increase the risk of school failure (Neuman et al., 2018).

Agirdag and Vanlaar (2018) criticize dichotomous categories of bilingualism, as used by, for example, OECD, and point to the need for evaluation of language exposure and use, as more reliable predictors of academic outcomes. The authors compared two competing views on bilingualism; the time-on-task perspective, which predicts a monolingual advantage in outcome; and the additive perspective on multilingualism, which predicts that transfer and switching between languages will have positive cognitive and linguistic effects and hence a bilingual advantage. Agirdag and Vanlaar (2018) failed to show a bilingual advantage. Bilinguals showed lower achievements in reading and mathematics than monolingual peers. Taking school and student characteristics into account reduced the achievement gap between monolingual and bilingual children, but not to an insignificant level. However, an in-depth analysis of the language exposure, taking into account which language the child used in different conversational contexts, provided more information. Bilingual children who regularly used the home language with their parents achieved in level with monolingual peers. In addition, speaking the majority language with friends was positively associated with academic achievements (Agirdag and Vanlaar, 2018). Similarly, Huang et al. (2018) found that using the second language in the spare time was more influential on language comprehension than using the language in the classroom. Thus, bilingual children who receive high-quality input in their mother tongue and who can use their school language for everyday conversation with friends are likely to perform at the same level as monolingual peers. In fact, Agirdag and Vanlaar (2018) were able to show that these children may even outperform monolingual peers in societies with a positive view on bilingualism.

The use of monolingual language norms and expectations may lead to an overidentification of language problems in bilingual populations (Lugo-Neris et al., 2015). Children who acquire a second language can sometimes be hard to distinguish from children with developmental language disorder (DLD), with both groups presenting with similar language profiles, at least at some point in development (Salameh et al., 1996; Windsor and Kohnert, 2004). The importance of assessing both

languages has long been stressed, but even when assessed in both languages, or in their strongest language, bilingual children from disadvantaged socioeconomic backgrounds are overidentified as having DLD (Barragan et al., 2018). In a sample of Spanish-English dual language learners, Barragan et al. (2018) found more than 50% to perform more than 1 SD below the mean on the Spanish version of the Clinical Evaluation of Language Fundamentals – Fourth Edition (CELF-4), that is, exceeding the recommended cut-off score for language disorder. The older children were more likely to show low performance on the expressive subtests (Recalling Sentences and Formulated Sentences) than the younger children, indicating a shift of language dominance at this age (Kohnert and Bates, 2002; Barragan et al., 2018). Norm-referenced tests risk overidentifying children from low-socioeconomic backgrounds, and separate norms may be necessary to improve the sensitivity and specificity of language assessments.

To sum up, bilingualism *per se* is not detrimental to children's language outcomes and academic achievements. However, a number of factors, associated with increased risk of academic underachievement, may accumulate in bilingual children. We aim to disentangle the relative contributions of bilingualism, socio-economic disadvantage, and suboptimal language exposure and use on core language abilities.

PURPOSE

The purpose of this study was to estimate the impact of bilingualism on CELF-4 Core scores in isolation and in combination with information on level of parental education, school characteristics (proportion of parents with tertiary education and proportion of students with Swedish as second language), and recreation center enrollment. We answer two questions:

1. How do monolingual and bilingual children perform on the Swedish CELF-4 Core?
2. How much unique and combined variance in CELF-4 Core scores can be attributed to bilingualism, level of parental education, school characteristics, and enrollment in the school's recreation center?

MATERIALS AND METHODS

Participants

CELF-4 Core scores were collected from 224 7- to 8-year-old children (M_{age} 90.8, SD 7.3, range 77–105 months), representing 57% of the students in first and second grade in six invited public schools from two school districts. The participants received education in Swedish with the exception of a weekly lesson of first language teaching for bilingual children, if requested by the parents (on a national level requested for 60% of eligible children; Swedish National Agency for Education, 2019). No preselection of participants was made on the basis of language risk or special education needs. The sample was

representative of the student cohort regarding the proportion of mono- and bilingual participants [$t(223) = 1.58$, $p = 0.12$]. The distribution of boys and girls (120 girls and 104 boys) did not differ significantly [$\chi^2(1) = 1.14$, $p = 0.29$].

The parents of all participants provided information on level of parental education, children's bilingualism status, and children's enrollment in the school's recreation center activities after school hours. Additional school characteristics (proportion of parents with tertiary education and proportion of students with Swedish as second language) were compiled from publicly available statistical data (Swedish National Agency for Education, 2019).

Assessment

Participating children were assessed with the Swedish version of the CELF-4 (Semel et al., 2013). Four subtests compose a core language score used as a screening in clinical decision-making. The subtest Concepts and Following Directions requires the child to point to pictures following increasingly complex oral instructions from the examiner. Word Structure assesses morphological ability in a sentence completion format, where the child is required to mark noun, verb, and adjective inflections. In Recalling Sentences, the task is to give a verbatim repetition of a sentence, without modifications. In Formulated Sentences, the child freely formulates a sentence appropriate to a picture stimulus, including a target word provided by the examiner.

Procedure

The study was carried out in accordance with the recommendations of the Ethics Review Board of Southern Sweden (approval number 2016/567) with written informed consent from the parents of all participants, in accordance with the Declaration of Helsinki. The teachers in participating schools and classrooms distributed parent consent forms. Parents who approved their child's participation filled out a form with information on language exposure and use, level of parental education, previous speech-language pathology (SLP) or special education services provided for the child, and enrollment in the school's recreation center activities after school hours. All examiners were native Swedish-speaking SLPs or final year SLP students specially trained for the purpose of the data collection. All testing was conducted during school hours in rooms adjacent to the child's classroom. The testing took approximately 40 min. The subtests were administered in a fixed sequence, and all verbal instructions were scripted, in order to reduce the risk of inter- and intra-rater inconsistencies.

Statistical Analyses

In accordance with the test manual, the raw scores from the subtests were converted to subscale scores with a mean of 10 and a SD of 3. The subscale scores were collapsed to form a core language score with a mean of 100 and a SD of 15, to allow comparison with the normative sample of the CELF-4.

From the sample of 224 participants, complete data on bilingualism status, level of parental education, and enrollment

TABLE 1 | Participants' demographic information.

Demographics		*n*	
School district	A	B	Total
Grade			
First	52	48	100
Second	68	56	124
Total	120	104	224
Level of parental education			
≤9 years (elementary school)	22	2	24
12 years (high school)	42	6	48
>12 years (university)	34	92	126
Missing	22	4	26
Bilingual			
Yes	87	19	106
No	33	85	118
Missing	0	0	0
Recreation center enrollment			
Yes	50	97	147
No	34	6	40
Missing	36	1	37
School characteristics			
Parents with tertiary education (%)	45[a] (56)	80[b] (57)	
Students with Swedish as second language (%)	43[a] (24)	25[b] (25)	

Missing = information missing in parental report.
[a]*School year 2017–2018. National averages in parentheses.*
[b]*School year 2018–2019. National averages in parentheses.*

in the school's recreation center were obtained for 170 participants (see **Table 1**). CELF-4 Core scores were obtained for 222 participants, with two children failing to participate in one of the CELF-4 subtests.

Publicly available data on the proportion of parents with tertiary education and proportion of students with Swedish as first language in the participating schools were ranked from lowest (1) to highest (6). The rank scores were summed to form an index of school characteristics (possible range 2–12). The highest index score was assigned to the school with highest proportion of parents with tertiary education and students with Swedish as first language.

Hierarchical regression was used to investigate the effect of the independent variables on the CELF-4 Core scores. Bilingualism was entered first into the model. In a second step, level of parental education was added to calculate the effect above and beyond that of bilingualism. In a final step, the index of school characteristics and enrollment in the school's recreation center were added to the regression model. Preliminary analyses ensured all assumptions of normality, linearity, multicollinearity, and homoscedasticity were met. All statistical analyses were performed using SPSS version 25 for Windows.

RESULTS

On the CELF-4 Core, the mean score for the sample was 77.99 (SD = 23.93, range 40–122), which is almost 1.5 SDs below the normative sample of the test. On a group level, monolingual

participants (n = 118) performed within the expected range (M = 91.81, SD = 16.8, range 49–122), whereas the bilingual participants (n = 104) performed below the normative range (M = 62.31, SD = 20.99, range 40–114). Although the participants, on an individual level, were not preselected on the basis of language or academic risk, 30% of monolingual participants, and 80% bilingual participants, performed more than 1 SD below the mean (≤85) on the CELF-4 Core index, the recommended cut-off score for language disorder.

Table 2 shows correlations between the dependent variable (CELF-4 Core) and the independent variables (Bilingualism, Level of parental education, School characteristics, Recreation center enrollment). All correlations were significant, indicating associations between the variables. To further explore the unique and shared variance in CELF-4 Core scores explained by the independent variables, all variables were entered into a hierarchical regression model (see **Table 3**). In the first model, bilingualism was entered as a single predictor, accounting for 38% of the variance in CELF-4 Core scores, $F(1, 171) = 104.96$, $p < 0.001$. In Model 2, level of parental education was added, increasing the proportion of explained variance to 52%, $F(2, 170) = 91.22$, $p < 0.001$. The unique variance explained by bilingualism decreased to 20%, while level of parental education explained 11.5% unique variance. Thus, the shared variance of bilingualism and level of parental education, as expected from the correlations in **Table 2**, are greater than or equal to the unique contribution. In the final model, school characteristics and enrollment in the school's recreation center were entered, explaining an additional 2% of the variance in CELF-4 Core scores, $F(4, 168) = 49.10$, $p < 0.001$. Again, the unique contribution of bilingualism decreased, to 9%, an indication of the overlapping multifactorial influence of the independent variables.

DISCUSSION

For Swedish speech-language pathologists, the CELF-4 represents one of few norm-referenced standardized language assessments. Consequently, clinicians rely heavily on the results from CELF-4 assessments when making diagnostic decisions. The purpose of this study was to examine how monolingual and bilingual participants perform on the CELF-4 Core. We report unexpectedly low results, with 30% of monolingual participants scoring below the recommended screening cut-off score for language disorder, according to the test manual (Semel et al., 2013), despite similar

TABLE 2 | Correlations between CELF-4 Core, bilingualism, level of parental education, school characteristics, and enrollment in the school's recreation center.

	1	2	3	4
1. CELF-4 Core				
2. Bilingualism	−0.62***			
3. Level of parental education	0.56***	−0.35***		
4. School characteristics	0.59***	−0.57***	0.53***	
5. Recreation center enrollment	0.41***	−0.44***	0.37***	0.41***

***$p < 0.001$.

TABLE 3 | Hierarchical regression model predicting CELF-4 Core scores.

Measure	Model 1			Model 2			Model 3		
	$R^2 = 0.38$			$R^2 = 0.52$			$R^2 = 0.54$		
				$\Delta R^2 = 0.14$			$\Delta R^2 = 0.02$		
	B	SEB	β	*B*	SEB	β	*B*	SEB	β
Bilingualism	−29.49	2.88	−0.62***	−22.97	2.71	−0.48***	−18.30	3.17	−0.38***
Level of parental education				13.43	1.93	0.40***	10.71	2.14	0.32***
School characteristics							1.68	0.66	0.18***
Recreation center enrollment							2.80	3.53	0.05

*** $p < 0.001$.

prerequisites as the normative sample. The relative size of the samples may be one possible explanation. The study sample is greater than the CELF-4 normative sample for 7- to 8-year olds. The results of the bilingual participants, who were excluded from normative sample, were less surprising, but equally alarming, with 80% scoring below the cut-off score for language disorder.

The background variables offer a deeper understanding of the results. Bilingualism, explaining 38% of the variance in CELF-4 Core scores when analyzed separately, loses most of its predictive force when taking socioeconomic and school factors into account. The hierarchical regression model reveals high levels of shared variance between bilingualism, level of parental education, school characteristics, and enrollment in the school's recreation center. For children who exhibit more than one risk factor, the effect is detrimental. Participants who speak Swedish as a second language, come from socioeconomically challenged home environments and who attend schools where many students share these circumstances are at an increased risk of low results on the CELF-4 Core, and, as a consequence, of being misidentified as having a language disorder. Although language support is required regardless the cause, children who experience suboptimal language learning conditions are likely to gain more from focused instruction on vocabulary and reading comprehension (Spencer et al., 2017). Individuals with language disorder, and the people around them, will also need to be equipped with compensatory strategies in order to be able to make necessary everyday adjustments to prevent the risk of language and communication breakdowns (Ebbels et al., 2019).

When analyzed in combination with socioeconomic and school factors, bilingualism only explained 9% of the CELF-4 Core scores. Consequently, separate norms for bilingual children or children from different socioeconomic circumstances would not provide a satisfactory solution nor would norm-referenced assessment in the first language. This would, in most cases, mean that the bilingual child once again is compared with a monolingual normative sample (for a discussion, see Scheidnes and Tuller, 2016). Instead, other types of language assessments should be more generally practiced. Dynamic assessment, focusing on the potential for language learning rather than providing a static assessment at one point in time, is one example (Hasson et al., 2013; Dockrel et al., 2015). Processing measures of language proficiency, for example, non-word repetition, have also shown higher sensitivity and specificity in bilingual populations than traditional language assessment (Thordardottir and Brandeker, 2013).

What, then, would increase the diagnostic accuracy in assessments of bilingual children? First, all assessments must take into account available demographic information, for example, level of parental education. Similar to the results presented here, Barragan et al. (2018) found more than 50% of children from low-income, bilingual backgrounds to perform in level with children with language disorder on CELF-4 assessments. We show the same applies to a high proportion of monolingual children with the same background.

Second, language assessments must evaluate bilingual children's opportunities to use their second language in different contexts and with different conversational partners. Agirdag and Vanlaar (2018) demonstrate the positive effect on academic results of speaking the second language with schoolmates and friends. Information on second language use with peers outside school hours, although not contributing significantly to the model as measured with enrollment in the school's recreation center, should be further investigated.

Third, schools are required to face the challenge of providing equitable education services to students of different language and socioeconomic backgrounds to make the curriculum content accessible for all. This calls for school environments with clearly defined areas within the classroom for different teaching activities, and high-quality teaching methods, for example, interactive book reading, structured conversations and targeted feedback, in order for the school hours to be used optimally (Dockrell et al., 2015). With these measures, all students will have a better chance of performing to their capacity within classrooms that accept and invite all voices and languages of the students to be heard (Rolstad et al., 2005).

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

REFERENCES

Agirdag, O., and Vanlaar, G. (2018). Does more exposure to the language of instruction lead to higher academic achievement? A cross-national examination. *Int. J. Biling.* 22, 123–137. doi: 10.1177/1367006916658711

Babayiğit, S. (2015). The relations between word reading, oral language, and reading comprehension in children who speak English as a first (L1) and second language (L2): a multigroup structural analysis. *Read. Writ.* 28, 527–544. doi: 10.1007/s11145-014-9536-x

Baker, C. (2011). *Foundations of bilingual education and bilingualism.* 5th Edn. Bristol, UK; Tonawanda, NY: Multilingual Matters.

Barragan, B., Castilla-Earls, A., Martinez-Nieto, L., Restrepo, M. A., and Gray, S. (2018). Performance of low-income dual language learners attending English-only schools on the clinical evaluation of language fundamentals – Fourth edition, Spanish. *Lang. Speech Hear. Serv. Sch.* 49, 292–305. doi: 10.1044/2017_LSHSS-17-0013

Brunner, M., Artelt, C., Krauss, S., and Baumert, J. (2007). Coaching for the PISA test. *Learn. Instr.* 17, 111–122. doi: 10.1016/j.learninstruc.2007.01.002

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Work. Papers Bilingual.* 19, 1–9.

Dockrell, J. E., Bakopoulou, I., Law, J., Spencer, S., and Lindsay, G. (2015). Capturing communication supporting classrooms: The development of a tool and feasibility study. *Child Lang. Teach. Ther.* 31, 271–286. doi: 10.1177/0265659015572165

Ebbels, S. H., McCartney, E., Slonims, V., Dockrell, J. E., and Norbury, C. F. (2019). Evidence-based pathways to intervention for children with language disorders. *Int. J. Lang. Comm. Disord.* 54, 3–19. doi: 10.1111/1460-6984.12387

Goldstein, H. (2004). International comparisons of student attainment: some issues arising from the PISA study. *Assess. Educ. Princ. Pol. Pract.* 11, 319–330. doi: 10.1080/0969594042000304618

Hasson, N., Camilleri, B., Jones, C., Smith, J., and Dodd, B. (2013). Discriminating disorder from difference using dynamic assessment with bilingual children. *Child Lang. Teach. Ther.* 29, 57–75. doi: 10.1177/0265659012459526

Hoff, E. (2013). Interpreting the early language trajectories of children from low-SES and language minority homes: implications for closing achievement gaps. *Dev. Psychol.* 49, 4–14. doi: 10.1037/a0027238

Huang, B. H., Chang, Y.-H. S., Zhi, M., and Niu, L. (2018). The effect of input on bilingual adolescents' long-term language outcomes in a foreign language instruction context. *Int. J. Biling.* doi: 10.1177/1367006918768311

Kohnert, K., and Bates, E. (2002). Balancing bilinguals II: lexical comprehension and cognitive processing in children learning Spanish and English. *J. Speech Lang. Hear. Res.* 45, 347–359. doi: 10.1044/1092-4388(2002/027)

Lugo-Neris, M. J., Peña, E. D., Bedore, L. M., and Gillam, R. B. (2015). Utility of a language screening measure for predicting risk for language impairment in bilinguals. *Am. J. Speech Lang. Pathol.* 24, 426–437. doi: 10.1044/2015_AJSLP-14-0061

Meroni, E. C., Vera-Toscano, E., and Costa, P. (2015). Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. *J. Policy Model* 37, 308–323. doi: 10.1016/j.jpolmod.2015.02.006

Neuman, S. B., Kaefer, T., and Pinkham, A. M. (2018). A double dose of disadvantage: language experiences for low-income children in home and school. *J. Educ. Psychol.* 110, 102–118. doi: 10.1037/edu0000201

O'Connor, M., O'Connor, E., Tarasuik, J., Gray, S., Kvalsvig, A., and Goldfeld, S. (2018). Academic outcomes of multilingual children in Australia. *Int. J. Speech Lang. Pathol.* 20, 393–405. doi: 10.1080/17549507.2017.1292546

OECD (2014). *PISA 2012 results: What students know and can do – student performance in mathematics, reading and science.* Vol. 1, Paris: OECD Publishing.

OECD (2015). *Students, computers and learning.* Paris: OECD Publishing.

OECD (2016). *PISA 2015 results.* Vol. I, Paris: OECD Publishing.

OECD (2018a). *International migration outlook 2018.* Paris: OECD Publishing.

OECD (2018b). *The resilience of students with an immigrant background.* Paris: OECD Publishing.

Pace, A., Alper, R., Burchinal, M. R., Golinkoff, R. M., and Hirsh-Pasek, K. (2019). Measuring success: within and cross-domain predictors of academic and social trajectories in elementary school. *Early Child. Res. Q.* 46, 112–125. doi: 10.1016/j.ecresq.2018.04.001

PIRLS (2016). *What makes a good reader: International findings from PIRLS 2016.* Boston: TIMSS & PIRLS International Study Center.

Pontoppidan, M., Keilow, M., Dietrichson, J., Solheim, O. J., Opheim, V., Gustafson, S., et al. (2018). Randomised controlled trials in Scandinavian educational research. *Educ. Res.* 60, 311–335. doi: 10.1080/00131881.2018.1493351

Rolstad, K., Mahoney, K., and Glass, G. V. (2005). The big picture: a meta-analysis of program effectiveness research on English language learners. *Educ. Pol.* 19, 572–594. doi: 10.1177/0895904805278067

Salameh, E.-K., Håkansson, G., and Nettelbladt, U. (1996). The acquisition of Swedish as second language in a group of arabic-speaking pre-school children: word order patterns and phrasal morphology. *Logoped. Phoniatr. Vocol.* 21, 163–170.

Scheidnes, M., and Tuller, L. (2016). Assessing successive bilinguals in two languages: a longitudinal look at English-speaking children in France. *J. Commun. Disord.* 64, 45–61. doi: 10.1016/j.jcomdis.2016.10.001

Semel, E., Wiig, E. H., and Secord, W. A. (2013). *Clinical evaluation of language fundamentals, fourth edition (CELF-4), Swedish version.* Stockholm, Sweden: Pearson Assessment.

Spencer, S., Clegg, J., Lowe, H., and Stackhouse, J. (2017). Increasing adolescents' depth of understanding of cross-curriculum words: an intervention study. *Int. J. Lang. Comm. Disord.* 52, 652–668. doi: 10.1111/1460-6984.12309

Swedish National Agency for Education (2006). *Equity trends in the Swedish school system – a quantitative analysis of variation in student performance and equity from a time perspective.* Stockholm: Skolverket.

Swedish National Agency for Education. (2019). Information on pupil characteristics. Available at: https://www.skolverket.se/skolutveckling/statistik (Accessed April 4, 2019).

Thordardottir, E., and Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *J. Commun. Disord.* 46, 1–16. doi: 10.1016/j.jcomdis.2012.08.002

Thordardottir, E., and Juliusdottir, A. G. (2013). Icelandic as a second language: a longitudinal study of language knowledge and processing by school-age children. *Int. J. Biling. Educ. Biling.* 16, 411–435. doi: 10.1080/13670050.2012.693062

TIMSS (2015). *TIMSS 2015 international results report*. Boston: TIMSS & PIRLS International Study Center.

Wang, S., Rubie-Davies, C. M., and Meissel, K. (2018). A systematic review of the teacher expectation literature over the past 30 years. *Educ. Res. Eval.* 24, 124–179. doi: 10.1080/13803611.2018.1548798

Windsor, J., and Kohnert, K. (2004). The search for common ground: part I. Lexical performance by linguistically diverse learners. *J. Speech Lang. Hear. Res.* 47, 877–890. doi: 10.1044/1092-4388(2004/065)

Check for updates

Moderate Reverberation Does Not Increase Subjective Fatigue, Subjective Listening Effort, or Behavioral Listening Effort in School-Aged Children

Erin M. Picou[1,2]*, Brianna Bean[1,2], Steven C. Marcrum[3], Todd A. Ricketts[2] and Benjamin W. Y. Hornsby[4]

[1] Hearing and Affect Perception Interest Laboratory, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States, [2] Dan Maddox Hearing Aid Research Laboratory, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States, [3] Department of Otolaryngology, University Hospital Regensburg, Regensburg, Germany, [4] Hearing and Communication Laboratory, Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, Nashville, TN, United States

Background noise and reverberation levels in typical classrooms have negative effects on speech recognition, but their effects on listening effort and fatigue are less well understood. Based on the Framework for Understanding Effortful Listening, noise and reverberation would be expected to increase both listening effort and fatigue. However, previous investigations of the effects of reverberation for adults have resulted in mixed findings. Some discrepancies in the literature might be accounted for by methodological differences; behavioral and subjective indices of listening effort do not often align in adults. The effects of sustained listening on self-reported fatigue in school-aged children are also not well understood. The purposes of this project were to (1) evaluate the effects of noise and reverberation on listening effort in school-aged children using behavioral and subjective measures, (2) compare subjective and behavioral indices of listening effort, and (3) evaluate the effects of reverberation on self-reported fatigue. Twenty typically developing children (10–17 years old) participated. Participants completed dual-task testing in two rooms that varied in terms of reverberation, an audiometric sound booth and a moderately reverberant room. In each room, testing was completed in quiet and in two levels of background noise. Participants provided subjective ratings of listening effort after completing the dual-task in each listening condition. Subjective ratings of fatigue were completed before and after testing in each level of reverberation. Results revealed background noise, not reverberation, increased behavioral and subjective listening effort. Subjective ratings of perceived performance, ease of listening, and desire to control the listening situation revealed a similar pattern of results as word recognition performance, making them poor candidates for providing an indication of behavioral listening effort. However, ratings of time perception were moderately correlated with behavioral listening effort. Finally, sustained listening for

approximately 25 min increased self-reported fatigue, although changes in fatigue were comparable in low and moderately reverberant environments. In total, these data offer no evidence that a moderate level of reverberation increases listening effort or fatigue, but the data do support the reduction of background noise in classrooms.

INTRODUCTION

For school-aged children, listening in classrooms can be challenging. Typical classroom environments are acoustically disadvantaged with signal-to-noise ratios (SNRs) ranging from −6 to +13 dB (Pearsons et al., 1977; Bradley and Sato, 2008; Sato and Bradley, 2008), whereas ideal SNRs for classrooms are considerably more favorable (e.g., +15 to +30; Berg, 1993; Bistafa and Bradley, 2000; Crandell and Smaldino, 2000). In addition, typical classrooms are likely to be more reverberant than is recommended, with measured classroom reverberation times of 600 ms (Crandell and Smaldino, 1994; Crukley et al., 2011) to 1200 ms (Crandell and Smaldino, 1994), whereas reverberation times of 400 to 500 ms or less are recommended (Finitzo-Hieber and Tillman, 1978; Bistafa and Bradley, 2000).

The perceptual consequences of listening in acoustically disadvantaged environments include not only reduced speech recognition, but also increased listening effort (e.g., Prodi et al., 2010). "Listening effort" is defined as the "deliberate allocation of resources to overcome obstacles in goal pursuit" when listening (Pichora-Fuller et al., 2016, pg. 11S). Given the important, negative consequences of sustained increases in listening effort, such as communicative disengagement (Hétu et al., 1988), reduced vocational involvement (Kramer et al., 2006), and mental fatigue (Hornsby, 2013), it is important to understand the factors that affect listening effort.

The Ease of Language Understanding (ELU) model (Rönnberg et al., 2008, 2013) provides a framework for understanding listening effort. Briefly, the model suggests that a listener compares language inputs to long-term memory stores. Understanding is easy or effortless if the language input matches a long-term memory store. Conversely, if a match is not immediate, cognitive resources must be deployed to facilitate a match. The Framework for Understanding Effortful Listening (FUEL; Pichora-Fuller et al., 2016), based on the model of limited attention proposed by Kahneman (1973), extends the ELU model by including elements of executive function that control a resource allocation policy. Specifically, cognitive resources can be allocated automatically (e.g., in response to sudden stimuli), intentionally (e.g., with explicit instruction), or evaluatively (e.g., to attain a goal). Assuming the allocation is consistent across listening conditions, both the ELU and FUEL frameworks suggest that factors interfere with the input-memory match, such as background noise and reverberation, would increase listening effort.

Consistent with the hypothesis that background noise interferes with an input-memory match and thus requires deployment of cognitive resources, investigators have repeatedly demonstrated increased listening effort in adults with the addition, or increased level, of background noise. Effects of background noise on listening effort have been demonstrated with memory paradigms (Surprenant, 1999; Murphy et al., 2000; Picou et al., 2011), physiologic measures (Zekveld et al., 2010, 2011; Mackersie and Cones, 2011), and behavioral reaction-time measures (Sarampalis et al., 2009; Fraser et al., 2010; Picou et al., 2013).

In school-aged children, the results of studies into the effects of background noise on listening effort are less consistent. Using behavioral reaction-time tasks, some investigators have reported that SNR improvements (i.e., decreasing background noise levels) reduces listening effort (Prodi et al., 2010; Gustafson et al., 2014; Lewis et al., 2016; Hsu et al., 2017; McGarrigle et al., 2019); however, the finding is not universal (Hicks and Tharpe, 2002; Howard et al., 2010; McGarrigle et al., 2017, 2019). Some of the discrepancy between the published findings and ELU and FUEL predictions might be related to the sensitivity of the various listening effort paradigms. If a task is not motivating or is too distracting, changes in listening effort will be less evident (Choi et al., 2008), as might have been the case in earlier investigations of effort in school-aged children (e.g., McFadden and Pittman, 2008). When utilizing secondary tasks that are moderately challenging, investigators have found changes in behavioral effort with changes in SNR (Hsu et al., 2017; Picou et al., 2017a).

According to the ELU and FUEL frameworks, another transmission factor expected to increase listening effort, and relevant to contemporary classrooms, is room reverberation. Reverberation effects are generally described as either "early" or "late," based on their time of arrival to a listener's ear. Early reflections, or those that arrive within 0.05 s after direct signal presentation (Bradley, 1986; Bradley et al., 1999), are integrated with direct signal energy (Haas, 1972; Nábelek and Robinette, 1978). Late reflections, however, are not integrated with the direct signal energy and instead result in masking and temporal smearing of the original signal. As a result, late reflections reduce speech recognition performance, particularly in the middle part of the performance-intensity function (Nábelek and Pickett, 1974; Finitzo-Hieber and Tillman, 1978; Neuman et al., 2010; Wróblewski et al., 2012). Thus, one would also expect increased listening effort associated with reverberation.

However, the observed effects of reverberation on listening effort are unclear. For adults with normal hearing, several investigators have reported that increased levels of reverberation result in increased listening effort, as measured via subjective ratings with recorded stimuli (Sato et al., 2008, 2012; Rennies et al., 2014). However, other investigators using behavioral

paradigms have failed to demonstrate increased listening effort with moderate increases in reverberation (Picou et al., 2016; Peng and Wang, 2019). For example, Picou et al. (2016) found that increasing reverberation (from <100 to 475 or to 834 ms), did not increase listening effort for adults with normal hearing. Explanations for the non-significant effects of reverberation remain elusive. It is possible reverberation affects listening effort only for some acoustic conditions, such as with multiple, moving talkers (Valente et al., 2012) or with longer reverberation times (e.g., T30 > 900 ms). It is also possible listening difficulties associated with listening to distortions are fundamentally different than the listening difficulties associated with noise masking, as suggested by Francis et al. (2016).

Importantly, there is scarce literature reporting on the effects of reverberation on listening effort for school-aged children. In terms of speech recognition, children are more vulnerable to the effects of reverberation than adults (Klatte et al., 2010b; Neuman et al., 2010; Valente et al., 2012; Wróblewski et al., 2012). In addition, evidence from real classrooms demonstrates negative effects of longer reverberation times (1000 compared to <500 ms) on students' phonological processing, noise annoyance ratings, and teacher relationships (Klatte et al., 2010a). Thus, it is possible that reverberation could increase listening effort in school-aged children, despite non-significant behavioral findings in adults.

Alternatively, Amlani and Russo (2016) found that adding acoustic paneling to reduce reverberation in a classroom *increased* listening effort, as measured using a recall-based, dual-task paradigm in 8 to 9-year-old children with normal hearing. The authors attributed this negative effect to a combination of loss of early reflections and seat positions outside the critical distance. Combined with the findings in adults, the data from Amlani and Russo provide support for the competing hypothesis that reverberation will not increase listening effort in school-aged children.

Note that increases in reverberation resulted in increased listening effort in adults using subjective paradigms but not using behavioral paradigms. This discrepancy might be attributable to the different listening effort methodologies. Physiology (e.g., pupillometry) and behavioral (e.g., recall and response time) measures have been shown to be sensitive, indirect, indicators of listening effort (McGarrigle et al., 2014; Strand et al., 2018). While subjective ratings are assumed to provide a more direct estimate of an individual's perceived listening effort, these ratings are often not associated with behavioral or physiologic measures (e.g., Feuerstein, 1992; Zekveld et al., 2010; Lemke and Besser, 2016; Picou et al., 2017b; Strand et al., 2018).

One explanation for the disparate findings is humans are not inherently disposed to accurately rate their listening effort; assigning a value to the "deliberate allocation of resources during listening" might be somewhat difficult. According to Kahneman and Frederick (2002), when faced with answering a difficult question (e.g., effort judgement), people answer an easier, substitute question, if a substitute attribute is highly accessible and reasonable. For effort judgements, some investigators have suggested participants use performance judgements as substitute attributes to make their ratings of

effort (e.g., Moore and Picou, 2018), since judgements of word recognition performance are easy and accurate (Cox et al., 1991; Cienkowski and Speaks, 2000).

According to Kahneman and Frederick (2002), if the target attribute is accessible or if there is no reasonable alternative substitute, people would be less likely to use a heuristic. Thus, instead of using language that includes the words "effort" or "work," it might be possible to use language that elicit judgements of "effort" that align with behavioral indices of listening effort. In adults, Picou et al. (2017b) and Picou and Ricketts (2017) identified that asking participants to judge the extent to which they wanted to control the listening situation ("want to do something to improve the situation, such as move to a quiet room or ask the talker to speak up") elicited subjective ratings that were more highly correlated with responses times in a dual-task paradigm than did asking participants "how hard" they had to work or how "tired" they were. That is, the desire to control the situation was a target attribute that was easy to answer and yet was still associated with behavioral listening effort.

Reports of subjective ratings of effort from school-aged children are surprisingly scarce. The limited data available suggest that, as with adults, behavioral measures of listening effort and subjective ratings can be discrepant (Hicks and Tharpe, 2002; Gustafson et al., 2014). For example, Gustafson et al. (2014) reported that digital noise reduction in hearing aids improved ratings of clarity and reduced listening effort (measured behaviorally using verbal response times), though the two outcomes were not correlated. Based on the findings in adults, it might be possible to use language in the ratings task to elicit responses from school-aged children that align with behavioral indices of listening effort. However, the questions used by Picou et al. (2017b) are likely not appropriate for school-aged children.

For the current study, the questions established by Picou et al. (2017b) were modified for language and content. Specifically, to evaluate a target attribute of "control," the question was reworded to have participants rate the degree to which they wanted to "turn up the lady's voice" (the study stimuli were spoken by a female talker). This question was a simpler version of the question used previously.

In addition to modifying the control question, a new question was developed to assess children's perception of the passage of time (i.e., "how long did that feel"). The sense of time passing is complex and multidimensional, but in some circumstances can be affected by cognitive load (Khan et al., 2006; Block et al., 2010). For example, in adults, simple laboratory tasks are perceived as taking longer than tasks that require deeper processing (Sucala et al., 2011). If someone is investing more resources during a listening task, fewer resources would be available for time awareness. Thus, if a task felt fast, it would indicate a participant was more cognitively engaged (exerting more listening effort) than if a condition felt slow. In total, the current study employed four subjective rating questions, two that are relatively straightforward, querying perceived performance and ease of listening, and two questions with the potential to associate with behavioral listening effort by probing related constructs, control and time.

A concept closely related to listening effort is mental fatigue. Fatigue is a multi-dimensional phenomenon that may be observed as a decrement in performance over time, or subjectively as a mood state, associated with feelings of tiredness, a lack of energy or motivation to continue on a task (Pichora-Fuller et al., 2016). Listening-related fatigue is thought to result, in part, from the application of sustained effort (Hornsby, 2013; Hornsby et al., 2016). However, evidence of fatigue as a result of sustained listening has not been empirically demonstrated in school-aged children. Two studies have evaluated potential fatigue in this population, both using a scale described by Bess and Hornsby (2014). The scale, referred to here as the "Right Now Fatigue Scale" is administered at various times throughout a test session and asks a participant to rate how they feel "right now" on five questions. The questions probe the degree to which a participant feels tired, that the task is easy, they are able to focus, they have trouble thinking, or their head hurts.

McGarrigle et al. (2017) used the survey, in addition to response-time and pupillometry indices, to evaluate listening-related fatigue and effort in two environments. The environments reflected a "typical" classroom with a poor SNR and an "ideal" classroom with a more favorable SNR. Outcomes were the same on all tasks after listening in both rooms. In addition, ratings of fatigue were generally low, suggesting participants did not experience listening-related fatigue. However, participants completed the Right Now Fatigue Scale only at the end of testing in each environment. It is possible listening-related fatigue would have been evident as a change in fatigue ratings relative to a pre-test score. In addition, the authors only analyzed a total fatigue score, calculated as the mean response to all five questions. It is not clear if all five questions are equally sensitive to listening-related fatigue.

Another study using the Right Now Fatigue Scale provides indirect evidence of listening-related fatigue. Bess and Hornsby (2014) reported descriptive changes in self-reported fatigue using mean scores from all five questions obtained at several time points throughout the course of a research visit lasting 2.5 to 3 h. Although fatigue scores were generally low, the authors described increased fatigue over the duration of the research visit, which included both active and passive listening tasks. However, the changes in fatigue were small and not analyzed statistically. Thus, it remains unclear if sustained, active listening affects fatigue in school-aged children. Furthermore, like McGarrigle et al. (2017), Bess and Hornsby only reported mean responses to all five questions on the scale. It is possible changes in fatigue would be larger with some questions (e.g., related to tiredness or task ease) than other questions (e.g., related to trouble thinking or head hurting). As noted above, the relative sensitivity of the five Right Now Fatigue Scale questions to listening-related fatigue have not been previously evaluated.

The purpose of this study was three-fold. The primary purpose was to evaluate the effects of noise and moderate reverberation on listening effort in school-aged children with normal hearing. Based on FUEL, it was hypothesized that noise and moderate reverberation would increase listening effort as evidenced by slower response times during a dual-task paradigm and by subjective ratings. It was also expected that the effects of

noise would be larger when the reverberation time was longer. A second purpose was to evaluate the relationship between subjective and behavioral measures of listening effort, with specific interest in questions that reconcile the noted discrepancy between behavioral and subjective indices. It was hypothesized that questions related to time and a desire to control the situation would be related to behavioral listening effort and questions related to performance and listening ease would be related to speech recognition scores. A third purpose was to evaluate the effect of reverberation on self-reported fatigue, taking into consideration the limitations of previous studies, notably the inclusion of a pre-test rating, evaluating fatigue after sustained, active listening, and analyzing responses to self-report questions separately. It was expected that the change in fatigue would be higher after sustained listening in moderate, compared to low, reverberation.

MATERIALS AND METHODS

Participants

Twenty school-aged children (five males) participated in the study (aged 10 to 17 years, $M = 13.25$, $SD = 2.34$). Participants were recruited via word of mouth and via e-mail solicitation to people who have opted in to receive e-mail notifications regarding research participation opportunities. All participants had normal hearing bilaterally, as evidenced by pure-tone, air conduction thresholds of 20 dB HL or better. In addition, all participants exhibited normal middle ear function on the day of testing, as indicated by normal middle ear pressure and compliance measured with 226 Hz tympanometry. Based on participant and parent/guardian self-report, all participants were typically developing with no known neurological, cognitive, vision, or developmental disorders.

All participants underwent speech in noise testing using the Bamford-Kowal-Bench, Speech in Noise test (BKB-SIN; Etymotic Research, 2005). The purpose of this test was to evaluate a participant's speech understanding in noise ability in order to establish the SNR to be used for the listening effort and fatigue procedures. For the listening effort and fatigue procedures, it was desirable to target specific performance levels (described below). The use of the BKB-SIN procedures allowed for setting of individualized SNRs without using the same stimuli that would be used later for experimental testing. Pilot testing was used to establish the relationship between BKB-SIN scores and SNRs necessary to approximate 84%- (easy) and 77%- (moderately difficult) word recognition performance with the experimental stimuli. Participants for pilot testing included adults and children (10–17 years old) with normal hearing bilaterally; these participants were not otherwise involved with the study.

Testing with BKB-SIN was accomplished bilaterally through supra-aural headphones (TDH-50) using standard test instructions in an audiometric sound booth. One passage pair was used for each participant. A passage pair consists of 10 sentences spoken by a male talker presented in a four-talker babble background noise. The SNR is progressively decreased in increments of 3 dB after each sentence. The starting SNR is

+21 dB and is progressively decreased in 3 dB steps to −6 dB. Specifically, the background noise level increases in 3 dB steps until the 8th sentence (0 dB SNR), for the remaining two sentences the level of the speech is decreased in 3 dB increments and the level of the noise is held constant. Consistent with test instructions, the level of the speech was set initially to be 70 dB HL (83 dB SPL). All stimuli during the BKBSIN test were routed from a compact disc player to an audiometer (Grason Stadler 61) and then to the headphones. After each sentence, the experimenter scored the number of keywords a participant correctly repeated back. Also based on test instructions, the SNR where participants were expected to understand 50% of speech (SNR-50) was calculated. SNR-50s recorded from study participants ranged from −2 to +4 dB ($M = 0.5$, $SD = 1.54$).

Procedures were approved by the Behavioral Sciences Committee at Vanderbilt University Medical Center's Institutional Review Board (IRB # 180919). All participants gave written informed assent and parents/guardians provided written informed consent. Participants were paid an hourly rate; most testing was accomplished in a single test visit lasting approximately 2 h. This project was pre-registered with the Center for Open Science (osf.io/9dj2q).

Behavioral Listening Effort

Behavioral listening effort was evaluated using a dual-task paradigm. The paradigm, described in detail by Picou et al. (2017a), included a primary task (monosyllable word recognition) and a secondary task (physical response to a visual probe). The monosyllable words, spoken by a female talker with an American English accent, were all nouns. The words were arranged into 8, 25-word lists based on pilot testing (completed with naïve adults with normal hearing). During presentation of the words, colored shapes (blue circle, blue triangle, yellow circle, or yellow triangle) were occasionally presented (18 out of 25 words). Participants' secondary task was to respond as quickly as possible by pressing a touchscreen monitor when the correct color/shape combination was displayed (blue circle and yellow triangle) and to not touch the screen when the incorrect shapes were presented (blue triangle and yellow circle). They were instructed to repeat every word, regardless of the visual probe. Half of the shapes were probes (blue circle and yellow triangle) and half were foils or non-probes (blue triangle and yellow circle). The order of probe and non-probe trials was randomized across word lists. During the trials where no visual shape was displayed (7 out of 25 trials), a small white fixation cross (1 cm × 1 cm) was presented on a black screen. Colored shapes were approximately 6.5 by 6.5 cm and were also presented on a black screen.

Subjective "Listening Effort"

Questions to elicit subjective ratings were developed for this study, each with a visual analog scale with verbal anchors at the end points. The questions were:

(1) How many words did you get right? (none of them – all of them);
(2) How easy was that? (not at all easy – very easy);

(3) How much did you want to turn up the lady's voice? (not at all – a lot);
(4) How long did that feel? (it felt fast – it took forever).

An on-line survey was created with the four questions and four visual analog scales to facilitate data collection. The survey was presented to a participant after each condition using an internet-enabled tablet (Nexxus 7) with the survey visible. Participants responded to the questions in the same order using a response slider, which had 100 increments between the anchors. The response numbers were not visible to participants. Higher scores indicated participants rated their performance as higher, the task easier, had a stronger desire to turn up the talker's voice, and had a longer perception of test time.

Self-Reported Fatigue

All five questions from the Right Now Fatigue Scale were used to evaluate self-reported fatigue. The questions were described by Bess and Hornsby (2014) and were later used experimentally by McGarrigle et al. (2017). The questions are thought to relate to the constructs underlying fatigue. When answering the questions, participants were instructed to consider how they feel "right now." Response options for all questions were "not at all (0)," "a little (1)," "some (2)," "quite a bit (3)," "a lot (4)." The questions were:

1- I feel tired;
2- It is easy for me to do these things;
3- My head hurts;
4- It's hard for me to pay attention;
5- I have trouble thinking.

The anchor response options included schematic drawings of children experiencing the question response (e.g., a head down on the desk for "a lot" on the tired question). The complete survey is displayed in Appendix A of McGarrigle et al. (2017). For this study, the questionnaire was converted to an on-line survey, separate from the subjective rating survey. The response options were radio buttons. Surveys were presented to participants twice in a given test room (i.e., the low or moderately reverberant room). The first survey was given just prior to dual-task testing in a given room ("pre-test") and again immediately following completion of all testing in the same room ("post-test").

Conditions

Participants completed dual-task testing and provided subjective ratings in six conditions, which varied by degree of reverberation (low and moderate) and background noise (quiet, easy, and moderately difficult). Testing in the low reverberation condition was completed in an audiometric test booth (T30 < 100 ms); testing in the moderate reverberation condition was completed in a moderately reverberant room (T30 = 834 ms). The T30 value is approximately equivalent to the RT60 measure; it is expressed as double the time it takes for energy to decay from 5 to 35 dB below the initial level (ISO 3382-1, 2009).

The background noise, when present, was a four-talker babble, as described in Picou et al. (2017a). Briefly, four female talkers simultaneously read sentences from the Connected Speech Test

(Cox et al., 1987). Each talker's voice originated from a single loudspeaker. The loudspeaker location of the talker changed after each sentence. The same sentence was never read by two talkers at the same time.

The background noise conditions were achieved by varying the level of the noise. In quiet, no background noise was present. In the other conditions, the level of background noise was chosen relative to a participant's BKB-SIN SNR-50 score to create an "easy" and a "moderately difficult" test condition. The use of individualized SNRs based on a participant's speech understanding in noise abilities ensured participants were listening in a performance range where the listening effort task would be sensitive to changes in SNR. Previous work demonstrates that response times during listening effort tasks exhibit an inverse U-shaped function (Wu et al., 2016), where response times progressively increase until a point of cognitive overload where participants exert less effort because cognitive demands exceed cognitive resources (e.g., Granholm et al., 1996; Zekveld et al., 2014). According to Wu et al. (2016), in adults, response times peak around 30–50% correct performance levels. It was desirable in this study to keep performance in a range where changes in SNR would not result in response times in the cognitive overload section of the performance-intensity function. Thus, word recognition performance levels were targeted to be 84 and 77% correct. Based on the aforementioned pilot testing, the "easy" condition was a SNR set to be 5 dB less favorable than the participant's BKB-SIN score. The mean noise level in the "easy" condition, hereafter referred to as the SNR84 condition, was 69.5 dB. The "moderately difficult" condition was a SNR set to be 9 dB less favorable than the participant's BKB-SIN score. The average background noise level in this condition, hereafter referred to as SNR77, was 73.5 dB SPL. The speech was always 65 dB SPL, resulting in mean SNRs of −4.5 and −8.5 dB for the SNR84 and SNR77 conditions, respectively.

Test Environment

In a sound booth (4 m × 4.3 m × 2.7 m), participants provided assent, a parent/guardian provided informed consent, and a researcher completed tympanometry, hearing testing, and BKB-SIN testing. In addition, dual-task testing and subjective ratings comprising the low reverberation conditions (T30 < 100 ms) were completed. Speech signals were presented via custom programming of experimental software (Presentation v 14, Neurobehavioral Systems), routed through an audiometer (Madsen Orbiter 922 v2), to a loudspeaker (Bowers and Wilkins 685 S2) 1.25 m in front of a listener (0°). The four background noise channels were presented via sound editing software (Adobe Audition CSS5) and a multichannel sound card (Layla Echo), to an amplifier (Russound DPA-6.12), and finally to loudspeakers (Bowers and Wilkins 685 S2). The loudspeakers were 1.25 m from the participant and were placed at 45, 135, 225, and 315°.

Dual-task testing was also completed in a moderately reverberant room (5.5 m × 6.5 m × 2.25 m), which has solid, random-incidences, walls and ceilings, and a concrete floor. Unoccupied and untreated, the T30 of this test space is approximately 2100 ms. Floor carpet and four ceiling acoustic blankets (Sound Spotter 124, 4 × 4) were used to limit

reverberation to the desired level (T30 = 834 ms). During testing, the speech was presented from a separate control room via custom programming of experimental software (Presentation v 12.0, Neurobehavioral Systems) and was routed to a self-powered loudspeaker (Tannoy 600A) 1.25 m in front of a participant (0°). The noise was routed from sound editing software (Adobe Audition v1.5) and a multichannel sound card (Layal Echo) through an amplifier (Crown) and to the four noise loudspeakers (Tannoy System 600). The loudspeakers were located 3.5 m from the participant at 45, 135, 225, and 315°. In both rooms, visual probes were displayed on a touchscreen monitor (Dell S2240T) placed directly in front of a participant. The monitor accepted touch responses via USB cable connected to the experimental control computer.

Procedures

Table 1 indicates the procedural order and approximate test time for study tasks. After informed consent and assent procedures, a participant underwent hearing and immittance testing using standard clinical procedures. Then, they completed dual-task testing in one of the two rooms. In a given room, participants first completed three practice conditions: (1) secondary-task only in quiet, (2) primary and secondary tasks combined in quiet, (3) primary and secondary tasks combined in background noise with a favorable SNR (1 dB less favorable than a participant's SNR-50 with expected word recognition performance of 98%, hereafter labeled SNR98). Immediately following these three practice conditions, participants performed the secondary task only in quiet again. This served as their room-specific baseline. Following these four conditions, participants completed the self-report fatigue questionnaire (pre-test fatigue). Then, each participant completed dual-task procedures in a given SNR. Following each 25-word list of dual-task testing in a given condition, the participant answered the four subjective ratings questions, answering the questions about their experience during the dual-task testing. Condition order (quiet, SNR84, and SNR77) within a room was randomized across participants. Each condition was tested twice; the second round of condition testing was initiated immediately after the first round was completed. After testing was completed in one room, participants answered again the five fatigue questions (post-test fatigue). Testing in a given room lasted approximately 25 min and breaks were discouraged during testing. After testing was fully completed in a room, participants took a 15-min break and switched rooms. Test order of rooms was counterbalanced across participants; half were tested in the low reverberant room first.

Data Analysis

Prior to analysis, word recognition scores were converted to rationalized arcsine units (RAU) according to the equations in Studebaker (1985). Word recognition scores, response times, and subjective ratings were evaluated separately using generalized linear models with two factors of interest: SNR (quiet, SNR84, and SNR77) and reverberation (low and moderate) and participant as a random factor. The relationship between response times and subjective ratings was explored using partial correlation analyses, statistically controlling for SNR

TABLE 1 | Order of study procedures.

	Procedure	Task	SNR	Number of words or questions	Approximate time (min)
1	Informed consent and assent				10
2	Audiometric evaluation and BKBSIN testing				10
3	Practice 1	Secondary task only	Quiet	25	3
4	Practice 2	Dual tasks	Quiet	25	3
5	Practice 3	Dual tasks	SNR98	25	3
6	Baseline	Secondary task only	Quiet	25	3
7	Fatigue survey pre-test	Questionnaire		5	2
8	Condition 1a	Dual tasks	Quiet, SNR84, or SNR77	25	3
9	"Listening effort" survey 1a	Questionnaire		4	1
10	Condition 2a	Dual tasks	Quiet, SNR84, or SNR77	25	3
11	"Listening effort" survey 2a	Questionnaire		4	1
12	Condition 3a	Dual tasks	Quiet, SNR84, or SNR77	25	3
13	"Listening Effort" Survey 3a	Questionnaire		4	1
14	Condition 1b	Dual tasks	Quiet, SNR84, or SNR77	25	3
15	"Listening effort" survey 1b	Questionnaire		4	1
16	Condition 2b	Dual tasks	Quiet, SNR84, or SNR77	25	3
17	"Listening effort" survey 2b	Questionnaire		4	1
18	Condition 3b	Dual tasks	Quiet, SNR84, or SNR77	25	3
19	"Listening effort" survey 3b	Questionnaire		4	1
20	Fatigue survey post-test	Questionnaire		5	1
21	Break				15
22	Repeat procedures 3–20 in the second room				40

Conditions were repeated twice within each level of reverberation (indicated by "a" and "b" below). Detailed procedures reflect testing in the first level of reverberation, which were then repeated in the second level of reverberation following the break.

(quiet, SNR84, or SNR77). In the correlation analyses, data were pooled across conditions; no correction was made to account for multiple data points from the same participant. Responses to the five self-reported fatigue questions were analyzed as a single score based on a participant's mean response to all five questions (with responses to question two reversed). In addition, questions were analyzed separately because it was not clear which, if any, of the questions would be sensitive to fatigue. In all cases, responses were analyzed using a generalized linear model with two factors of interest (pre-test/post-test, low reverberation/moderate reverberation). All analyses were conducted with IBM SPSS Statistics 25.

RESULTS

Word Recognition Performance

Analysis of the transformed word recognition scores collected during the dual-task paradigm, displayed in **Figure 1A** (left panel), revealed a significant main effect of SNR $[F(2,73.37) = 231.70, p < 0.0001, \eta_p^2 = 0.86]$. The main effect of Reverberation ($p = 0.62$, $\eta_p^2 = 0.002$) and the Reverberation × SNR interaction ($p = 0.68$, $\eta_p^2 = 0.01$) were non-significant. The mean difference in performance between the low and moderate reverberation conditions was 0.94 RAU (95% CI: −2.79 to 4.67). Results of follow-up pairwise comparison testing, displayed in **Table 2**, revealed word recognition performance was significantly different in all SNRs ($p < 0.001$). These

data demonstrate adding background noise and increasing the background noise both significantly reduced word recognition performance, but increasing reverberation did not affect performance.
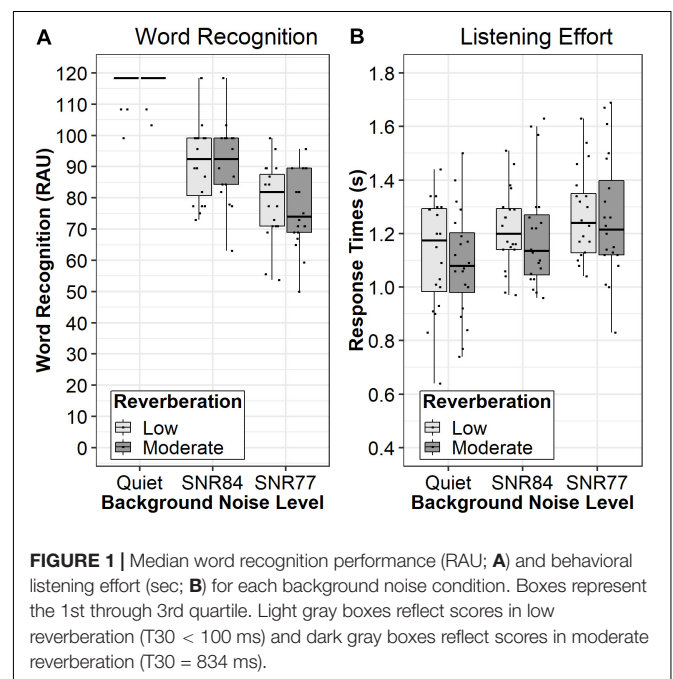
FIGURE 1 | Median word recognition performance (RAU; **A**) and behavioral listening effort (sec; **B**) for each background noise condition. Boxes represent the 1st through 3rd quartile. Light gray boxes reflect scores in low reverberation (T30 < 100 ms) and dark gray boxes reflect scores in moderate reverberation (T30 = 834 ms).

TABLE 2 | Mean differences between background noise conditions (quiet, SNR84, and SNR77) with each of the outcomes (word recognition, response times, and four subjective ratings).

	Word recognition scores (RAU)			Response times (ms)	
	SNR84	SNR77		SNR84	SNR77
Quiet	−24.94	−38.94	**Quiet**	98.10	155.61
p-Value	<0.001	<0.001	p-Value	0.025	.001
95% CI	−20.75 to 29.13	−34.87 to −43.01	95% CI	12.40 to 183.80	64.54 to 246.69
SNR84		−14.00	**SNR84**		57.51
p-Value		<0.0001	p-Value		0.183
95% CI		−8.58 to −19.43	95% CI		027.71 to 142.74
	Ratings of performance			**Ratings of control**	
	SNR84	SNR77		SNR84	SNR77
Quiet	−15.34	−25.86	**Quiet**	37.40	49.48
p-Value	<0.0001	<0.0001	p-Value	<0.0001	<0.0001
95% CI	−8.98 to −21.70	−18.78 to −32.95	95% CI	27.71 to 47.09	39.86 to 59.09
SNR84		−10.53	**SNR84**		12.08
p-Value		0.013	p-Value		0.012
95% CI		−2.27 to −18.78	95% CI		2.77 to 21.38
	Ratings of ease			**Ratings of time**	
	SNR84	SNR77		SNR84	SNR77
Quiet	−25.94	−39.60	**Quiet**	14.80	21.06
p-Value	<0.0001	<0.0001	p-Value	<0.0001	<0.0001
95% CI	−16.72 to −35.16	−30.16 to −49.04	95% CI	5.16 to 24.44	12.56 to 29.56
SNR84		−13.66	**SNR84**		6.26
p-Value		0.013	p-Value		0.236
95% CI		−2.98 to −24.34	95% CI		−4.20 to 16.72

Negative values indicate scores in quiet were higher than scores in noise or that scores in SNR84 were higher than scores in the SNR77 condition. Actual p-Values and 95% CI of the difference are also provided.

Behavioral Listening Effort

Mean baseline response times were 1036.8 ms (std. error = 47.46) and 1099.3 ms (std. error = 53.9) in the moderate and low reverberant conditions, respectively. They were not significantly different from each other [$F_{(1,37.41)} = 0.76$, $p = 0.39$]. Analysis of the response times during the dual-task paradigm, displayed in **Figure 1B** (right panel), revealed a significant main effect of SNR [$F_{(2,69.88)} = 5.94$, $p < 0.005$, $\eta_p^2 = 0.15$]. The main effect of Reverberation ($p = 0.57$, $\eta_p^2 = 0.003$) and the Reverberation × SNR interaction ($p = 0.97$, $\eta_p^2 < 0.001$) were non-significant. The mean difference in performance between the low and moderate reverberation conditions was 20.5 ms ($p = 0.57$, 95% CI: −50.4 to 91.4). Results of follow-up pairwise comparison testing, displayed in **Table 2**,

revealed significant response time differences only between quiet and noise conditions ($p < 0.05$). Taken together, these data demonstrate the addition of background noise increased behavioral listening effort, but further increases in background noise level or increased reverberation did not increase behavioral listening effort.

Subjective "Listening Effort"
Performance

Ratings of performance (how many words did you get right?) are displayed in **Figure 2A** (top left panel). Analysis revealed a significant main effect of SNR [$F_{(2,62.99)} = 31.69$, $p < 0.0001$, $\eta_p^2 = 0.50$]. The main effect of Reverberation ($p = 0.71$, $\eta_p^2 < 0.01$) and the Reverberation × SNR interaction ($p = 0.13$, $\eta_p^2 < 0.01$) were not significant. The mean difference in ratings between the low and moderate reverberation conditions was 1.09 (95% CI: −4.81 to 6.99). Follow-up pairwise comparison testing results, displayed in **Table 2**, revealed ratings of performance were significantly different in all SNRs ($p < 0.001$). This pattern of results is the same as the pattern of results for word recognition performance.

Ease of Listening

Ratings of ease of listening (how easy was that?) are displayed in **Figure 2B** (bottom left panel). Analysis results revealed a significant main effect of SNR [$F_{(2,66.94)} = 39.45$, $p < 0.0001$, $\eta_p^2 = 0.54$]. The main effect of Reverberation ($p = 0.38$, $\eta_p^2 < 0.01$) and the Reverberation × SNR interaction ($p = 0.23$, $\eta_p^2 = 0.04$) were non-significant. The mean difference in ratings between the low and moderate reverberation conditions was 3.52 (95% CI: −0.41 to 11.50). Follow-up pairwise comparison testing results, displayed in **Table 2**, revealed ratings of ease were significantly different in all SNRs ($p < 0.05$). This pattern of results is the same as the pattern of results for word recognition performance and perceived performance.

Control

Ratings of control (how much did you want to turn up the lady's voice?) are displayed in **Figure 2C** (top right panel). Analysis results revealed a significant main effect of SNR [$F_{(2,77.61)} = 55.87$, $p < 0.0001$, $\eta_p^2 = 0.50$]. The main effect of Reverberation ($p = 0.52$, $\eta_p^2 < 0.01$) and the Reverberation × SNR interaction ($p = 0.20$, $\eta_p^2 = 0.04$) were non-significant. The mean difference in ratings between the low and moderate reverberation conditions was 2.50 (95% CI: −5.25 to 10.25). Follow-up pairwise comparison results, displayed in **Table 2**, revealed ratings of control were significantly different in all SNRs ($p < 0.05$). This pattern of results is the same as the pattern of results for word recognition performance, perceived performance, and ease of listening ratings.

Time

Ratings of a listener's sense of time (how long did that feel?) are displayed in **Figure 2D** (bottom right panel). Analysis results revealed a significant main effect of SNR [$F_{(2,79.71)} = 13.31$, $p < 0.001$, $\eta_p^2 = 0.25$]. The main effect of Reverberation ($p = 0.15$,
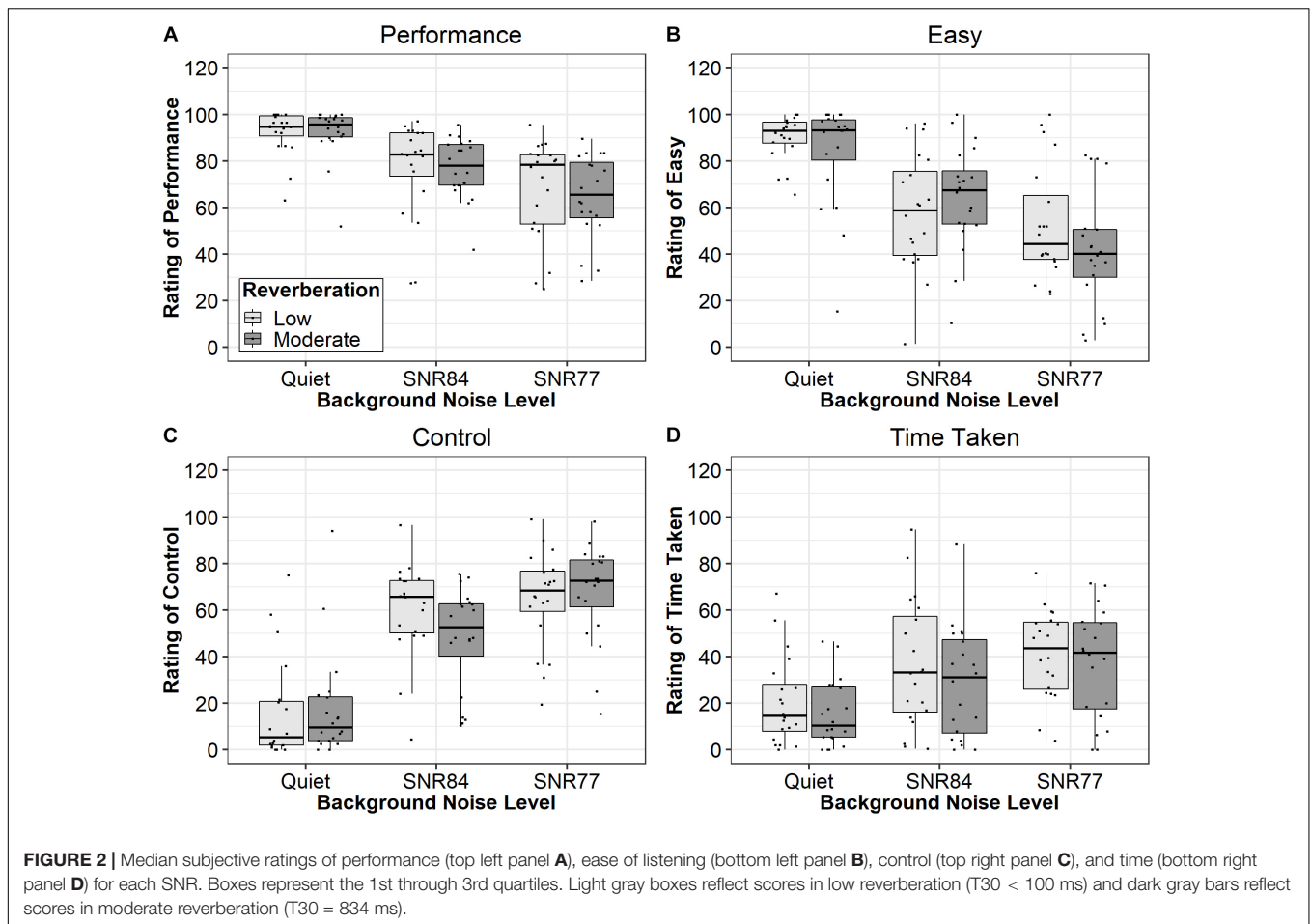
FIGURE 2 | Median subjective ratings of performance (top left panel **A**), ease of listening (bottom left panel **B**), control (top right panel **C**), and time (bottom right panel **D**) for each SNR. Boxes represent the 1st through 3rd quartiles. Light gray boxes reflect scores in low reverberation (T30 < 100 ms) and dark gray bars reflect scores in moderate reverberation (T30 = 834 ms).

TABLE 3 | Partial correlation coefficients (and p-Values in parentheses) examining the relationships between word recognition performance (RAU), response times during the secondary task (ms), and ratings of ease of listening, control, and time, while controlling for condition (quiet, SNR84, and SNR77).

	Word recognition	Performance	Ease	Control	Time
Response times	−0.44 (*<0.001*)	−0.15 (*0.06*)	−0.09 (*0.29*)	0.001 (*0.99*)	0.17 (*0.03*)
Word recognition		0.26 (*<0.01*)	0.18 (*0.02*)	−0.22 (*<0.01*)	−0.14 (*0.08*)
Performance			0.48 (*<0.001*)	−0.44 (*<0.001*)	−0.11 (*0.19*)
Ease				−0.61 (*<0.001*)	−0.34 (*<0.001*)
Control					0.25 (*<0.01*)

Ratings were on a 100-point scale. For all correlations, n = 160 and df = 157.

$\eta_p^2 = 0.02$) and the Reverberation × SNR interaction ($p = 0.94$, $\eta_p^2 < 0.01$) were not significant. The mean difference in ratings between the low and moderate reverberation conditions was 5.65 (95% CI: −2.12 to 13.42). Follow-up pairwise comparison results, displayed in **Table 2**, revealed ratings of giving up were significantly different in the noise conditions compared to quiet condition ($p < 0.01$). This pattern of results is the same as the pattern of results for response times during the secondary task.

Relationship Between Variables

Partial correlations were conducted between word recognition scores (RAU), response times (ms), and responses to each of the four questions while controlling for test SNR. Results, displayed in **Table 3**, reveal that the word recognition performance was significantly correlated with ratings of performance, ease of listening, and control [$r(157) = 0.18$ to 0.26], in addition to response times [$r(157) = 0.44$]. Word recognition performance was not correlated with ratings of time. Response times were correlated only with word recognition performance and with ratings of time [$r(157) = 0.17$]. These data demonstrate ratings of time are related to response times, whereas ratings of control, ease of listening and perceived performance are related to word recognition performance.

To evaluate the accuracy of subjective ratings of performance, a repeated measures analysis of variance was conducted

with three within-participant factors: outcome variable (word recognition performance in percent correct, rating of perceived accuracy), reverberation (low and moderate), and SNR (quiet, SNR84, and SNR77). This analysis was not planned *a priori* and thus not included in the pre-registration. Results indicated a significant main effect of outcome [$F(1,19) = 15.20$, $p < 0.01$, $\eta_p^2 = 0.44$] and a main effect of SNR [$F(2,18) = 34.48$, $p < 0.001$, $\eta_p^2 = 0.79$]. The main effect of Reverberation and all the interactions were non-significant ($p > 0.50$). These results demonstrate participants underestimated their word recognition performance ($M = 9.69$, 95% CI: 4.49 to 14.89), but the magnitude of the underestimation was consistent across conditions. That is, across conditions, participants rated their performance as 9.69 percentage points lower than their actual word recognition performance.

Self-Reported Fatigue

Mean responses to all five self-reported fatigue questions, in addition to the mean fatigue score (with question two reversed) are displayed in **Figure 3**. When the mean of all five responses was used to indicate self-reported fatigue, analyses revealed a significant main effect of Time [$F(1,85.73) = 4.86$, $p < 0.05$,

$\eta_p^2 = 0.05$] and no significant effect of Reverberation ($p = 0.86$, $\eta_p^2 < 0.001$) or Reverberation $\times$ Time interaction ($p = 0.90$, $\eta_p^2 < 0.001$). The mean difference between pre- and post-test was 0.3 points (95% CI: 0.03 to 0.57), a 35.8% increase relative to pre-test ratings. Analysis of the question about tiredness revealed a significant main effect of Time [$F(1,71.05) = 4.90$, $p < 0.05$, $\eta_p^2 = 0.06$], a 56.1% increase in reported tiredness relative to pre-test ratings. The main effect of Reverberation ($p = 0.92$, $\eta_p^2 < 0.01$) and the Reverberation $\times$ Time interaction ($p = 0.92$, $\eta_p^2 < 0.01$) were not significant. These results demonstrate ratings of tiredness were significantly higher after sustained listening (M difference = 0.58). None of the other questions resulted in ratings that were significantly different in the post-test compared to the pre-test ($p > 0.10$, $\eta_p^2 < 0.03$). The mean differences between the pre- and post-tests ranged from 0.18 to 0.35 points. These data indicate increases in self-reported fatigue resulting from a sustained listening task, as measured by the overall score and by rating of tiredness, was independent of level of reverberation. Exploratory analysis with an additional variable, test order (first room versus second room), revealed an identical pattern of results, suggesting test order did not affect ratings of fatigue.

FIGURE 3 | Median responses to the self-reported fatigue questions related to feeling tired **(A)**, task ease **(B)**, head hurting **(C)**, paying attention **(D)**, difficulty thinking **(E)**, and total fatigue score **(F)**. Boxes represent the 1st through 3rd quartiles. Light gray boxes reflect scores in low reverberation (T30 <100 ms) and dark gray boxes reflect scores in moderate reverberation (T30 = 834 ms).

FIGURE 4 | Comparison of data from young adults (replotted data from Picou et al., 2016; **A**) and school-aged children **(B)**. Median percent listening effort relative to baseline is displayed where the response time during dual-task testing (RT_Dual_Task) is reflected as the percent increase relative to baseline testing (RT_Baseline). Specifically, percent listening effort is calculated as 100*(RT_Dual_Task-RT_Baseline)/RT_Baseline. Boxes represent the 1st through 3rd quartiles.

DISCUSSION

The purpose of this project was three-fold: (1) to evaluate the effects of noise and reverberation on behavioral listening effort and subjective ratings of performance, ease of listening, desire to control, and perception of time, (2) to evaluate the relationship between behavioral and subjective indices of listening effort and (3) to evaluate the effects of reverberation on self-reported listening-related fatigue. Each purpose will be considered in turn.

Effects of Noise and Reverberation on Listening Effort

Based on FUEL and ELU, it was expected that background noise and reverberation would both increase listening effort. However, the current results do not fully confirm this hypothesis. Although the addition of background noise increased listening effort, behavioral listening effort was the same in the low (T30 < 100 ms) and moderate reverberation conditions (T30 = 834 ms). There are a number of possible explanations that could account for the non-significant findings.

First, it is possible the dual-task was not sensitive to changes in reverberation, as dual-task results in children might be less valid compared to other methodologies (Choi et al., 2008;

McGarrigle et al., 2019), unlike in adults where dual-task paradigms are accepted measures of behavioral effort (e.g., Gagne et al., 2017). To compare the results of this study with the results of an earlier study with young adults (Picou et al., 2016), percent change in listening effort was calculated using the following formula:

$$Percent\ Listening\ Effort = \frac{100 * (RT_{Dual\_task} - RT_{Baseline})}{RT_{Baseline}}$$

where RT_{Dual_Task} is the secondary task response time in a given condition and $RT_{Baseline}$ is the secondary task response time without the primary task. **Figure 4** displays percent listening effort for adults (Picou et al., 2016) and school-aged children (current study). For both groups, introducing background noise increased listening effort, whereas increasing the background noise and increasing reverberation time did not increase listening effort. Thus, the pattern of results with the school-aged children, although more variable, was similar to the findings in adults.

Second, the results of the study are limited to the specific acoustic conditions evaluated, which include relatively short stimuli (words rather than sentences or passages), SNRs that resulted in good word recognition performance (mean lowest performance 77%), a relatively small test room (moderate reverberation room was approximately 80 m³), and a speaker

and listener inside the critical distance. Larger rooms are more likely to have longer reverberation times (Knecht et al., 2002) and potentially more detrimental reverberation effects because late reflections comprise a greater portion of the total reverberant energy. Furthermore, if a listener is outside the critical distance from a loudspeaker, or the distance at which reverberant and direct energy are equivalent (Peutz, 1971; Egan, 1988), reverberant energy will dominate the signal, potentially increasing the effects of reverberation on listening effort. These acoustic factors warrant consideration in future work.

Third, and related, the moderate reverberation time was only 834 ms. For the school-aged children in this study, this level of reverberation was insufficient to affect word recognition performance. Thus, it might not be expected to affect listening effort either. Interestingly, Picou et al. (2016) used the same reverberation time with the young adults whose data are presented in **Figure 4**. In the earlier study, moderate reverberation did reduce word recognition performance, whereas it did not for the children in the current study. The reason for the discrepancy is not clear and might be related to the increased variability in the school-aged children or to the typical experiences of children, who routinely listen in reverberant classroom environments. Regardless of the explanation, it seems clear that future work is necessary to evaluate the limits of the non-significant reverberation effects on listening effort.

Fourth, participants were tested in reverberant rooms and were permitted to move their heads during testing. Conversely, investigators who previously demonstrated increases in listening effort with increased reverberation used recorded signals convolved with impulse room responses (Sato et al., 2008, 2012; Rennies et al., 2014). This methodology allows for testing across a wide range of reverberation times in a controlled manner, but unnaturally eliminates head movements. Head movements can help listeners resolve ambiguous cues (Wallach, 1939, 1940) and improve their SNR (Grange and Culling, 2016). Thus, it is possible that in real rooms, the negative consequences of reverberation on listening could also be alleviated with head movements.

Fifth, the participant age range was large (10–17 years). It is possible the effects of reverberation on listening effort are more likely to be evident in one group of listeners than another, although it is not clear which group of listeners might be more likely to demonstrate changes in effort with increased reverberation. Relative to older children, younger children are more likely to demonstrate worse speech recognition performance in noise (Klatte et al., 2010b; Neuman et al., 2010) and in reverberation (Neuman and Hochberg, 1983), so they might also be more vulnerable to the effects of noise and reverberation on listening effort. Conversely, the younger children tend to be more variable on some measures of listening effort (Picou et al., 2017a) and the additional variability might limit the possibility of demonstrating significant effects of reverberation on listening effort. Exploratory analysis with the current data set revealed a similar pattern of results with children when divided into four age groups (10–11, 12–13, 14–15, and 16–17 years). However, the sample size in each age group

precludes full investigation into the developmental effects of reverberation or SNR on listening effort, warranting future study.

Finally, it is possible that moderate reverberation does not increase behavioral listening effort, contrary to the expectations outlined in the existing frameworks. This final possibility is based on converging lines of emerging evidence, such as increased listening effort with the addition of acoustic paneling (Amlani and Russo, 2016), non-significant effects with behavioral paradigms (Picou et al., 2016; Peng and Wang, 2019), and differential physiological effects of noise and distortion (Francis et al., 2016). In some cases, the reverberation affected word recognition performance but did not have a comparable detrimental effect on listening effort (Peng and Wang, 2016, 2019; Picou et al., 2016). If future studies continue to demonstrate results contrary to framework predictions, it will be necessary to update the FUEL (Pichora-Fuller et al., 2016) and ELU framework (Rönnberg et al., 2013). Perhaps factors that affect signal transmission, such as noise and reverberation, should be considered separately, rather than assuming that all interferers with signal transmission increase listening effort.

Alternatively, the frameworks may need to be clarified to include the possibility that the long-term memory stores against which incoming speech signals are compared do not exclusively represent clean memory traces. Instead, it is possible that with experience (e.g., listening in classrooms), listeners can update or expand long-term memory representations to include distorted versions of speech. This possibility is consistent with an episodic theory of lexical access, which suggests perceptual details of speech (e.g., talker gender, speaking rate) are encoded in memory along with linguistic information (e.g., Goldinger, 1998; Grossberg, 2003) and with the observed effects of experience with reverberant stimuli (e.g., Zahorik and Brandewie, 2016). These hypotheses are speculation beyond the scope of this article, but they warrant further investigation.

Subjective Ratings of "Listening Effort"

A second purpose of this study was to evaluate the relationship between subjective and behavioral indices of listening effort in school-aged children. The results of the current study demonstrate that children's responses of perceived performance are significantly related to their ratings of actual performance. These data are consistent with findings in adult listeners, whose rated and actual performance are highly correlated (Cox et al., 1991; Cienkowski and Speaks, 2000; Saunders et al., 2004). Somewhat unlike adult listeners with normal hearing whose perceived and actual performances are nearly identical (e.g., Cox et al., 1991; Saunders et al., 2004), the school-aged children in this study consistently underestimated their performance by approximately 10 percentage points. This might reflect a lack of confidence in their understanding ability or the measurement methodology. The visual analog scale used for collecting subjective ratings included verbal anchors at the end points; no numbers were provided along the scale. Thus, participants were blinded to the score they were reporting.

The results of the current study also demonstrate that ratings of "ease of listening" are more closely related to actual and perceived performance than to the behavioral measure

of listening effort. This finding is consistent with the adult literature dissociation between ratings of "ease" or "effort" and behaviorally measured listening effort (Feuerstein, 1992; Hicks and Tharpe, 2002; Lemke and Besser, 2016). As suggested by these authors, among others, the emerging pattern of results discourage investigators from using "ease of listening" as a proxy for behavioral listening effort.

Instead of ease of listening, it was expected that a question related to desire to control the situation (turn up the lady's voice) would relate to behavioral listening effort, consistent with ratings of control in adults (Picou and Ricketts, 2017; Picou et al., 2017b). However, ratings of control revealed a pattern of results identical to those of word recognition performance, ratings of performance, and ratings of ease, suggesting participants were using their performance as a basis for rating their desire to control the listening situation. Because self-control has been related to willingness to accept background noise (Nichols and Gordon-Hickey, 2012), it is possible the difference between the results for children and adults is related to the development and understanding of self-control. It is also possible that ratings of control are affected differentially in quiet and in noise, where overall level of the speech might contribute to ratings in quiet but noise level dominates ratings in noise. Regardless of the explanation, it appears ratings of control were not an effective indirect, subjective measure of behavioral listening effort for children in this study.

Instead, subjective ratings of time perception were the only ratings associated with behavioral listening effort, as indicated by a significant correlation (**Table 3**) and by the same pattern of results as the response time data (see **Figure 2D**, bottom right panel compared to **Figure 1B**, right panel). Interestingly, the direction of the relationship between behavioral effort and subjective ratings of time to complete the task was unexpected. In adults, a decrease in perceived time is associated with higher task demands (Block et al., 2010; Sucala et al., 2011). Thus, it would be expected ratings of time would be negatively associated with response times during the dual-task paradigm; ratings of time would *increase* when listening effort *decreased*. The unexpected direction of the relationship might be related to the participant ages in the current study. Previous results demonstrate there are developmental effects of time perception; younger children are less sensitive to the effects of time (Zélanti and Droit-Volet, 2011). Thus, future work is warranted to investigate the interaction between the association between ratings of time, behavioral listening effort, and participant age.

Self-Reported Listening-Related Fatigue

The third purpose of this study was to evaluate the effects of reverberation on self-reported, listening fatigue. Results revealed increased self-reported listening fatigue with the fatigue question that addressed feeling tired and with a mean score reflecting responses to all five questions. The current data demonstrate that a relatively short, sustained listening task (approximately 25 min) can induce feelings of mental fatigue in both low and moderate reverberant conditions. Participants rated their tiredness as 0.58 points higher, or a 56% increase relative to pre-testing, after a relatively short, sustained listening activity. However, the effect

was the same in both environments, consistent with the listening effort data and with the findings of McGarrigle et al. (2017).

The results of this study also demonstrate that the five questions in the Right Now Fatigue Scale described by Bess and Hornsby (2014) are not equally sensitive to the effects of sustained listening. The only question that was sensitive to pre/post-test differences was the one related to tiredness. These data suggest that additional work is needed to validate a "right now" fatigue scale that is appropriate for use with children.

CONCLUSION

In summary, the findings of the current study have three important implications. First, in the modest range of SNRs and reverberation times evaluated, the current data do not support the conclusion that increased reverberation results in increased listening effort or fatigue. Instead, only the addition of background noise increased listening effort. These findings suggest the need for future careful investigation into the acoustic limits across which these findings hold true (e.g., longer reverberation times, larger rooms, greater speaker to listener distances). These data, coupled with emerging reports, question the assertion that moderate reverberation is a significant factor related to increases in listening effort. If confirmed, an update to the existing frameworks for understanding listening effort might be warranted. Second, the study results demonstrate that school-aged children's ratings of perceived performance are similar to their actual performance in controlled laboratory conditions. Moreover, their ratings of "ease of listening" are also related to their word recognition performance. Participants' perceived test time was the best candidate for a proxy of behavioral listening effort, but more work is necessary to evaluate the validity and reliability of the question. Finally, a relatively brief, focused listening task can induce listening-related fatigue, as indicated by subjective ratings of "tiredness" and an overall right now fatigue score. In total, these data offer no evidence that increasing reverberation to moderate levels increases listening effort or fatigue, but the data do support the reduction of background noise in classrooms.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Behavioral Health Sciences Committee of the Vanderbilt University Medical Center's Institutional Review Board with written informed consent from all participants' parents/guardians and written informed assent of all participants. The informed consent and assent were in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

EP, BB, BH, and TR designed the study. BB was primarily responsible for the data collection. EP analyzed the data. EP, BB, SM, TR, and BH wrote and edited the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Amlani, A., and Russo, T. (2016). Negative effect of acoustic panels on listening effort in a classroom environment. *J. Am. Acad. Audiol.* 27, 805–815. doi: 10.3766/jaaa.15096

Berg, F. S. (1993). *Acoustics and Sound Systems in Schools*. San Diego, CA: Singular Publishing Group.

Bess, F. H., and Hornsby, B. W. (2014). The complexities of fatigue in children with hearing loss. *Perspect. Hear. Hear. Disord. Child.* 24, 25–39. doi: 10.1044/hhdc24.2.25

Bistafa, S. R., and Bradley, J. S. (2000). Reverberation time and maximum background-noise level for classrooms from a comparative study of speech intelligibility metrics. *J. Acoust. Soc. Am.* 107, 861–875. doi: 10.1121/1.428268

Block, R. A., Hancock, P. A., and Zakay, D. (2010). How cognitive load affects duration judgments: a meta-analytic review. *Acta Psychol.* 134, 330–343. doi: 10.1016/j.actpsy.2010.03.006

Bradley, J. S. (1986). Speech intelligibility studies in classrooms. *J. Acoust. Soc. Am.* 80, 846–854. doi: 10.1121/1.393908

Bradley, J. S., Reich, R., and Norcross, S. (1999). A just noticeable difference in C50 for speech. *Appl. Acoust.* 58, 99–108. doi: 10.1016/S0003-682X(98)00075-9

Bradley, J. S., and Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *J. Acoust Soc Am.* 123, 2078–2086. doi: 10.1121/1.2839285

Choi, S., Lotto, A., Lewis, D., Hoover, B., and Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *J. Speech Lang. Hear. Res.* 51, 1042–1054. doi: 10.1044/1092-4388(2008/076)

Cienkowski, K. M., and Speaks, C. (2000). Subjective vs. Objective intelligibility of sentences in listeners with hearing loss. *J. Speech Lang. Hear. Res.* 43, 1205–1210. doi: 10.1044/jslhr.4305.1205

Cox, R., Alexander, G., and Gilmore, C. (1987). Development of the connected speech test (CST). *Ear Hear.* 8(5 Suppl.), 119s–126s.

Cox, R. M., Alexander, G. C., and Rivera, I. M. (1991). Comparison of objective and subjective measures of speech intelligibility in elderly hearing-impaired listeners. *J. Speech Hear. Res.* 34, 904–915. doi: 10.1044/jshr.3404.904

Crandell, C., and Smaldino, J. (2000). "Classroom acoustics and amplification," in *Audiology Volume II: Treatment*, eds M. Valente, R. Roeser, and H. Hosford-Dunn (New York, NY: Thieme Medical Publishers), 382–410.

Crandell, C. C., and Smaldino, J. J. (1994). An update of classroom acoustics for children with hearing impairment. *Volta Rev.* 96, 291–306. doi: 10.4103/1463-1741.104894

Crukley, J., Scollie, S., and Parsa, V. (2011). An exploration of non-quiet listening at school. *J. Educ. Audiol.* 17, 23–35.

Egan, M. D. (1988). *Architectual Acoustics*. New York, NY: McGraw-Hill.

Etymotic Research (2005). *BKB-SIN[TM] Speech in Noise Test Version 1.03 (compact disc)*. Elk Grove Village, IL: Etymotic Research.

Feuerstein, J. (1992). Monaural versus binaural hearing: ease of listening, word recognition, and attentional effort. *Ear Hear.* 13, 80–86. doi: 10.1097/00003446-199204000-00003

Finitzo-Hieber, T., and Tillman, T. W. (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *J. Speech Hear. Res.* 21, 440–458. doi: 10.1044/jshr.2103.440

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., and Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. *Front. Psychol.* 7:263. doi: 10.3389/fpsyg.2016.00263

Fraser, S., Gagne, J., Alepins, M., and Dubois, P. (2010). Evaluating the effort expended to understand speech in noise using a dual-task paradigm: the effects of providing visual speech cues. *J. Speech Lang. Hear. Res.* 53, 18–33. doi: 10.1044/1092-4388(2009/08-0140)

Gagne, J.-P., Besser, J., and Lemke, U. (2017). Behavioral assessment of listening effort using a dual-task paradigm: a review. *Trends Hear.* 21:2331216516687287. doi: 10.1177/2331216516687287

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychol. Rev.* 105, 251–279. doi: 10.1037/0033-295X.105.2.251

Grange, J. A., and Culling, J. F. (2016). The benefit of head orientation to speech intelligibility in noise. *J. Acoust. Soc. Am.* 139, 703–712. doi: 10.1121/1.4941655

Granholm, E., Asarnow, R. F., Sarkin, A. J., and Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology* 33, 457–461. doi: 10.1111/j.1469-8986.1996.tb01071.x

Grossberg, S. (2003). Resonant neural dynamics of speech perception. *J. Phon.* 31, 423–445. doi: 10.1016/s0095-4470(03)00051-2

Gustafson, S., McCreery, R., Hoover, B., Kopun, J. G., and Stelmachowicz, P. (2014). Listening effort and perceived clarity for normal-hearing children with the use of digital noise reduction. *Ear Hear.* 35, 183–194. doi: 10.1097/01.aud.0000440715.85844.b8

Haas, H. (1972). The influence of a single echo on the audibility of speech. *J. Audio Eng. Soc.* 20, 146–159.

Hétu, R., Riverin, L., Lalande, N., Getty, L., and St-cyr, C. (1988). Qualitative analysis of the handicap associated with occupational hearing loss. *Br. J. Audiol.* 22, 251–264. doi: 10.3109/03005368809076462

Hicks, C., and Tharpe, A. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *J. Speech Lang. Hear. Res.* 45, 573–584. doi: 10.1044/1092-4388(2002/046)

Hornsby, B. W. (2013). The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. *Ear Hear.* 34, 523–534. doi: 10.1097/AUD.0b013e31828003d8

Hornsby, B. W., Naylor, G., and Bess, F. H. (2016). A taxonomy of fatigue concepts and their relation to hearing loss. *Ear Hear.* 37, 136S–144S. doi: 10.1097/AUD.0000000000000289

Howard, C. S., Munro, K. J., and Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *Int. J. Audiol.* 49, 928–932. doi: 10.3109/14992027.2010.520036

Hsu, B. C.-L., Vanpoucke, F., and van Wieringen, A. (2017). Listening effort through depth of processing in school-age children. *Ear Hear.* 38, 568–576. doi: 10.1097/AUD.0000000000000436

ISO 3382-1 (2009). *Acoustics—Measurement of Room Acoustic Parameters—Part 1: Performance Spaces*. Brussels: International Organization for Standardization.

Kahneman, D. (1973). *Attention and Effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., and Frederick, S. (2002). "Representativeness revisited: attribute substitution in intuitive judgment," in *Heuristics and Biases: The Psychology of Intuitive Judgment*, eds T. Gilovich, D. Griffin, and D. Kahneman (New York, NY: Cambridge University Press), 49–81. doi: 10.1017/cbo9780511808098.004

Khan, A., Sharma, N. K., and Dixit, S. (2006). Effect of cognitive load and paradigm on time perception. *J. Indian Acad. Appl. Psychol.* 32, 37–42.

Klatte, M., Hellbrück, J., Seidel, J., and Leistner, P. (2010a). Effects of classroom acoustics on performance and well-being in elementary school children: a field study. *Environ. Behav.* 42, 659–692. doi: 10.1177/0013916509336813

Klatte, M., Lachmann, T., and Meis, M. (2010b). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise Health* 12, 270–282. doi: 10.4103/1463-1741.70506

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., and Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms: predictions and measurements. *Am. J. Audiol.* 11, 65–71. doi: 10.1044/1059-0889(2002/009)

Kramer, S., Kapteyn, T., and Houtgast, T. (2006). Occupational performance: comparing normally-hearing and hearing-impaired employees using the amsterdam checklist for hearing and work. *Int. J. Audiol.* 45, 503–512. doi: 10.1080/14992020600754583

Lemke, U., and Besser, J. (2016). Cognitive load and listening effort: concepts and age-related considerations. *Ear Hear.* 37, 77S–84S. doi: 10.1097/AUD.0000000000000304

Lewis, D., Schmid, K., O'Leary, S., Spalding, J., Heinrichs-Graham, E., and High, R. (2016). Effects of noise on speech recognition and listening effort in children with normal hearing and children with mild bilateral or unilateral hearing loss. *J. Speech Lang. Hear. Res.* 59, 1218–1232. doi: 10.1044/2016_jslhr-h-15-0207

Mackersie, C. L., and Cones, H. (2011). Subjective and psychophysiological indices of listening effort in a competing-talker task. *J. Am. Acad. Audiol.* 22, 113–122. doi: 10.3766/jaaa.22.2.6

McFadden, B., and Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Lang. Speech Hear. Serv. Sch.* 39, 342–351. doi: 10.1044/0161-1461(2008/032)

McGarrigle, R., Dawes, P., Stewart, A. J., Kuchinsky, S. E., and Munro, K. J. (2017). Measuring listening-related effort and fatigue in school-aged children using pupillometry. *J. Exp. Child Psychol.* 161, 95–112. doi: 10.1016/j.jecp.2017.04.006

McGarrigle, R., Gustafson, S. J., Hornsby, B. W., and Bess, F. H. (2019). Behavioral measures of listening effort in school-age children: examining the effects of signal-to-noise ratio, hearing loss, and amplification. *Ear Hear.* 40, 381–392. doi: 10.1097/AUD.0000000000000623

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., et al. (2014). Listening effort and fatigue: what exactly are we measuring? A british society of audiology cognition in hearing special interest group 'white paper'. *Int. J. Audiol.* 53, 433–445. doi: 10.3109/14992027.2014.890296

Moore, T. M., and Picou, E. M. (2018). A potential bias in subjective ratings of mental effort. *J. SpeechLang. Hear. Res.* 61, 2405–2421. doi: 10.1044/2018_JSLHR-H-17-0451

Murphy, D., Craik, F., Li, K., and Schneider, B. (2000). Comparing the effects of aging and background noise on short-term memory performance. *Psychol. Aging* 15, 323–334. doi: 10.1037/0882-7974.15.2.323

Nábělek, A. K., and Pickett, J. M. (1974). Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-Impaired listeners. *J. Speech Hear. Res.* 17, 724–739. doi: 10.1044/jshr.1704.724

Nábělek, A. K., and Robinette, L. (1978). Influence of the precedence effect on word identification by normally hearing and hearing-impaired subjects. *J. Acoust. Soc. Am.* 63, 187–194. doi: 10.1121/1.381711

Neuman, A. C., and Hochberg, I. (1983). Children's perception of speech in reverberation. *J. Acoust. Soc. Am.* 73, 2145–2149.

Neuman, A. C., Wróblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1044/jshr.1704.724

Nichols, A. C., and Gordon-Hickey, S. (2012). The relationship of locus of control, self-control, and acceptable noise levels for young listeners with normal hearing. *Int. J. Audiol.* 51, 353–359. doi: 10.3109/14992027.2011.645074

Pearsons, K. S., Bennett, R., and Fidell, S. (1977). *Speech Levels in Various Noise Environments*. Washington, DC: US Environmental Protection Agency.

Peng, Z. E., and Wang, L. M. (2016). Effects of noise, reverberation and foreign accent on native and non-native listeners' performance of English speech comprehension. *J. Acoust. Soc. Am.* 139, 2772–2783. doi: 10.1121/1.4948564

Peng, Z. E., and Wang, L. M. (2019). Listening effort by native and nonnative listeners due to noise, reverberation, and talker foreign accent during english speech perception. *J. Speech Lang. Hear. Res.* 62, 1068–1081. doi: 10.1044/2018_JSLHR-H-17-0423

Peutz, V. (1971). Articulation loss of consonants as a criterion for speech transmission in a room. *J. Audio Eng. Soc.* 19, 915–919.

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., et al. (2016). Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear.* 37, 5S–27S. doi: 10.1097/AUD.0000000000000312

Picou, E. M., Charles, L. M., and Ricketts, T. A. (2017a). Child–adult differences in using dual-task paradigms to measure listening effort. *Am. J. Audiol.* 26, 143–154. doi: 10.1044/2016_AJA-16-0059

Picou, E. M., Moore, T. M., and Ricketts, T. A. (2017b). The effects of directional processing on objective and subjective listening effort. *J. Speech Lang. Hear. Res.* 60, 199–211. doi: 10.1044/2016_JSLHR-H-15-0416

Picou, E. M., Gordon, J., and Ricketts, T. A. (2016). The effects of noise and reverberation on listening effort for listeners with normal hearing. *Ear Hear.* 37, 1–13. doi: 10.1097/AUD.0000000000000222

Picou, E. M., and Ricketts, T. A. (2017). The relationship between speech recognition, behavioral listening effort, and subjective ratings. *Int. J. Audiol.* 57, 457–467. doi: 10.1080/14992027.2018.1431696

Picou, E. M., Ricketts, T. A., and Hornsby, B. W. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear Hear.* 34, e52–e64. doi: 10.1097/AUD.0b013e31827f0431

Picou, E. M., Ricketts, T. A., and Hornsby, B. W. Y. (2011). Visual cues and listening effort: individual variability. *J. Speech Lang. Hear. Res.* 54, 1416–1430. doi: 10.1044/1092-4388(2011/10-0154)

Prodi, N., Visentin, C., and Farnetani, A. (2010). Intelligibility, listening difficulty and listening efficiency in auralized classrooms. *J. Acoust. Soc. Am.* 128, 172–181. doi: 10.1121/1.3436563

Rennies, J., Schepker, H., Holube, I., and Kollmeier, B. (2014). Listening effort and speech intelligibility in listening situations affected by noise and reverberation. *J. Acoust. Soc. Am.* 136, 2642–2653. doi: 10.1121/1.4897398

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theory, data, and clinical implications. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47, S99–S105. doi: 10.1080/14992020802301167

Sarampalis, A., Kalluri, S., Edwards, B., and Hafter, E. (2009). Objective measures of listening effort: effects of background noise and noise reduction. *J. SpeechLang. Hear. Res.* 52, 1230–1240. doi: 10.1044/1092-4388(2009/08-0111)

Sato, H., and Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *J. Acoust. Soc. Am.* 123, 2064–2077. doi: 10.1121/1.2839283

Sato, H., Morimoto, M., Sato, H., and Wada, M. (2008). Relationship between listening difficulty and acoustical objective measures in reverberant sound fields. *J. Acoust. Soc. Am.* 123, 2087–2093. doi: 10.1121/1.2885750

Sato, H., Morimoto, M., and Wada, M. (2012). Relationship between listening difficulty rating and objective measures in reverberant and noisy sound fields for young adults and elderly persons. *J. Acoust. Soc. Am.* 131, 4596–4605. doi: 10.1121/1.4714790

Saunders, G. H., Forsline, A., and Fausti, S. A. (2004). The performance-perceptual test and its relationship to unaided reported handicap. *Ear Hear.* 25, 117–126. doi: 10.1097/01.AUD.0000120360.05510.E5

Strand, J. F., Brown, V. A., Merchant, M. B., Brown, H. E., and Smith, J. (2018). Measuring listening effort: convergent validity, sensitivity, and links with cognitive and personality measures. *J. SpeechLang. Hear. Res.* 61, 1463–1486. doi: 10.1044/2018_JSLHR-H-17-0257

Studebaker, G. A. (1985). A "rationalized" arcsine transform. *J. Speech Hear. Res.* 28, 455–462. doi: 10.1044/jshr.2803.455

Sucala, M., Scheckner, B., and David, D. (2011). Psychological time: interval length judgments and subjective passage of time judgments. *Curr. Psychol. Lett.* 2, 1–9.

Surprenant, A. (1999). The effect of noise on memory for spoken syllables. *Int. J. Psychol.* 34, 328–333. doi: 10.1080/002075999399648

Valente, D. L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., and Lewis, D. E. (2012). Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. *J. Acoust. Soc. Am.* 131, 232–246. doi: 10.1121/1.3662059

Wallach, H. (1939). On sound localization. *J. Acoust. Soc. Am.* 10, 270–274. doi: 10.1121/1.1915905

Wallach, H. (1940). The role of head movements and vestibular and visual cues in sound localization. *J. Exp. Psychol.* 27, 339–368. doi: 10.1037/h0054629

Wróblewski, M., Lewis, D. E., Valente, D. L., and Stelmachowicz, P. G. (2012). Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. *Ear Hear.* 33, 731–744. doi: 10.1097/AUD.0b013e31825aecad

Wu, Y.-H., Stangl, E., Zhang, X., Perkins, J., and Eilers, E. (2016). Psychometric functions of dual-task paradigms for measuring listening effort. *Ear Hear.* 37, 660–670. doi: 10.1097/aud.0000000000000335

Zahorik, P., and Brandewie, E. J. (2016). Speech intelligibility in rooms: effect of prior listening exposure interacts with room acoustics. *J. Acoust. Soc. Am.* 140, 74–86. doi: 10.1121/1.4954723

Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2010). Pupil response as an indication of effortful listening: the influence of sentence intelligibility. *Ear Hear.* 31, 480–490. doi: 10.1097/AUD.0b013e3181d4f251

Zekveld, A., Rudner, M., Kramer, S., Lyzenga, J., and Rönnberg, J. (2014). Cognitive processing load during listening is reduced more by decreasing voice similarity than by increasing spatial separation between target and masker speech. *Front. Neurosci.* 8:88. doi: 10.3389/fnins.2014.00088

Zekveld, A. A., Kramer, S. E., and Festen, J. M. (2011). Cognitive load during speech perception in noise: the influence of age, hearing loss, and cognition on the pupil response. *Ear Hear.* 32, 498–510. doi: 10.1097/AUD.0b013e31820512bb

Zélanti, P. S., and Droit-Volet, S. (2011). Cognitive abilities explaining age-related changes in time perception of short and long durations. *J. Exp. Child Psychol.* 109, 143–157. doi: 10.1016/j.jecp.2011.01.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Consonant and Vowel Confusions in Well-Performing Children and Adolescents With Cochlear Implants, Measured by a Nonsense Syllable Repetition Test

Arne Kirkhorn Rødvik[1,2]*, Ole Tvete[2], Janne von Koss Torkildsen[1], Ona Bø Wie[1,2], Ingebjørg Skaug[3] and Juha Tapio Silvola[1,2,4]

[1] Department of Special Needs Education, Institute of Educational Sciences, University of Oslo, Oslo, Norway, [2] Cochlear Implant Unit, Department of Otorhinolaryngology, Division of Surgery and Clinical Neuroscience, Oslo University Hospital, Oslo, Norway, [3] Cochletten Foundation, Oslo, Norway, [4] Ear, Nose, and Throat Department, Division of Surgery, Akershus University Hospital, Lørenskog, Norway

Although the majority of early implanted, profoundly deaf children with cochlear implants (CIs), will develop correct pronunciation if they receive adequate oral language stimulation, many of them have difficulties with perceiving minute details of speech. The main aim of this study is to measure the confusion of consonants and vowels in well-performing children and adolescents with CIs. The study also aims to investigate how age at onset of severe to profound deafness influences perception. The participants are 36 children and adolescents with CIs (18 girls), with a mean (*SD*) age of 11.6 (3.0) years (range: 5.9–16.0 years). Twenty-nine of them are prelingually deaf and seven are postlingually deaf. Two reference groups of normal-hearing (NH) 6- and 13-year-olds are included. Consonant and vowel perception is measured by repetition of 16 bisyllabic vowel-consonant-vowel nonsense words and nine monosyllabic consonant-vowel-consonant nonsense words in an open-set design. For the participants with CIs, consonants were mostly confused with consonants with the same voicing and manner, and the mean (*SD*) voiced consonant repetition score, 63.9 (10.6)%, was considerably lower than the mean (*SD*) unvoiced consonant score, 76.9 (9.3)%. There was a devoicing bias for the stops; unvoiced stops were confused with other unvoiced stops and not with voiced stops, and voiced stops were confused with both unvoiced stops and other voiced stops. The mean (*SD*) vowel repetition score was 85.2 (10.6)% and there was a bias in the confusions of [i:] and [y:]; [y:] was perceived as [i:] twice as often as [y:] was repeated correctly. Subgroup analyses showed no statistically significant differences between the consonant scores for pre- and postlingually deaf participants. For the NH participants, the consonant repetition scores were substantially higher and the difference between voiced and unvoiced consonant repetition scores considerably lower than for the participants with CIs. The participants with CIs obtained

scores close to ceiling on vowels and real-word monosyllables, but their perception was substantially lower for voiced consonants. This may partly be related to limitations in the CI technology for the transmission of low-frequency sounds, such as insertion depth of the electrode and ability to convey temporal information.

INTRODUCTION

Provided with adequate access to environments in which speech is the common mode of communication, the majority of profoundly deaf children implanted in their sensitive period (before age 3.5–4.0 years) will develop intelligible speech and functional hearing for oral language (Kral and Sharma, 2012; Leigh et al., 2013; Dettman et al., 2016). Early implanted children follow similar development in speech and language as normal-hearing (NH) children do (e.g., the systematic review by Bruijnzeel et al., 2016). However, early implanted children with good speech perception ability do not discriminate minute details of speech, such as voicing, frication, and nasality, as well as their NH peers, even in quiet surroundings (Tye-Murray et al., 1995; Geers et al., 2003).

The present study aims to reveal possible systematic misperceptions of speech sounds in detail for children and adolescents with cochlear implants (CIs) and to investigate how age at onset of severe to profound (pre-, peri-, and postlingual) deafness influences their confusion of speech sounds and features. In the following, we will outline the maturation of the auditory system and the fundamentals of speech processing in CIs, before presenting the rationale for our test design and giving a brief introduction to the Norwegian language.

The human cochlea is fully developed at birth, but the brain's auditory pathways and centers, from the brain stem to the auditory cortex, continue to develop. Conditions for the acquisition of language are optimal in a sensitive period, which can be estimated by measuring the cortical P1 latency response as an index of maturation of the auditory pathway in populations with abnormal auditory experience, such as congenital profound deafness. Sharma et al. (2002a,b,c) found that the optimal sensitive period for cochlear implantation in profoundly deaf children lasts until approximately 3.5–4 years of age, and it is important that children receive auditory stimulation within this critical period. These children can still benefit from CIs until the eventual end of the overall sensitive period, at approximately 6.5–7.0 years of age (Kral and Sharma, 2012). However, later implantation in congenitally deaf children normally results in difficulties with acquiring oral speech and language skills.

As normal maturation of the auditory system depends on adequate auditory input in very early childhood, detection of hearing loss by otoacoustic emissions and/or auditory brainstem responses right after birth is crucial. Immediate programming of hearing aids (HAs) for infants with discovered mild to moderate hearing loss, or of CIs for the profoundly deaf among them, will facilitate stimulation of the brain's auditory pathways in the sensitive period. Clinical findings indisputably show that children with hearing impairments who receive appropriate and early intervention achieve much better hearing and better oral language performance than those who start the process later (Wilson and Dorman, 2008; Niparko et al., 2010; Wie, 2010).

The gradual development and maturation of the auditory system can be seen in outcomes of auditory tests into the late teenage years, with individual variability within a given age (Maxon and Hochberg, 1982; Fischer and Hartnegg, 2004). Children's peripheral hearing is established before their speech. However, the development of the ability to discriminate speech sounds, as well as vocabulary and language, takes many years.

Auditory sensitivity in audiometric tests, in absence of noise or other masking stimuli, is known to improve between infancy and early school age (Olsho et al., 1988; Trehub et al., 1988). Litovsky (2015) suggests that the reason for this improvement is that the tasks used to measure perception of pure-tones do not separate the effects of cognitive ability, motivation, memory, and variability in neural representation of the stimuli. For real-word tests, top-down processing allows for decoding based on context and is facilitated by the lexical content present in real-word stimulus materials or by the intrinsic language proficiency. To diminish the influence of these factors in the present study, auditory skills are measured by a nonsense syllable repetition test (NSRT), which is idealized to measure the perception of speech sounds with only minor influence from top-down processing and with minimal stress on working memory. This test should therefore establish a more correct expression of the true auditory perception skills of a child with CIs.

CI users are often classified into pre-, peri-, and postlingually deaf. In the present study, prelingual deafness is defined as congenital, profound deafness or onset of severe to profound deafness before the age of 12 months. According to the widely used definition by the World Health Organization [WHO] (2019), severe hearing loss is characterized by a pure-tone average (PTA)[1] between a 60 and 80 dB hearing level (HL), and profound hearing loss is characterized by a PTA above 80 dB HL. In prelingually deaf children, the auditory system is immature when hearing is initiated by a CI, whose stimulus signal is different from the signal generated by the inner hair cells in a normal cochlea. The earlier the age at implantation, the faster the adaptation to the novel signal, and the better the speech perception outcomes (Niparko et al., 2010; Tobey et al., 2013; Liu et al., 2015). Furthermore, prelingually deaf children with CIs can be divided into two groups: those who have had no or minimal access to sound and hence acquired very little oral language before implantation (these children are often congenitally deaf

---

[1]PTA is defined as average hearing loss on the frequencies 1,000, 2,000, 3,000, and 4,000 Hz, according to the National Institute for Occupational Safety, and Health [NIOSH] (1996).

and receive a CI before age 1), and those who have acquired oral language and benefited from HAs due to residual hearing, receiving a CI at a higher age.

The children with onset of severe to profound deafness between 1 and 3 years of age are classified as perilingually deaf. postlingual deafness is defined as progressive or sudden hearing loss and onset of severe to profound deafness after age 3 years, with a benefit from HAs and acquired oral language before onset of deafness (Myhrum et al., 2017).

Although language acquisition is a gradual process, the breakpoint of age 1 year for distinguishing between pre- and perilingual deafness is precisely defined for practical reasons. This age corresponds to when infants usually start saying their first words (Darley and Winitz, 1961; Locke, 1983, p. 8). In postlingually deaf adults and children, the neural pathways in the brain have been shaped by acoustic sound perception before onset of deafness. The degree of success with a CI is dependent on how the brain compares the new signal with what was heard previously.

For both the pre-, peri-, and postlingually deaf, auditory deprivation will occur after a period of lack of sensory input. This process entails a degeneration of the auditory system, both peripherally and centrally (Feng et al., 2018), including a degradation of neural spiral ganglion cells (Leake and Hradek, 1988). If profound deafness occurs in the sensitive period before 3.5–4.0 years of age, it arrests the normal tonotopic organization of the primary auditory cortex. This arrest can, however, be reversed after reactivation of afferent input by a CI (Kral, 2013).

The hearing-impaired participants in this study are aided by CIs, which consist of a speech processor on the ear and a surgically implanted electrode array in the cochlea with up to 22 electrical contacts. A speech signal input is received by the built-in speech processor microphone and translated into sequences of electrical pulses in the implant by a stimulation strategy. The main purpose of every such strategy is to set up an electrical signal in the auditory nerve using electrical stimulation patterns in the electrode array to mimic the signal in a normal ear. These patterns vary somewhat between stimulation strategies and implant manufacturers, but they all attempt to convey spectral (frequency-related) and temporal information of the original signal through the implant (Wouters et al., 2015).

The spectral information of the speech signal (e.g., the first and second formant, F1 and F2) is conveyed by the multichannel organization of the implants, by mimicking the tonotopic (place) organization of the cochlea from low frequencies in the apex to high frequencies in the base. This information is implemented in all stimulation strategies from the main (in terms of market share) implant manufacturers today, listed in alphabetical order: Advanced Bionics (Stäfa, Switzerland), Cochlear (Sydney, NSW, Australia), Med-El (Innsbruck, Austria), and Oticon Medical/Neurelec (Vallauris, France).

The temporal information of the speech signal is commonly decomposed into envelope (2–50 Hz), periodicity (50–500 Hz), and temporal fine structure (TFS; 500–10,000 Hz), for instance described by Wouters et al. (2015). The envelope is the slow variations in the speech signal. Periodicity corresponds with the vibrations of the vocal cords, which conveys fundamental frequency (F0) information. TFS is the fast fluctuations in the signal, and contributes to pitch perception, sound localization, and binaural segregation of sound sources.

All stimulation strategies represent high-frequency sounds only by place coding. Moreover, the stimulation rate in every implant is constant, varying between 500 and 3,500 pulses per second for the different manufacturers. Low-frequency sounds can be represented by both temporal and place coding.

In the present study, the consonant and vowel repetition scores and confusions were measured using an NSRT with recorded monosyllabic consonant-vowel-consonant (CVC) and bisyllabic vowel-consonant-vowel (VCV) nonsense words, named nonsense syllables in this article, in an open-set design. By open-set design, we mean that the responses are not made through a forced choice of alternatives, but rather by repetition of what is perceived. The nonsense syllables follow the phonotactic rules of the participants' native language, which in our case is Norwegian (e.g., Coady and Aslin, 2004). To avoid straining the working memory, each stimulus unit was limited to 1 or 2 syllables (Gathercole et al., 1994). In the following, the rationale for the test design is presented.

Speech perception tests for children with CIs are traditionally performed with live or recorded real words or sentences in quiet or in noise (e.g., Harrison et al., 2005; Zeitler et al., 2012; Ching et al., 2018). Such tests indisputably measure the children's language skills in addition to their auditory skills.

There are two methods of making speech perception tests more difficult in order for the test subjects not to perform at ceiling. One is to degrade the speech signal by altering its temporal and spectral information, for instance by adding background noise to the test words or applying high- or low-pass filtering. Perception of speech in background noise is more difficult than in quiet due to factors such as diminished temporal coding (Henry and Heinz, 2012). The other method is to use more challenging test units, such as words without lexical meaning, and assess details in the perception of individual speech sounds under optimal listening conditions. The use of an NSRT in quiet allows for directly studying feature information transmission as opposed to tests relying on a degraded speech signal. In real life, listeners are faced with challenging situations similar to NSRTs when they try to catch an unfamiliar name or are confronted with new vocabulary. New and difficult words are perceived as nonsense syllables until they become internalized as meaningful units.

The measurement of consonant and vowel scores in children with CI's via recorded nonsense syllables has rarely been reported in scientific literature. A systematic review and meta-analysis by Rødvik et al. (2018), found only two studies of this kind (Tyler, 1990; Arisi et al., 2010). Tyler (1990) included five children who were asked to choose between several written alternatives when they identified each nonsense syllable. Their mean (SD) age at testing was 8.5 (1.6) years, and they obtained a mean (SD) consonant identification score of 30% (13%) (range: 19–50%). The reason for this relatively low score was probably the high age at implantation for these prelingually ($N = 2$) and postlingually ($N = 3$) deaf children [mean (SD) = 7.4 (1.9) years]. Arisi et al. (2010) included 45 adolescents with a mean (SD) age of 13.4 (2.6)

years, who obtained a mean (*SD*) consonant identification score of 53.5 (33.6)%. All participants marked their choices with a pen on printed text.

We chose a test with verbal repetition of the test words, to ensure that the test scores would neither be influenced by the test subjects' reading or writing ability nor their computer skills, and that they were not required to relate to anything other than their own hearing and speech as well as their own established phoneme inventory. This design provided detailed information about speech perception and listening capacity for acoustic properties.

Furthermore, an open-set test design was chosen, in which the participants did not know which or how many test units would be presented to them. The participants were thus not limited in their responses and would find no external clues when interpreting what they heard. Previous studies have reported robust effects of competition between items in the mental lexicon and of speaker variability in open-set but not in closed-set tests (e.g., Sommers et al., 1997; Clopper et al., 2006). Moreover, open-set test designs have relatively small learning effects compared to main closed-set test designs and can therefore be performed reliably at desired intervals (Drullman, 2005, p. 8).

Open-set test designs also have some disadvantages. For example, they often result in lower overall performance than closed-set test designs and may be challenging to use with low-performing adults and young children. Moreover, they require a substantial effort in post-test analysis if each response is to be transcribed phonetically. Alternatively, responses may be scored simply as correct or incorrect for routine-testing in a clinical practice.

Norwegian is a Northern Germanic language, belonging to the Scandinavian language group. There is no official common Norwegian pronunciation norm, as oral Norwegian is a collection of dialects, and Norwegians normally speak the dialect of their native region. Norwegian has two lexical tones (except for certain dialects), which span across bisyllabic words and are used as a distinguishing, lexical factor. The tones' melodies are indigenous to each dialect and are recognized as a dominant and typical prosodic element of the dialect, distinguishing it from other dialects. Norwegian has a semi-transparent orthography, meaning that there is not a consistent one-to-one correspondence between letters and phonemes, like for instance in Finnish, but a much more transparent relation between phonemes and letters than in English (Elley, 1992). In the present study, only speech sounds common for all Norwegian dialects are included; see **Table 1** and **Figure 1** for an overview.

The overall objective of the present study was to measure the perception of speech sounds in well-performing children and adolescents with CIs with an NSRT.

The two sub-objectives were as follows:

Objective 1: To identify the most common vowel and consonant confusions and the most common confusions of the phonetic features voicing, frication, stopping, nasality, and laterality in a sample of well-performing children and adolescents with CIs.

Objective 2: To investigate how age at onset of severe to profound (pre-, peri-, and postlingual) deafness in children

and adolescents with CIs influences their confusion of speech sounds and features.

MATERIALS AND METHODS

Abbreviations and acronyms are presented in **Table 2**.

Participants

Informed written consent was obtained from all participants and their legal guardians, according to the guidelines in the Helsinki declaration (World Medical Association [WMA], 2017). The project was approved by the ethical committee of the regional health authority in Norway (REC South East) and by the data protection officer at Oslo University Hospital.

Participants With CIs

Thirty-six children and adolescents with CIs (18 girls) participated in this study. Their age range was 5.9–16.0 years [mean (*SD*) = 11.6 (3.0) years]. Oral language was the main communication mode for all participants. The study sample included 29 prelingually and 7 postlingually deaf participants using the CI stimulation strategies FS4 (*N* = 4), FSP (*N* = 7), and CIS + (*N* = 2) from Med-El and ACE (*N* = 23) from Cochlear (abbreviations are explained in **Table 2**).

The following inclusion criteria were met for all of these participants: minimum 6 months of implant use, more than 3 months since the activation of the second CI (if they had one), and unchanged processor settings for at least the last 2 months. Furthermore, the participants were required to obtain a score of more than 50% on the HIST monosyllable test in free-field (Øygarden, 2009) and to spontaneously pronounce 100% of all the Norwegian speech sounds correctly. Subjects with a contralateral HA were excluded.

All the included participants were enrolled in the CI program at Oslo University Hospital and were recruited for the present study as part of their ordinary follow-up appointments. Individual demographic information is shown in **Supplementary Table S1**, and individual test results are listed in **Supplementary Table S2**.

Reference Groups

The two reference groups of NH participants were: seventeen 6-year-olds (7 girls; [mean (*SD*) age = 5.9 (0.3) years; range: 5.3–6.3 years]), and twelve 13-year-olds (7 girls; [mean (*SD*) age = 13.0 (0.3) years; range: 12.5–13.3 years]). Six years was an appropriate lower age limit in the reference group, as the majority of children of this age were able to pronounce all the speech sounds correctly in their own dialect. The NH 6-year-olds were mainly recruited from kindergartens near the hospital, and the 13-year-olds were recruited from a primary school nearby.

Normal hearing was confirmed by pure-tone audiometry showing audiometric thresholds at 20 dB (HL) or better on frequencies between 125 and 8,000 Hz. We chose a level of uncertainty of 5 dB, according to the *SDs* of measured audiometric thresholds in a large group of NH listeners in a study by Engdahl et al. (2005). Thus, also children and

TABLE 1 | Simplified IPA chart displaying the speech sounds used in the NSRT.

	Place of articulation														
	Bilabial		Labiodental		Dental		Post-alveolar		Palatal		Velar		Glottal	
Manner of articulation	U	V	U	V	U	V	U	V	U	V	U	V	U	V
Stops	[p]	[b]			[t]	[d]					[k]	[g]		
Fricatives			[f]		[s]		[ʃ]		[j]				[h]	
Nasals		[m]				[n]						[ŋ]		
Lateral					[l]									

U = unvoiced; V = voiced.

FIGURE 1 | Simplified vowel chart displaying the long Norwegian vowels used in the NSRT, plotted according to the two first formant frequencies, F1 and F2 [modified after Kristoffersen, 2000 (2000, p. 17)].
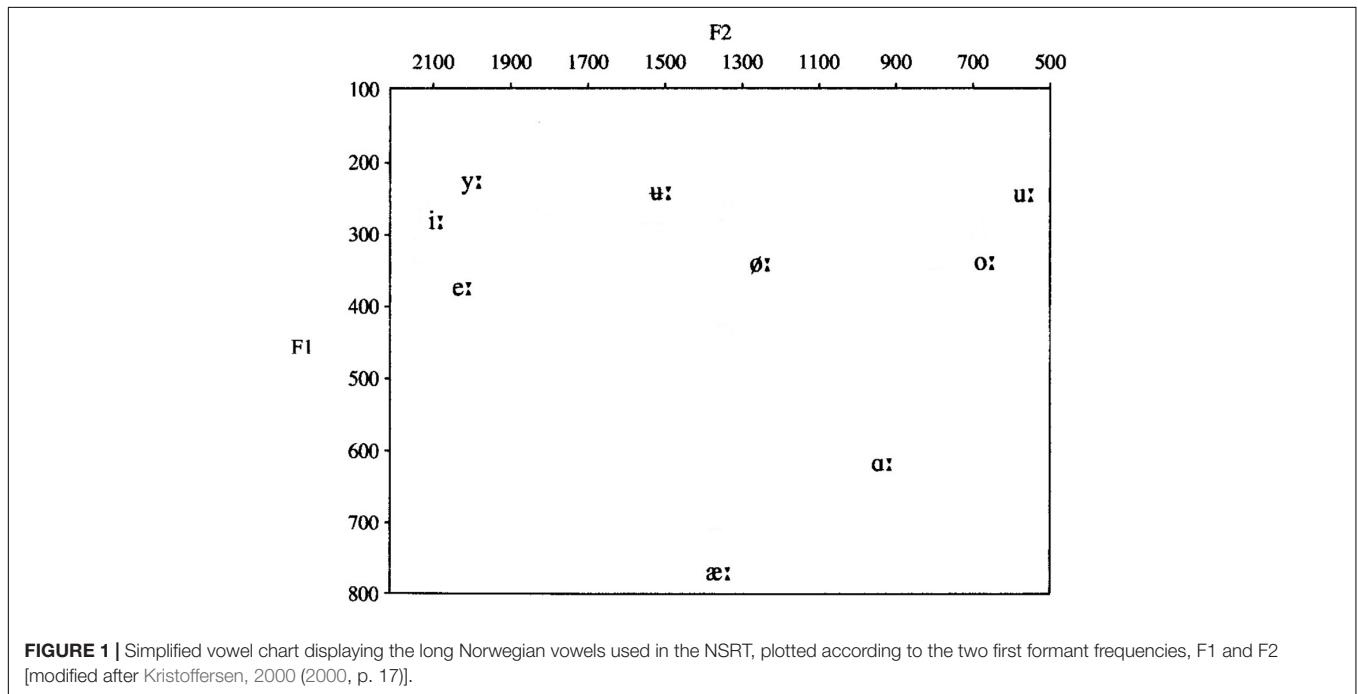
adolescents with hearing thresholds at 25 dB were included. The middle-ear status of the reference groups was checked with tympanometry and otomicroscopy by an ear, nose, and throat specialist before audiometry.

Inclusion Criteria for All Groups

All participants were required to have Norwegian as their native language and to obtain a 100% score on a pronunciation test of all the target speech sounds in the NSRT.

Test Descriptions

The Nonsense Syllable Repetition Test

The NSRT contains the 16 consonant sounds that are common for all Norwegian dialects, [p, t, k, s, ʃ, f, h, b, d, ɡ, j, v, n, m, ŋ, l], and 11 additional consonant sounds that are used in local Norwegian dialects. To avoid dialect background as a confounding factor in our study, only the first-mentioned 16 consonants were included in the analyses, as they were familiar to all participants. The consonants were placed in a bisyllabic VCV context with the three main cardinal vowels in Norwegian, /ɑː, iː,

u:/ (see **Supplementary Table S3**). **Table 1** presents a simplified IPA chart of the included consonants, classified by manner and place of articulation, and by voicing/non-voicing.

The NSRT also contains the nine Norwegian long vowels, [ɑː, eː, iː, uː, ʉː, yː, æː, øː, ɔː], presented in a monosyllabic CVC context with /b/ as the chosen consonant (see the vowel chart in **Figure 1** and an overview of the nonsense syllables in **Supplementary Table S3**).

None of the CVC or VCV combinations presented in the test had lexical meaning in Norwegian. Recording and preparation of the test was mainly done with the computer program Praat (Boersma and Weenink, 2018) and is described in **Supplementary Data Sheet S1** and Introduction provides the rationale for using a repetition test with nonsense syllables in an open-set design.

Real-Word Monosyllable Test

The perception of real-word monosyllables was measured by the HIST monosyllable test in free-field, a test with 50 Norwegian phonetically balanced words, which produces a percent score

TABLE 2 | List of acronyms and abbreviations.

Number	Abbreviation/ acronym	Meaning
1	ACE	Advanced combination encoder (stimulation strategy from Cochlear)
2	CI	Cochlear implant
3	CIS	Continued interleaved sampling (generic stimulation strategy)
4	CM	Confusion matrix
5	CVC	Consonant-vowel-consonant
6	F0, F1, F2	Fundamental frequency, first formant, and second formant
7	FSP/FS4/FS4-p	Fine structure processing (stimulation strategies from Med-El)
8	HA	Hearing aid
9	HIST	Høgskolen i Sør-Trøndelag (real-word monosyllable test)
10	HL	Hearing level
11	NH	Normal-hearing
12	NSRS	Nonsense syllable repetition score
13	NSRS-C	Nonsense syllable repetition score – consonants
14	NSRS-C$_{voi}$	Nonsense syllable repetition score – voiced consonants
15	NSRS-C$_{unvoi}$	Nonsense syllable repetition score – unvoiced consonants
16	NSRS-C$_{aCa}$	Nonsense syllable repetition score – consonants in the aCa context
17	NSRS-C$_{iCi}$	Nonsense syllable repetition score – consonants in the iCi context
18	NSRS-C$_{uCu}$	Nonsense syllable repetition score – consonants in the uCu context
19	NSRS-C$_{pre}$	Nonsense syllable repetition score – consonants repeated by prelingually deaf
20	NSRS-C$_{post}$	Nonsense syllable repetition score – consonants repeated by postlingually deaf
21	NSRS-V	Nonsense syllable repetition score – vowels
22	NSRS-V$_{pre}$	Nonsense syllable repetition score – vowels repeated by prelingually deaf
23	NSRS-V$_{post}$	Nonsense syllable repetition score – vowels repeated by postlingually deaf
24	NSRT	Nonsense syllable repetition test
25	PTA	Pure-tone average
26	REC	Regional ethical committee
27	T, T$_{max}$, T$_{rel}$	Speech transmission index (absolute, maximum, and relative)
28	TFS	Temporal fine structure
29	VCV	Vowel-consonant-vowel
30	VOT	Voice onset time

(Øygarden, 2009). The test words were presented at 65 dB(A), and 1 out of 12 lists was chosen.

Pronunciation Test

A sample of "Norsk fonemtest" (Norwegian test of phonemes; Tingleff, 2002) with 28 of its 104 pictures, was used to assess the participants' ability to pronounce all Norwegian consonants and vowels correctly. The selected test items presented the target

phoneme in the medial position to match their position in the NSRT. Only those who obtained a 100% score on this test were included in the study.

Procedure and Design

The test words were presented from a SEAS 11F-LGWD 4.5″ loudspeaker (Moss, Norway), in an anechoic chamber via the computer program SpchUtil, v. 5 (Freed, 2001). The hard disk recorder Zoom H4n (Hauppauge, NY, United States) was used to record the repeated test words and the naming of the pictures. The distance between the loudspeaker and the participants was 1.5 m, and the equivalent sound level in listening position was 65 dB(A).

Testing of Children and Adolescents With CIs

The NSRT was conducted by playing the recorded CVC and VCV nonsense syllables in randomized order and recording participants' verbal repetitions. The participants were exposed to auditory stimuli only and could not rely on lipreading. They were informed that words with no meaning would be presented to them, but they were not given any further details about how many, which words, and in which consonant or vowel context the speech sounds would be presented.

The participants were instructed to repeat what they heard and to guess if they were unsure, in order to achieve a 100% response rate. Each speech stimulus was presented only once, and the participants were not allowed to practice before being tested or provided with feedback during the testing.

The ecological validity of the testing was optimized by having the participants use the everyday settings of their speech processors instead of switching off front-end sound processing, which has been done in similar studies (e.g., Wolfe et al., 2011). The speech processors were quality checked before testing, and new programming was not performed prior to the testing.

Unaided pure-tone audiometry was performed to check for residual hearing, if these results were not present in the patient's file. Otomicroscopy was performed by an ear, nose, and throat specialist if the participant had residual hearing in one ear or if middle-ear problems were suspected.

Fifty HIST monosyllabic test words in free-field were conducted with all the participants with CIs.

Testing of Normal-Hearing Children and Adolescents

The test setup for the NH reference groups corresponded to that for the participants with CIs, except that the HIST monosyllable test was not conducted, because listeners with normal hearing typically perform at the ceiling level on this test.

Phonetic Transcription and Scoring

The recordings of the participants' repetitions were transcribed by two independent, trained phoneticians, who were blind to the purpose of the study and to what kind of participant groups they transcribed. The transcribers performed a broad phonetic transcription of the nonsense syllables in the test, including primary and secondary stress, and lexical tone, but not suprasegmentals.

The transcriptions of the two phoneticians were compared, and in the case of disagreement between the transcribers, the first author listened to the recordings and picked the transcription that he judged to be correct. The mean (*SD*; range) exact percent agreement between the two transcribers was 82.8 (6.6; 66.7–98.2)% for the participants with CIs and 89.2 (7.5; 68.4–100)% for the NH reference groups.

The repetitions of each target speech sound were scored as either correct (1) or incorrect (0). The total scores were calculated by dividing the number of correctly repeated responses by the total number of stimuli, for the consonants, averaged for the three vowel contexts (NSRS-C), for the vowels (NSRS-V), for the consonants in aCa, iCi, and uCu contexts (NSRS-C_{iCi}, NSRS-C_{aCa}, and NSRS-C_{uCu}), and for the voiced and unvoiced consonants averaged for the three vowel contexts (NSRS-C_{voi} and NSRS-C_{unvoi}). The consonant and vowel scores for the subgroups of prelingually and postlingually deaf were calculated by dividing the number of correctly repeated responses by the total number of stimuli for each subgroup (NSRS-C_{pre}, NSRS-C_{post}, NSRS-V_{pre}, and NSRS-V_{post}). The nonsense syllable repetition score (NSRS) was produced by calculating a weighted mean of NSRS-V and NSRS-C, in which the weights were determined by the number of different vowels (9) and consonants (16) in the test [NSRS = (NSRS-V × 9 + NSRS-C × 16)/25].

Analysis

The 12 variables mentioned in the previous section (#12–23 in **Table 2**) were constructed to score the performance on the NSRT for the three groups of participants, and means, medians, and standard deviations were calculated for all variables. The consonant speech features voicing, stopping, frication, nasality, and laterality were examined separately in the analyses. Assumptions of a normal distribution were violated due to checking of the data with the Shapiro–Wilk test, possibly due to a ceiling effect in some of the variables. Therefore, scores from the participants with CIs were compared by the non-parametric Wilcoxon signed rank z test for related samples, for the following variables:

- Voiced and unvoiced consonant scores (NSRS-C_{voi} and NSRS-C_{unvoi}).
- The HIST real-word monosyllable score and the NSRS.
- NSRS-C_{aCa}, NSRS-C_{iCi}, and NSRS-C_{uCu}.
- The consonant and vowel scores (NSRS-C and NSRS-V).
- Consonant and vowel scores for the pre- and postlingually deaf (NSRS-C_{pre}, NSRS-C_{post}, NSRS-V_{pre}, and NSRS-V_{post}).

Comparisons of NSRS-C and NSRS-V, and NSRS-C_{voi} and NSRS-C_{unvoi}, were also performed for the NH 6- and 13-year olds. Correlations were calculated with Spearman's rho (*ρ*).

Scores on all variables were compared between the CI users and the NH 6-year-olds, and between the NH 6-year-olds and the NH 13-year-olds, with the Mann–Whitney *U* test for independent samples. To determine statistical significance, an alpha (α) level of 0.05 was chosen for all tests.

Box-and-whiskers were used to display the score distribution for HIST monosyllables, NSRS-V, NSRS-C_{unvoi}, and NSRS-C_{voi} for the three participant groups (see **Figure 2**). All statistical analyses were performed by SPSS v. 24.0 (SPSS Inc., Chicago, IL, United States). A Holm-Bonferroni correction was used to correct for multiple comparisons in all the statistical tests.

Information Transmission for Subgroup Comparisons of Speech Sound Features

The speech sound confusions were organized into confusion matrices (CMs). The CM for the consonant confusions was submitted to an information transfer analysis. This method was introduced by Miller and Nicely (1955) and is an application of the information measure by Shannon (1948) to obtain data from a speech repetition task and measure the covariance of input and output in a stimulus-response system. The method produces a measure of mean logarithmic probability. The logarithm is taken to the base 2, and the measure can thus be called the average number of binary decisions needed to specify the input, or the number of bits of information per stimulus. The method has been used in a large number of studies of the speech sound perception of implantees (e.g., Tye-Murray et al., 1990; Tyler and Moore, 1992; Doyle et al., 1995; Sheffield and Zeng, 2012; Yoon et al., 2012).

The advantage of using this unit instead of recognition scores of correct and incorrect repetitions that are measured binarily is that the repetition errors within the same category of speech sounds obtain higher scores than repetition errors between different categories.

The information transmission (T) was calculated with the formula below:

$$T = -\sum_i \sum_j \frac{n_{ij}}{n} \log_2 \frac{\frac{n_i}{n} \frac{n_j}{n}}{\frac{n_{ij}}{n}}$$

Here, *i* and *j* are the stimulus number and response number (the column and row numbers of the CM, respectively), n_{ij} is the cell value, n_i is the row sum, n_j is the column sum, and *n* is the total sum.

The relative transmission, T_{rel}, is given by $T_{rel} = T/T_{max}$, in which T_{max} is the maximum transmission of information. T_{max} describes the transmission if all the speech sounds were repeated correctly and no stimulus/response pairs were missing, and *T* is the absolute transmission. T_{rel} was calculated for the speech sound feature contrasts voicing versus non-voicing, nasality versus non-nasality, frication versus stopping, and nasality versus the lateral [l] for the subgroups of the prelingually (N = 29) and postlingually (N = 7) deaf.

The information transmissions for the subgroups were compared by collapsing the CMs in **Table 6** and analyzing them by χ^2 statistics. Fisher's exact test was applied if the number in one of the quadrants in the 2 × 2 tables was lower than 5. Our null hypothesis was that the information transmission was equally large for both pre- and postlingually deaf participants. A histogram was constructed to visualize the transmission of speech sound features for the two groups (**Figure 3**).
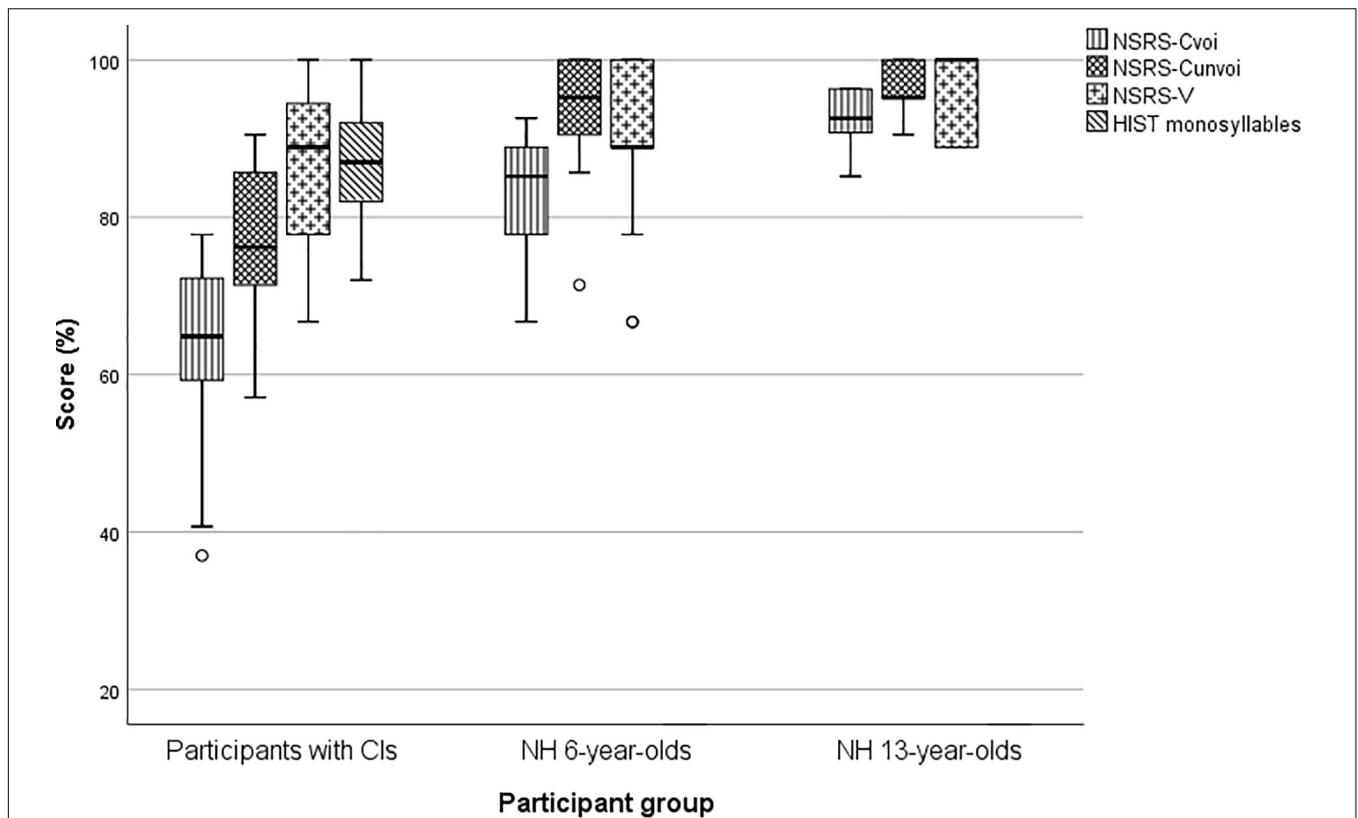
FIGURE 2 | Unvoiced and voiced consonant scores, vowel scores, and monosyllable scores for the three participant groups. The small circles are outliers that represent scores larger than 1.5 times the interquartile range of the box.
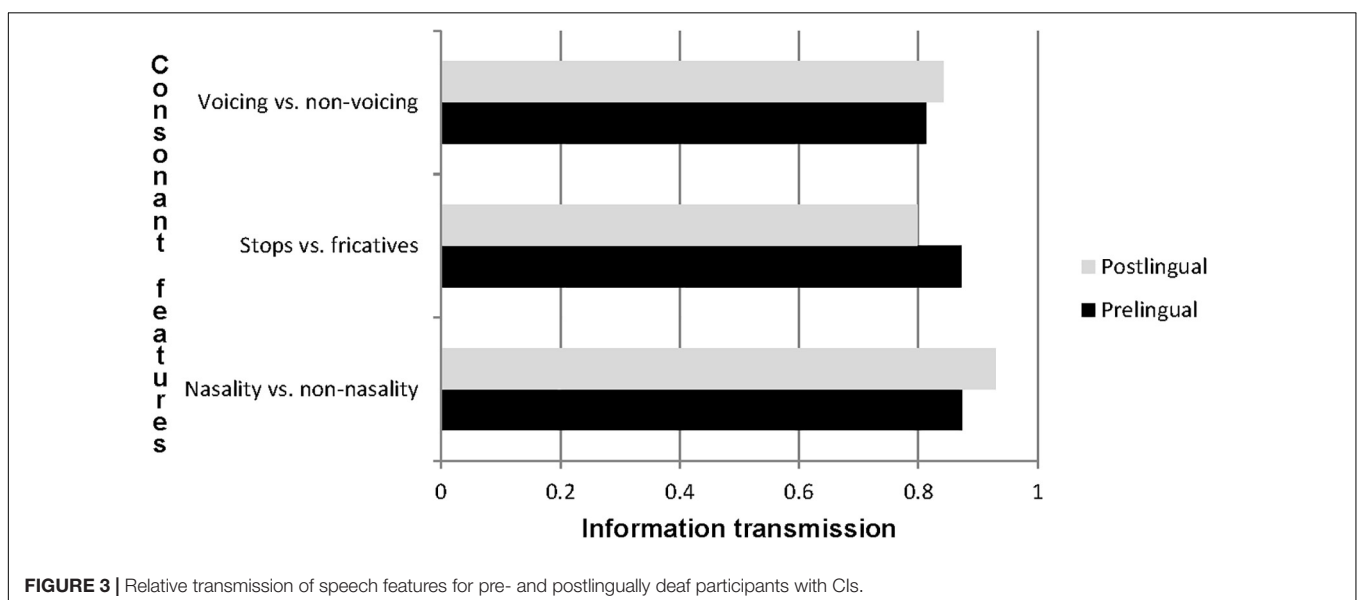
FIGURE 3 | Relative transmission of speech features for pre- and postlingually deaf participants with CIs.

RESULTS

Study Characteristics

The medians of the three groups of participants are displayed in **Table 3**, and comparisons of the participants with CIs and

the NH 6-year-olds, and of the NH 6-year-olds and the NH 13-year-olds with independent sample Mann–Whitney tests, are displayed in **Table 4**. The results show, as expected, that the NH 6-year-olds had significantly higher scores than the participants with CIs on all variables, except on the NSRS-V. The comparisons

TABLE 3 | *M, Md*, and *SD* of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

	CI users (*N* = 36)			NH 6-year-olds (*N* = 17)			NH 13-year-olds (*N* = 12)		
Variable (%)	*M* (*SD*)	*Md*	Range	*M* (*SD*)	*Md*	Range	*M* (*SD*)	*Md*	Range
NSRS	75.2 (8.0)	77.3	56.0–89.3	87.6 (5.8)	88.0	72.0–94.7	94.8 (2.0)	94.7	90.7–97.3
NSRS-C	69.6 (8.0)	70.8	50.0–83.3	86.9 (6.1)	87.5	72.9–93.8	94.4 (2.7)	95.8	89.6–97.9
NSRS-C$_{aCa}$	78.0 (8.6)	81.3	56.3–93.8	90.1 (7.3)	87.5	75.0–100	97.9 (3.1)	100	93.8–100
NSRS-C$_{iCi}$	69.3 (12.3)	71.9	25.0–87.5	89.3 (4.8)	87.5	81.3–100	96.4 (5.0)	100	87.5–100
NSRS-C$_{uCu}$	61.5 (13.1)	62.5	31.3–93.8	81.3 (12.1)	87.5	56.3–100	89.1 (3.9)	87.5	81.3–93.8
NSRS-C$_{voi}$	63.9 (10.6)	64.9	37.0–77.8	82.6 (7.5)	85.2	66.7–92.6	92.6 (3.5)	92.6	85.2–96.3
NSRS-C$_{unvoi}$	76.9 (9.3)	76.2	57.1–90.5	92.4 (7.5)	95.2	71.4–100	96.8 (3.1)	95.2	90.5–100
NSRS-C$_{pre}$	69.1 (7.8)	70.8	50.0–81.3	–	–	–	–	–	–
NSRS-C$_{post}$	71.4 (9.0)	70.8	56.3–83.3	–	–	–	–	–	–
NSRS-V	85.2 (10.9)	88.9	66.7–100	88.9 (11.1)	88.9	66.7–100	95.4 (5.7)	100	88.9–100
NSRS-V$_{pre}$	86.2 (10.1)	88.9	66.7–100	–	–	–	–	–	–
NSRS-V$_{post}$	81.0 (13.9)	88.9	66.7–100	–	–	–	–	–	–
HIST monosyllable score	86.9 (6.7)	87.0	72.0–100	–	–	–	–	–	–

of the medians of the NH 6- and 13-year-olds show a significantly higher score for the 13-year-olds for all variables except NSRS-C$_{uCu}$, NSRS-C$_{unvoi}$, and NSRS-V.

In **Table 5** the medians for the three groups of participants were compared with Wilcoxon's signed rank test and Mann-Whitney's U-test, and furthermore, correlations between the HIST score and NSRS-C$_{voi}$, NSRS-C$_{unvoi}$, and NSRS-V were shown. For the children with CIs, statistically significant differences were found for NSRS-V versus NSRS-C, NSRS-C$_{unvoi}$ versus NSRS-C$_{voi}$, NSRS-C$_{aCa}$ versus NSRS-C$_{iCi}$, and NSRS-C$_{aCa}$ versus NoSRS-C$_{uCu}$. No statistically significant differences were found for NSRS-C$_{iCi}$ versus NSRS-C$_{uCu}$, NSRS-C$_{pre}$ versus NSRS-C$_{post}$, and NSRS-V$_{pre}$ versus NSRS-V$_{post}$. For the NH participants, no statistically significant difference was found, except for the comparison of NSRS-C$_{unvoi}$ and NSRS-C$_{voi}$ for the NH 6-year-olds.

Consonant Confusions

Tables 6, **7** show the CMs for the 16 consonants in aCa, iCi, and uCu contexts for the 36 participants with CIs. The consonants are grouped primarily as voiced and unvoiced and secondarily according to manner of articulation. Of the consonant stimuli, 223 (12.9%) were repeated as consonant clusters or as consonants other than the ones listed in the CM and were excluded from the analyses. These are listed in the unclassified category of the CM.

The consonant CM in **Table 6** shows a devoicing bias for the stops. Unvoiced consonants are in general most frequently confused with other unvoiced consonants and voiced consonants are most frequently confused with other voiced consonants, except for the voiced stops, which are frequently repeated as unvoiced stops. Furthermore, there are highly populated clusters of correct repetitions around voiced and unvoiced stops, voiced and unvoiced fricatives, and nasals.

Table 7 shows that the highest proportion of correct repetitions was within the manner-groups of unvoiced fricatives; 90.5% of these were repeated as the same, or as another unvoiced fricative, and of unvoiced stops; 85.8% were repeated as the same, or as another unvoiced stop. Among the nasals, 81.2% were

repeated as the same, or as another nasal, among the voiced fricatives, 79.2% were repeated as the same, or as another voiced fricative, and among the voiced stops, 79.3% were repeated as the same, or as another voiced stop. The highest proportion of consonant confusions was found for the lateral [l], with a correct score of only 61.1%.

The correct repetition scores of the categories of speech features in **Figure 4** ranged from 60% to 80%, except for the nasals, which had a score slightly below 50%. The most common confusions were between consonants with the same manner and same voicing (Type 1 confusions). The least common confusions were between consonants with a different manner and opposite voicing (Type 3 confusions). The number of unclassified confusions, which includes consonant clusters and consonant sounds other than the stimuli, was also substantial, particularly for the lateral [l].

TABLE 4 | Comparisons of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

	CI users vs. NH 6-year-olds*				NH 6-year-olds vs. NH 13-year-olds**			
Variable (%)	*U*	*z*	*p*	*r*	*U*	*z*	*p*	*r*
NSRS	47.0	−4.94	<0.001	0.68	20.5	−0.64	<0.001	0.12
NSRS-C	23.0	−5.41	<0.001	0.74	19.5	−3.73	<0.001	0.69
NSRS-C$_{aCa}$	84.5	−4.30	<0.001	0.59	34.0	−3.17	0.002	0.59
NSRS-C$_{iCi}$	22.5	−5.47	<0.001	0.75	35.0	−3.12	0.002	0.58
NSRS-C$_{uCu}$	85.5	−4.25	<0.001	0.58	60.0	−1.96	0.050***	0.36
NSRS-C$_{voi}$	40.0	−5.10	<0.001	0.70	18.5	−3.76	<0.001	0.70
NSRS-C$_{unvoi}$	61.0	−4.70	<0.001	0.65	65.5	−1.69	0.091	0.31
NSRS-V	264.5	−0.83	0.404	0.11	68.5	−1.62	0.105	0.30

*The columns show the results of comparisons of means with the Mann–Whitney independent samples U-test between participants with CIs and NH 6-year olds. **The columns show the results of comparisons of means with the Mann–Whitney independent samples U-tests between NH 6- and 13-year-olds. ***The comparison was non-significant after adjusting for multiple testing. The medians and sample sizes that were used in the analyses can be found in **Table 3**.

TABLE 5 | Comparisons of the study variables for the participants with CIs, the NH 6-year-olds, and the NH 13-year-olds.

Comparison	Participant group	Statistical test	ρ	U	z	p	r
HIST vs. NSRS-C$_{unvoi}$	CI	S	0.26	–	–	0.13	–
HIST vs. NSRS-C$_{voi}$	CI	S	0.41	–	–	0.013*	–
HIST vs. NSRS-V	CI	S	0.18	–	–	0.31	–
HIST vs. NSRS	CI	W	–	–	−4.90	< 0.001	0.82
NSRS-V vs. NSRS-C	CI	W	–	–	−5.12	< 0.001	0.85
	NH6	W	–	–	−0.78	0.43	0.19
	NH13	W	–	–	−0.32	0.75	0.09
NSRS-C$_{unvoi}$ vs. NSRS-C$_{voi}$	CI	W	–	–	−4.46	< 0.001	0.74
	NH6	W	–	–	−3.15	0.002	0.76
	NH13	W	–	–	−2.60	0.009*	0.75
NSRS-C$_{aCa}$ vs. NSRS-C$_{iCi}$	CI	W	–	–	−3.96	< 0.001	0.66
	NH6	W	–	–	−0.18	0.86	0.04
	NH13	W	–	–	−0.97	0.33	0.27
NSRS-C$_{aCa}$ vs. NSRS-C$_{uCu}$	CI	W	–	–	−4.75	< 0.001	0.79
	NH6	W	–	–	−2.64	0.008*	0.64
	NH13	W	–	–	−2.99	0.003*	0.86
NSRS-C$_{iCi}$ vs. NSRS-C$_{uCu}$	CI	W	–	–	−2.76	0.006*	0.46
	NH6	W	–	–	−2.51	0.012*	0.61
	NH13	W	–	–	−2.72	0.006*	0.79
NSRS-C$_{pre}$ vs. NSRS-C$_{post}$	CI	M-W U	–	85.00	−0.66	0.51	0.11
NSRS-V$_{pre}$ vs. NSRS-V$_{post}$	CI	M–W U	–	80.00	−0.91	0.36	0.15

*CI = cochlear implant; NH6 = NH 6-year-olds; NH13 = NH 13-year-olds; S = Spearman's correlation test; W = Wilcoxon's signed rank test; M–W U = Mann–Whitney's U-test for independent samples. *Not significant after adjusting for multiple testing. The medians and sample sizes that were used in the analyses can be found in **Table 3**.*

TABLE 6 | Confusion matrix for 36 participants with CIs; consonants in the aCa, iCi, and uCu contexts added together.

			Response																	
			Unvoiced							Voiced										
			S			F				S			F		Na			L		
Stimulus			/p/	/t/	/k/	/s/	/ʃ/	/f/	/h/	/b/	/d/	/g/	/j/	/v/	/n/	/m/	/ŋ/	/l/	U	N
Unvoiced	S	/p/	86	6				10		3	1								2	108
		/t/		84	4						2								18	108
		/k/	5	4	89							1							9	108
	F	/s/				93	5	4											6	108
		/ʃ/				13	75												20	108
		/f/	1			14	13	73	4										3	108
		/h/					3	13	81		2			1					8	108
Voiced	S	/b/	13	1	1				1	66	11			4					11	108
		/d/		6							85	3							14	108
		/g/			9				1		2	90		2					4	108
	F	/j/							2		2		88						16	108
		/v/							1		1	1		83		1			21	108
	N	/n/													77	9	2		20	108
		/m/													29	66	2	1	10	108
		/ŋ/										1		43	21	16	4		23	108
	L	/l/								1	1	2						66	38	108
																		Total sum		1,728

S = stops; F = fricatives; Na = nasals; L = lateral [l]; U = unclassified speech sounds and consonant clusters.

TABLE 7 | Confusion matrix of consonant repetitions for participants with CIs, collapsed with regard to manner and place of articulation (percentage of correctly repeated stimulus features in each cell).

Stimulus			Response (%)															
			Unvoiced					Voiced										
			S	F				S	F		Na			L				
			/p/	/t/	/k/	/s/	/ʃ/	/f/	/h/	/b/	/d/	/g/	/j/	/v/	/n/	/m/	/ŋ/	/l/	U	Sum (%)	N

Stimulus			/p/ /t/ /k/	/s/ /ʃ/ /f/ /h/	/b/ /d/ /g/	/j/ /v/	/n/ /m/ /ŋ/	/l/	U	Sum (%)	N
Unvoiced	S	/p/ /t/ /k/	85.8	3.1	2.2				9.0	100	324
	F	/s/ /ʃ/ /f/ /h/	0.2	90.5	0.5	0.2			8.6	100	432
Voiced	S	/b/ /d/ /g/	9.3	0.6	79.3	1.9			9.0	100	324
	F	/j/ /v/		1.4	1.9	79.2 0.5			17.1	100	216
	N	/n/ /m/ /ŋ/				0.3	81.2	2.2	16.4	100	324
	L	/l/			0.9	0.9	1.9	61.1	35.2	100	108
									Total sum		1,728

S = stops; F = fricatives; Na = nasals; L = lateral [l]; U = unclassified speech sounds and consonant clusters.

FIGURE 4 | Percentages of correct consonant repetitions and of five types of consonant confusions for participants with CIs. The upper bar describes the complete material of consonant confusions and the eight bars below the horizontal line describe subsets of the material. The units on the horizontal axis are the percentage scores of correct and incorrect repetitions. The bars with a horizontal pattern visualize correct repetitions. Type 1 is confusion between consonants with the same manner and the same voicing. Type 2 is confusion between consonants with the same manner and the opposite voicing. Type 3 is confusion between consonants with a different manner and opposite voicing. Type 4 is confusion between consonants with a different manner and the same voicing. Type 5 is unclassified confusions.

The NH participants repeated almost all the consonants correctly, as shown in **Supplementary Tables S4, S5, S7**, and **S8**. However, we observed an important exception for the 6-year-olds: 10 (19.6%) of the /ŋ/ stimuli were confused with /m/. The 13-year-olds also had an unexpectedly high number of misperceptions of /ŋ/ (7; 19.4%).

Vowel Confusions

Only two cases of unclassified vowels were found among the nine vowels in the bVb context for the 36 participants with CIs (**Table 8**). An [i:]-[y:] perception bias was revealed; [y:] was more frequently repeated as [i:] (67%) than as [y:] (31%).

The CMs for the NH children and adolescents (**Supplementary Tables S6, S9**) show that almost all vowels were repeated correctly. The vowel CM for the 6-year-olds in **Supplementary Table S6** shows some randomly distributed errors, in addition to 6 (35%) of the /y:/ stimuli repeated as either /i:/ or /ʉ:/. There were fewer vowel misperceptions for the 13-year-olds than for the 6-year-olds, but even so, 3 (25%) of the /y:/ stimuli were repeated as /i:/, as displayed in **Supplementary Table S9**.

Perception of Consonant Features Compared by Information Transmission and Chi Square Statistics Between the Pre- and Postlingually Deaf

Figure 3 shows that nasality versus non-nasality had the highest information transmission, and voicing versus non-voicing had the lowest. The information transmission of speech features did not display large differences between pre- and postlingually deaf participants.

Chi square testing showed no statistically significant differences between the transmission of voicing and non-voicing ($\chi^2 = 1.16$; $p = 0.28$), nor between the transmission of nasality and non-nasality ($\chi^2 = 0.41$; $p = 0.52$), nor between the transmission of stops and fricatives ($\chi^2 = 1.12$; $p = 0.29$). **Supplementary Table S10** displays the three 2×2 matrices that these analyses are based on.

DISCUSSION

The objective of this study was to assess the effectiveness of CIs by obtaining a measure of speech sound confusions in well-performing children and adolescents with CIs, using an NSRT, and to investigate whether the perception of speech features differs between the pre- and postlingually deaf. The study was cross-sectional, and it included 36 participants with CIs and 2 reference groups (17 NH 6-year-olds and 12 NH 13-year-olds).

An important finding was that unvoiced consonants were significantly less confused than voiced consonants for the participants with CIs. Moreover, there was a devoicing bias for the stops; unvoiced stops were confused with other unvoiced stops and not with voiced stops, and voiced stops were confused with both unvoiced stops and other voiced stops. Another major finding was that there was no significant difference between the perception of speech sound features for pre- and postlingually deaf CI users.

A central issue when assessing consonant confusions in participants with CIs is to investigate the underlying reasons. Are the confusions caused by limitations in the implants, are they due to immature cognitive development, or can they be explained by other factors? The difference between the NSRS and the HIST real-word monosyllable score suggests that the participants with CIs rely substantially on their language proficiency and the top-down processing introduced by lexical content present in real-word stimulus material. The finding is in line with a study on NH individuals by Findlen and Roup (2011), who investigated dichotic speech recognition performance for nonsense and real-word CVC syllables, and found that performance with nonsense CVC syllables was significantly poorer. Findlen and Roup's study is to the authors' knowledge the only previous investigation of recognition differences between real-word and nonsense CVC syllable stimuli that have similar phonetic content but differ in lexical content.

The moderate correlation between NSRS-C_{voi} and HIST monosyllables suggests that problems with perceiving the real-word monosyllables could partly be explained by difficulties in perceiving the voiced consonants.

TABLE 8 | Confusion matrix of vowel repetitions in the bVb context for participants with CIs.

	Response										
Stimulus	/bɑ:b/	/be:b/	/bi:b/	/bu:b/	/bʉ:b/	/by:b/	/bæ:b/	/bø:b/	/bɔ:b/	U	N
/bɑ:b/	35									1	36
/be:b/		35	1								36
/bi:b/			36								36
/bu:b/				36							36
/bʉ:b/			2		30	4					36
/by:b/			24		1	11					36
/bæ:b/	1						35				36
/bø:b/		2	1		5	1		26		1	36
/bɔ:b/				1					35		36
										Total sum	324

U = unclassified.

The Results of the Participants With CIs Related to Those of the NH Reference Groups

As expected, the scores on the NSRT were higher for the NH 13-year-olds than for the NH 6-year-olds for all variables. However, the differences were not significant for NSRS-C$_{uCu}$, NSRS-C$_{unvoi}$, and NSRS-V, probably because NH 13-year-olds usually have a more developed phonemic lexicon and higher phonemic awareness, or because of age-related differences in attentiveness during the task. We compared the scores of the participants with CIs only to those of the NH 6-year-olds, as these two groups are closest in hearing age. Significant differences were found between the groups of NH 6-year-olds and CI users for all variables except for the NSRS-V, which was just as high for both groups. This may be due to the long duration and high energy of the vowels in the NSRT.

For the NH groups, there were no statistically significant differences in any of the comparisons, except for unvoiced versus voiced consonant score for the NH 6-year-olds. Since this difference was not found for the NH 13-year-olds, this can probably be explained by language immaturity and fatigue.

For the participants with CIs, the difference between voiced and unvoiced consonant scores seems to be mostly due to the fact that unvoiced stops in Norwegian, /p, t, k/, are strongly aspirated and hence have a substantially longer voice onset time (VOT)[2] than the voiced stops, /b, d, ɡ/ (Halvorsen, 1998). For both CI users and the NH 6-year-olds, the low, voiced consonant score is likely due to the nasals, /m, n, ŋ/, being confused with one another, and by /l/ having a low recognition score.

The Most Common Confusions of Consonants and Vowels for Participants With CIs

Most consonant confusions observed in the present study can be explained by acoustic similarity in manner and voicing, a conclusion that has also been reached in many previous studies (e.g., Fant, 1973; Dorman et al., 1997; Dinino et al., 2016).

A bias toward unvoiced stops was found, a phenomenon that only occurred for the CI group and hence probably is implant related. This may be related to two main issues: (1) implants convey the F0 in voiced sounds rather poorly due to missing temporal information in the electrical signal for most implant models and to the electrode's insertion depth possibly being too shallow to cover the whole cochlea (Hamzavi and Arnoldner, 2006; Svirsky et al., 2015; Caldwell et al., 2017) and (2) the VOT makes the unvoiced stops much easier to perceive than the voiced stops due to the aspirated pause between the stop and the following vowel in the VCV syllables.

The subgroups of voiced and unvoiced stops can be distinguished by the presence of a silent gap in the unvoiced stops (Lisker, 1981). For Norwegian unvoiced stops, as for unvoiced stops in most Germanic languages, aspiration is a salient feature: a distinct final auditory breathy pause that is created by closing

the vocal cords from a maximally spread position, lasting longer than the occluded phase of the stop articulation (Kristoffersen, 2000). Stops can be difficult to identify, since they are very short and unvoiced stops have little acoustic energy. In identifying stops, CI users usually rely considerably on the spectral properties of the surrounding vowels, such as locus and length of the formant transitions, spectral height and steepness, and VOT (Välimaa et al., 2002).

Moreno-Torres and Madrid-Cánovas (2018) found a voicing bias for the stops for children with CIs, which is the opposite of the results of the present study. Their study design is, however, considerably different from the present study, as the children were Spanish-speaking and were tested with added, speech-modulated noise, which may create a perception of voicing. Also, Spanish does not have aspiration as a salient feature of unvoiced stops, as Norwegian has. Studies with English and Flemish participants have found a devoicing bias similar to our study (e.g., van Wieringen and Wouters, 1999; Munson et al., 2003).

The least correctly repeated consonant was the lateral [l], which elicited many confusions in the unclassified category of the CMs and had the largest difference in correct scores between the participants with CIs and the NH 6-year-olds. Since all the NH participants were recruited from the same dialect area, Standard East Norwegian, many of them confused [l] with [ɭ], which is also part of their speech sound inventory. Remarkably, [l] was almost never confused with the nasals for any of the participant groups.

The nasals, [m, n, ŋ], were often confused with one another by the participants with CIs, and this – together with the [l]-confusions – comprise most of the difference between the NSRS-C$_{voi}$ and NSRS-C$_{unvoi}$. It seems that nasality adds a new obstacle to consonant recognition. This may be due to the prominence of low frequencies around 250 Hz in the nasals' spectrum; the nasal murmur, also called the nasal formant (F1). The CIs render low frequencies rather poorly compared to high frequencies (Caldwell et al., 2017; D'Alessandro et al., 2018). Perceptual experiments with NH listeners have shown that nasal murmur and the formant transitions are both important for providing information on place of articulation (e.g., Kurowski and Blumstein, 1984). The transitions of F2 are particularly important; [m] is preceded or succeeded by an F2 transition toward a lower frequency, [n] provides little transition change, and [ŋ] is preceded or succeeded by an F2 transition toward a higher frequency.

Although the NH 6- and 13-year-olds perceived almost all consonants and vowels correctly, they confused /ŋ/ with /m/ in 19.6 and 19.4% of the cases, respectively. This confusion was almost exclusively found in the uCu-context. The reason for this tendency might be twofold. First, the tongue body is very retracted for the Norwegian [u:], with a narrow opening of the mouth and in a position close to the tongue position of [ŋ], making the formant transition audibly indistinct. Second, the listeners might primarily be focused on recognizing letters when performing this type of task. There is no unique letter in Norwegian rendering the speech sound [ŋ], and participants may not on the spur of the moment consider this speech sound an alternative, and instead decide on the one that they find

---

[2]VOT is the time between air release and vocal-cord vibration.

acoustically more similar to the other nasals, [m] and [n], which both correspond to single letters of the alphabet.

The most prevalent vowel confusion for the participants with CIs was [y:] perceived as [i:]. The main reason for this confusion is probably that the F1s of these vowels are low (∼250 Hz) and almost coinciding, and the F2 of [i:] is only slightly higher than of [y:]. These vowels are thus closely located in the vowel chart in **Figure 1**. However, [i:] was never perceived as [y:], probably because [i:] in Norwegian is about 10 times more prevalent than [y:] (Øygarden, 2009, p. 108), and when in doubt, the participants would be likely to choose the most common of the two speech sounds.

Vowels are known to be more easily perceived than consonants, due to their combination of high intensity and long duration. Norwegian vowels are distinguishable by F1 and F2 alone, as opposed to vowels in other languages, which may also be distinguished by higher formants. Vowels are never distinguished by F0.

Comparison Between the Pre- and Postlingually Deaf Participants

Between the pre- and postlingually deaf participants, we found no significant differences for the consonant and vowel scores, and no significant differences for the speech feature contrasts voicing versus non-voicing, nasality versus non-nasality, and stopping versus frication. All but three participants were provided with CIs in their optimal ($N = 28$) or late ($N = 5$) sensitive period. Four of the prelingually deaf participants who received CIs in their late sensitive period had used bilateral HAs and developed language in the period between onset of deafness and implantation, and their auditory pathways had presumably been effectively stimulated in this period.

For postlingually deaf CI users, the vowel formants conveyed by the implant tend to be misplaced in the cochlea compared to its natural tonotopy. This may be a reason why acoustically similar vowels are easily confused for the CI users than for the NH listeners.

The mechanisms of brain plasticity and the consequences of age at onset of deafness (pre-, peri-, and postlingual) are important factors for both auditory and linguistic development. Buckley and Tobey (2011) found that the influence of cross-modal plasticity on speech perception ability is greatly influenced by age at acquisition of severe to profound (pre- or postlingual) deafness rather than by the duration of auditory deprivation before cochlear implantation. In our study, brain plasticity at implantation may be a more relevant prognostic factor for the development of speech perception skills than age at onset of deafness, because of the large individual variations in age at implantation and HA use before implantation.

The Impact of Vowel and Consonant Context on Recognition

The results of the perception of consonants in different vowel contexts indicated that formant transitions played a larger role for the participants with CIs than for the NH participants, since the influence of vowel context on the consonant score was statistically significant for the CI group but not for the NH groups. This is in accordance with Donaldson and Kreft (2006), who found that the average consonant recognition scores of adult CI users were slightly but significantly higher (6.5%) for consonants presented in an aCa or uCu context than for consonants presented in an iCi context. The vocal tract is more open for [ɑ:] than for [i:] and [u:], making the formant transition more pronounced and the consonants therefore more easily perceptible. The Norwegian [u:] is much more retracted than the English [u:], and thus closer to the velar speech sounds, making their formant transitions more challenging to perceive.

The nine long vowels were presented in only one consonant context, with /b/, as vowel perception is based on steady-state formants rather than on formant transitions.

Inclusion Criteria and Test Design

By only including well-performing participants with CIs (score above 50% on the HIST monosyllable test and 100% correct spontaneous pronunciation score of all the Norwegian speech sounds), we were able to reveal systematic details in speech sound confusions. If poorer-performing participants with CIs had been included, a great deal of noise would have been added to the CMs, as the unclassified category would have become much larger.

In the present study, other higher language skills are of minor importance, as the NSRT is limited to speech sounds and syllables. We therefore had no inclusion criterion regarding language skills. Since the participants with CIs and the NH 6-year-olds had a similar mean hearing age, some perception problems may be related to their developmental stage in speech perception ability, in addition to being implant related.

As our study required that the participants respond verbally, a closed-set test was not a practical option. Moreover, we consider an open-set test design to be more ecologically valid than a closed-set test design, as repetition of unknown syllables is a common activity for children and one with which they are familiar when acquiring new vocabulary in their everyday life.

Limitations and Strengths

As expected, we obtained ceiling effects on both the vowel and consonant scores for the NH reference groups. For the participants with CIs, there were ceiling effects only on the vowel scores. This explains lack of statistical significance in many of the comparisons, and is in line with previous studies. For instance, Rødvik et al. (2018) have shown that NSRTs rarely result in ceiling effects when measuring consonant perception for CI users but may do so for vowel perception. It is well known that vowels are easier to perceive than consonants, due to longer duration and higher intensity. All nine Norwegian vowels exist in a long and a short version, and in the NSRT, only long vowels were included, making them audibly very distinct.

An important reason for the ceiling effect on the vowel and HIST scores for the participants with CIs is probably our criterion of only including well-performing CI users who had scores above 50% on HIST. The ceiling effect on the HIST score has probably also weakened the correlations with consonant and vowels scores in the CI users.

Since the test lists of the NSRT counted as many as 90 CVC and VCV words, fatigue and lack of concentration may have influenced the results, especially for the younger children. We

randomized the word order to prevent the same words from always appearing at the end of the test list and thus avoiding systematic errors.

This study used a convenience sample due to a limited time window for recruiting participants, who were assessed in conjunction with their regular CI checkup. This design has limitations as far as internal matching regarding, for instance age, gender, age at onset of deafness, duration of implant use, age at implantation, or implant model is concerned. Using a convenience sample may, however, also be considered a strength, as the participants represent a completely random sample of Norwegian-speaking children with CIs, since all implanted children in Norway have received their CI at the same clinic, Oslo University Hospital.

The two groups of pre- and postlingually deaf participants are very different in size, and the participants are very different with regard to level of hearing loss after onset of deafness, HA use before implantation, and age at implantation. Ideally, these factors should have been controlled for, so the evidence present to compare these groups may therefore have been weak.

Recommendations for Future Research and Clinical Use

This study provides information regarding typical misperceptions of speech sounds in participants with CIs, which may be useful as a basis for further research, focusing on its consequences for CI programming. The information will also be very useful when planning listening and speech therapy for the implantees.

The study might also be used as a basis for the development, validation, and norming of a simplified version of the NSRT to be included in the standard test battery in audiology clinics. Children with CIs tested regularly with the NSRT would be provided with individual feedback on what needs to be targeted in the programming of their CIs and in their listening therapy sessions. Pre- and post-testing with the NSRT can be used as a quality control tool of the programming. A clinical NSRT would also meet the increasing challenge of assessing speech perception in patients with different language backgrounds, as it can be adjusted for different languages by modifying it to only include speech sounds existing in a particular language.

A close examination of the CMs of each individual CI user may possibly be employed when deciding whether to reprogram the CIs or simply adjust the approach in listening therapy, since speech sounds within the same manner-group in the CMs are in general more acoustically similar than speech sounds in different manner groups. Hence, a rule-of-thumb may be that in case of confusions within the same manner-group, start with listening therapy, and in case of confusions between two manner-groups, reprogramming of the implant may be useful.

CONCLUSION

For the participants with CIs, consonants were mostly confused with consonants with the same voicing and manner. In general, voiced consonants were more difficult to perceive than unvoiced consonants, and there was a devoicing bias for the stops. The vowel repetition score was higher than the consonant repetition score. Additionally there was a [i:]-[y:] confusion bias, as [y:] was perceived as [i:] twice as often as [y:] was repeated correctly.

The subgroup analyses showed no statistically significant differences between consonant repetition scores for the pre- and postlingually deaf participants.

Although the children with CIs obtained scores close to 100% on vowels and real-word monosyllables, none of them obtained scores for voiced consonants above 78%. This is likely to be related to limitations in CI technology for the transmission of low-frequency sounds, such as insertion depth of the electrode and ability to convey temporal information.

AUTHOR'S NOTE

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The study was approved by the Regional Ethical Committees for Medical and Health Research Ethics – REC South East, Oslo, Norway. This study was carried out in accordance with the recommendations of "helseforskningsloven" (Health Research Law) §9, §10, §11, and §33, and cf. "forskningsetikkloven" (Research Ethics Law) §4, approved by the REC South East, with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the REC South East. Additional considerations regarding vulnerable populations such as minors: the speech perception testing of the children included in the project implied no risk for them, and no additional measures were necessary.

AUTHOR CONTRIBUTIONS

AR designed the study, analyzed the data, and wrote the manuscript. OT was responsible for the analyses and for technical matters regarding the CI, JT was responsible for methodological, structural, and linguistic matters, OW was responsible for audiological and educational matters, IS was responsible for phonetic and speech therapeutic matters, and JS was responsible for study design and medical matters. All authors discussed the results and suggested revisions of the manuscript at all stages.

ACKNOWLEDGMENTS

REFERENCES

Arisi, E., Forti, S., Pagani, D., Todini, L., Torretta, S., Ambrosetti, U., et al. (2010). Cochlear implantation in adolescents with prelinguistic deafness. *Otolaryngol. Head Neck Surg.* 142, 804–808. doi: 10.1016/j.otohns.2010.02.016

Boersma, P., and Weenink, D. (2018). *Praat: Doing phonetics by computer [Computer program]. Version 6.0.37.* Available at: http://www.praat.org (accessed March 14, 2018).

Bruijnzeel, H., Ziylan, F., Stegeman, I., Topsakal, V., and Grolman, W. (2016). A systematic review to define the speech and language benefit of early (<12 months) pediatric cochlear implantation. *Audiol. Neurotol.* 21, 113–126. doi: 10.1159/000443363

Buckley, K. A., and Tobey, E. A. (2011). Cross-modal plasticity and speech perception in pre- and postlingually deaf cochlear implant users. *Ear Hear.* 32, 2–15. doi: 10.1097/AUD.0b013e3181e8534c

Caldwell, M. T., Jiam, N. T., and Limb, C. J. (2017). Assessment and improvement of sound quality in cochlear implant users. *Laryngoscope Investig. Otolaryngol.* 2, 119–124. doi: 10.1002/lio2.71

Ching, T. Y., Zhang, V. W., Flynn, C., Burns, L., Button, L., Hou, S., et al. (2018). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57, 70–80.

Clopper, C. G., Pisoni, D. B., and Tierney, A. T. (2006). Effects of open-set and closed-set task demands on spoken word recognition. *J. Am. Acad. Audiol.* 17, 331–349. doi: 10.3766/jaaa.17.5.4

Coady, J. A., and Aslin, R. N. (2004). Young children's sensitivity to probabilistic phonotactics in the developing lexicon. *J. Exp. Child Psychol.* 89, 183–213. doi: 10.1016/j.jecp.2004.07.004

D'Alessandro, H. D., Ballantyne, D., Boyle, P. J., De Seta, E., DeVincentiis, M., and Mancini, P. (2018). Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. *Ear Hear.* 39, 679–686. doi: 10.1097/AUD.0000000000000525

Darley, F. L., and Winitz, H. (1961). Age of first word: review of research. *J. Speech Hear. Disord.* 26, 272–290. doi: 10.1044/jshd.2603.272

Dettman, S. J., Dowell, R. C., Choo, D., Arnott, W., Abrahams, Y., Davis, A., et al. (2016). Long-term communication outcomes for children receiving cochlear implants younger than 12 months: a multicenter study. *Otol. Neurotol.* 37, e82–e95. doi: 10.1097/MAO.0000000000000915

Dinino, M., Wright, R. A., Winn, M. B., and Bierer, J. A. (2016). Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces. *J. Acoust. Soc. Am.* 140, 4404–4418. doi: 10.1121/1.4971420

Donaldson, G. S., and Kreft, H. A. (2006). Effects of vowel context on the recognition of initial and medial consonants by cochlear implant users. *Ear Hear.* 27, 658–677. doi: 10.1097/01.aud.0000240543.31567.54

Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. *J. Acoust. Soc. Am.* 102, 2403–2411. doi: 10.1121/1.419603

Doyle, K. J., Mills, D., Larky, J., Kessler, D., Luxford, W. M., and Schindler, R. A. (1995). Consonant perception by users of nucleus and clarion multichannel cochlear implants. *Am. J. Otol.* 16, 676–681.

Drullman, R. (2005). *D-7-1: Speech Recognition Tests for Different Languages Report FP6–004171 HEARCOM Hearing in the Communication Society.* Sankt Augustin: Fraunhofer Institute for Applied Information Technology.

Elley, W. B. (1992). *How in the World Do Students Read? IAE Study of Reading Literacy.* The Hague: International Association for the Evaluation of Educational Achievement.

Engdahl, B., Tambs, K., Borchgrevink, H. M., and Hoffman, H. J. (2005). Screened and unscreened hearing threshold levels for the adult population: results from the Nord-Trøndelag hearing loss study. *Int. J. Audiol.* 44, 213–230. doi: 10. 1080/14992020500057731

Fant, G. (1973). *Speech Sounds and Features.* Cambridge, MA: MIT Press.

Feng, G., Ingvalson, E. M., Grieco-Calub, T. M., Roberts, M. Y., Ryan, M. E., Birmingham, P., et al. (2018). Neural preservation underlies speech improvement from auditory deprivation in young cochlear implant recipients. *Proc. Natl. Acad. Sci. U.S.A.* 115, E1022–E1031. doi: 10.1073/pnas.171760 3115

Findlen, U. M., and Roup, C. M. (2011). Dichotic speech recognition using CVC word and nonsense CVC syllable stimuli. *J. Am. Acad. Audiol.* 22, 13–22. doi: 10.3766/jaaa.22.1.3

Fischer, B., and Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia* 10, 105–118. doi: 10.1002/ dys.268

Freed, D. (2001). *Speech Utility [Computer Program].* San Francisco, CA: House Ear institute.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., and Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory* 2, 103–127. doi: 10.1080/09658219408258940

Geers, A. E., Brenner, C., and Davidson, L. (2003). Factors associated with development of speech perception skills in children implanted by age five. *Ear Hear.* 24, 24S–35S. doi: 10.1097/01.aud.0000051687.99218.0f

Halvorsen, B. (1998). *Timing Relations in Norwegian Stops.* Ph. D dissertation. University of Bergen: Bergen.

Hamzavi, J., and Arnoldner, C. (2006). Effect of deep insertion of the cochlear implant electrode array on pitch estimation and speech perception. *Acta Oto Laryngol.* 126, 1182–1187. doi: 10.1080/00016480600672683

Harrison, R. V., Gordon, K. A., Mount, R. J., and Maurer, D. (2005). Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. *Dev. Psychobiol.* 46, 252–261. doi: 10.1002/dev.20052

Henry, K. S., and Heinz, M. G. (2012). Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nat. Neurosci.* 15, 1362–1365. doi: 10.1038/nn.3216

Kral, A. (2013). Auditory critical periods: a review from system's perspective. *Neuroscience* 247, 117–133. doi: 10.1016/j.neuroscience.2013.05.021

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Kristoffersen, G. (2000). *The Phonology of Norwegian.* Oxford: Oxford University Press.

Kurowski, K., and Blumstein, S. E. (1984). Perceptual integration of the murmur and formant transitions for place of articulation in nasal consonants. *J. Acoust. Soc. Am.* 76, 383–390. doi: 10.1121/1.391139

Leake, P. A., and Hradek, G. T. (1988). Cochlear pathology of long term neomycin induced deafness in cats. *Hear. Res.* 33, 11–33. doi: 10.1016/0378-5955(88) 90018-4

Leigh, J., Dowell, R., Briggs, R., and Dettman, S. (2013). Communication development in children who receive a cochlear implant by 12 months of age. *Otol. Neurotol.* 34, 443–450. doi: 10.1097/MAO.0b013e3182814d2c

Lisker, L. (1981). On generalizing the rapid-rabid distinction based on silent gap duration. *Haskins Lab. Status Rep. Speech Res.* 65, 251–259.

Litovsky, R. Y. (2015). "The human auditory system: fundamental organization and clinical disorders," in *Handbook of Clinical Neurology*, ed. G. G. Celesia (Amsterdam: Elsevier Science).

Liu, H., Liu, S., Kirk, K. I., Zhang, J., Ge, W., Zheng, J., et al. (2015). Longitudinal performance of spoken word perception in Mandarin pediatric cochlear implant users. *Int. J. Pediatr. Otorhinolaryngol.* 79, 1677–1682. doi: 10.1016/j. ijporl.2015.07.023

Locke, J. L. (1983). *Phonological Acquisition and Change.* London: Academic Press.

Maxon, B. A., and Hochberg, B. I. (1982). Development of psychoacoustic behavior: sensitivity and discrimination. *Ear Hear.* 3, 301–308. doi: 10.1097/00003446-198211000-00003

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* 27, 338–352. doi: 10.1121/1.1907526

Moreno-Torres, I., and Madrid-Cánovas, S. (2018). Recognition of Spanish consonants in 8-talker babble by children with cochlear implants, and by children and adults with normal hearing. *J. Acoust. Soc. Am.* 144:69. doi: 10.1121/1.5044416

Munson, B., Donaldson, G. S., Allen, S. L., Collison, E. A., and Nelson, D. A. (2003). Patterns of phoneme perception errors by listeners with cochlear implants as a function of overall speech perception ability. *J. Acoust. Soc. Am.* 113, 925–935. doi: 10.1121/1.1536630

Myhrum, M., Strøm-Roum, H., Heldahl, M. G., Rødvik, A. K., Eksveen, B., Landsvik, B., et al. (2017). Sequential bilateral cochlear implantation in children: outcome of the second implant and long-term use. *Ear Hear.* 38, 301–313. doi: 10.1097/AUD.0000000000000383

National Institute for Occupational Safety, and Health [NIOSH] (1996). *DRAFT: Criteria for a Recommended Standard: Occupational Noise Exposure Revised Criteria.* Cincinnati, OH: National Institute for Occupational Safety, and Health.

Niparko, J. K., Tobey, E. A., Thal, D. J., Eisenberg, L. S., Wang, N.-Y., Quittner, A. L., et al. (2010). Spoken language development in children following cochlear implantation. *J. Am. Med. Assoc.* 303, 1498–1506.

Olsho, L. W., Koch, E. G., Carter, E. A., Halpin, C. F., and Spetner, N. B. (1988). Pure-tone sensitivity of human infants. *J. Acoust. Soc. Am.* 84, 1316–1324. doi: 10.1121/1.396630

Øygarden, J. (2009). *Norwegian Speech Audiometry.* Ph. D Dissertation. Norwegian University of Science and Technology: Trondheim.

Rødvik, A. K., Torkildsen, J. V. K., Wie, O. B., Storaker, M. A., and Silvola, J. T. (2018). Consonant and vowel identification in cochlear implant users measured by nonsense words: a systematic review and meta-analysis. *J. Speech Lang. Hear. Res.* 61, 1023–1050. doi: 10.1044/2018_JSLHR-H-16-0463

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Techn. J.* 27, 379–423.

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002a). A sensitive period for the development of the central auditory system in children with cochlear implants: implications for age of implantation. *Ear Hear.* 23, 532–539. doi: 10.1097/00003446-200212000-00004

Sharma, A., Dorman, M. F., and Spahr, A. J. (2002b). Rapid development of cortical auditory evoked potentials after early cochlear implantation. *Neuroreport* 13, 1365–1368. doi: 10.1097/00001756-200207190-00030

Sharma, A., Spahr, A. J., Dorman, M. F., and Todd, N. W. (2002c). Early cochlear implantation in children allows normal development of central auditory pathways. *Ann. Otol. Rhinol. Laryngol.* 111, 38–41. doi: 10.1177/00034894021110s508

Sheffield, B. M., and Zeng, F. G. (2012). The relative phonetic contributions of a cochlear implant and residual acoustic hearing to bimodal speech perception. *J. Acoust. Soc. Am.* 131, 518–530. doi: 10.1121/1.3662074

Sommers, S. M., Kirk, I. K., and Pisoni, B. D. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal-hearing, and cochlear implant listeners. I: the effects of response format. *Ear Hear.* 18, 89–99. doi: 10.1097/00003446-199704000-00001

Svirsky, M. A., Fitzgerald, M. B., Sagi, E., and Glassman, E. K. (2015). Bilateral cochlear implants with large asymmetries in electrode insertion depth: implications for the study of auditory plasticity. *Acta Oto Laryngol.* 135, 354–363. doi: 10.3109/00016489.2014.1002052

Tingleff, H. (2002). *Norsk Fonemtest [Norwegian Phoneme Test].* Oslo: Damm.

Tobey, E. A., Thal, D., Niparko, J. K., Eisenberg, L. S., Quittner, A. L., and Wang, N.-Y. (2013). Influence of implantation age on school-age language performance in pediatric cochlear implant users. *Int. J. Audiol.* 52, 219–229. doi: 10.3109/14992027.2012.759666

Trehub, S. E., Schneider, B. A., Morrongiello, B. A., and Thorpe, L. A. (1988). Auditory sensitivity in school-age children. *J. Exp. Child Psychol.* 46, 273–285. doi: 10.1016/0022-0965(88)90060-4

Tye-Murray, N., Lowder, M., and Tyler, R. S. (1990). Comparison of the F0F2 and F0F1F2 processing strategies for the Cochlear Corporation cochlear implant. *Ear Hear.* 11, 195–200. doi: 10.1097/00003446-199006000-00005

Tye-Murray, N., Spencer, L., and Gilbert-Bedia, E. (1995). Relationships between speech production and speech perception skills in young cochlear-implant users. *J. Acoust. Soc. Am.* 98, 2454–2460. doi: 10.1121/1.413278

Tyler, R. S. (1990). Speech perception with the nucleus cochlear implant in children trained with the auditory/verbal approach. *Am. J. Otol.* 11, 99–107.

Tyler, R. S., and Moore, B. C. J. (1992). Consonant recognition by some of the better cochlear-implant patients. *J. Acoust. Soc. Am.* 92, 3068–3077. doi: 10.1121/1.404203

Välimaa, T. T., Määttä, T. K., Löppönen, H. J., and Sorri, M. J. (2002). Phoneme recognition and confusions with multichannel cochlear implants: consonants. *J. Speech Lang. Hear. Res.* 45, 1055–1069. doi: 10.1044/1092-4388(2002/085)

van Wieringen, A., and Wouters, J. (1999). Natural vowel and consonant recognition by Laura cochlear implantees. *Ear Hear.* 20, 89–103. doi: 10.1097/00003446-199904000-00001

Wie, O. B. (2010). Language development in children after receiving bilateral cochlear implants between 5 and 18 months. *Int. J. Pediatr. Otorhinolaryngol.* 74, 1258–1266. doi: 10.1016/j.ijporl.2010.07.026

Wilson, B. S., and Dorman, M. F. (2008). Cochlear implants: current designs and future possibilities. *J. Rehabilitat. Res. Dev.* 45, 695–730. doi: 10.1682/jrrd.2007.10.0173

Wolfe, C. J., Schafer, C. E., John, C. A., and Hudson, C. M. (2011). The effect of front-end processing on cochlear implant performance of children. *Otol. Neurotol.* 32, 533–538. doi: 10.1097/MAO.0b013e318210b6ec

World Health Organization [WHO] (2019). *Grades of Hearing Impairment.* Geneva: World Health Organization.

World Medical Association [WMA] (2017). *WMA Declaration of Helsinki-Ethical Principles for Medical Research Involving Human Subjects.* Ferney-Voltaire: World Medical Association.

Wouters, J., Francart, T., and Mcdermott, H. J. (2015). Sound coding in cochlear implants: from electric pulses to hearing. *IEEE Signal Process. Magaz.* 32, 67–80. doi: 10.1109/msp.2014.2371671

Yoon, Y. S., Li, Y., and Fu, Q. J. (2012). Speech recognition and acoustic features in combined electric and acoustic stimulation. *J. Speech Lang. Hear. Res.* 55, 105–124. doi: 10.1044/1092-4388(2011/10-0325)

Zeitler, D. M., Anwar, A., Green, J. E., Babb, J. S., Friedmann, D. R., Roland, T. Jr., et al. (2012). Cochlear implantation in prelingually deafened adolescents. *Arch. Pediatr. Adolesc. Med.* 166, 35–41.

Check for updates

The Others Are Too Loud! Children's Experiences and Thoughts Related to Voice, Noise, and Communication in Nordic Preschools

Anita McAllister[1,2]*, Leena Rantala[3] and Valdís Ingibjörg Jónsdóttir[4]

[1] CLINTEC, Division of Speech and Language Pathology, Karolinska Institutet, Solna, Sweden, [2] Functional Area Speech and Language Pathology, Karolinska University Hospital, Stockholm, Sweden, [3] Department of Logopedics, University of Tampere, Tampere, Finland, [4] "Thad er malid", Voice Pathology, University of Akureyri, Akureyri, Iceland

Background: High noise levels affect hearing, voice use, and communication. Several studies have reported high noise levels in preschools and impaired voice quality in children. Noise and poor listening conditions impair speech comprehension in children more than in adults and even more for children with hearing or language impairment, attention deficits, or another first language.

Aim: The aim of this study was to explore how children in Finland, Sweden, and Iceland describe the preschool environment in relation to noise, voice, and verbal communication; what were their experiences, knowledge and ideas in relation to voice, noise, and communication. Children's awareness of effects of noise, reactions, and coping strategies were also studied. In addition, country and gender differences were analyzed.

Methods: Eighteen Icelandic, 14 Finnish, and 16 Swedish children were interviewed using a common interview-guide. Swedish and Finnish children were interviewed in focus groups and Icelandic children individually. All interviews were transcribed verbatim and analyzed thematically by the native speaker. The interviews were translated to English to be re-analyzed for inter-judge reliability of identified themes. Inter-judge reliability was calculated using percentage absolute agreement.

Results: The interviews resulted in 1052 utterances, 471 from focus groups, and 581 from individual interviews. Three themes were identified, *Experiences, Environment*, and *Strategies* with two to three subcategories. Inter-judge agreement for the themes was excellent, 92–98%. Experiences occurred in 55% of the utterances. The subcategories were bodily and emotional experiences and experiences of hearing and being heard. Environment occurred in 20% of the utterances, with subcategories indoor vs. outdoor and noise. Strategies was found in 15%, with subcategories games and problem oriented actions. The only significant difference between the countries was for the theme Strategies where the Swedish children produced more utterances than the Finnish. No gender differences were found.

Conclusion: Children are aware of high noise levels and mainly blame other children for making noise and shouting. They describe reactions and strategies related to noise like impaired communication and effects on hearing but are less aware of effects on voice. Expressed thoughts were similar across countries. No gender differences were found.

Keywords: communication, experience, environment, strategies, risk factors, awareness, voice

INTRODUCTION

High background noise levels are well documented in preschools and schools (e.g., Sala et al., 2002; Shield and Dockrell, 2004; Sala and Rantala, 2016). Despite these environments being shared between children and adults, most studies have investigated effects of noise exposure only on teachers. Results from these studies report that high background noise levels affect general well-being (Kristiansen et al., 2013), stress (Basner et al., 2014), and of course communication and hearing (McKellin et al., 2007; Klatte et al., 2010). High background noise levels also increase vocal loudness for the speaker, known as the Lombard effect (Lane and Tranel, 1971). Increased vocal loudness increases vocal loading and reported subjective symptoms including vocal fatigue (Vilkman, 2004; Whitling et al., 2017). Long term, increased vocal loading may lead to vocal nodules (Szabo Portela et al., 2018) and impaired voice quality (Södersten et al., 2005; Ternström et al., 2006; Rantala et al., 2015; Szabo Portela et al., 2018). In preschool children, higher noise exposure also revealed an affected voice quality, with higher perceptual assessments of hoarseness, breathiness, and hyperfunction (McAllister et al., 2009). However, few studies have reported effects on children's speech and voice in relation to different settings (e.g., Sederholm, 1995; Sederholm et al., 1995; McAllister et al., 2009; Kallvik et al., 2015). Even fewer have reported on children's own perception of their soundscape.

In a field study of eleven 5-year old Swedish children from three preschools, voice use, and noise exposure were recorded using individually worn equipment, including two omnidirectional electret condenser microphones (TCM 110) at equal distance from the mouth and a DAT recorder. Mean background noise across children and preschools was 82.6 dB LAeq equivalent level, ranging from 81.5 to 83.6 dB LAeq for the three preschools (McAllister et al., 2009). Background noise was related to the children's activity and peaked during lunch time, where one preschool exceeded 85 dB LAeq based on four 1 h recordings. This is alarming even if only registered during lunch time and not during the whole day. In the EU safety directives for workers, hearing protection should be provided in environments with noise levels at or above 80 dB LAeq for 8 h (European Parliament, 2003). However, preschool children are not included in the directives for workers since preschool attendance is not mandatory.

In preschools language learning, communication, and other social activities take place. For this to happen both children and adults need to be able to talk and hear each other. Studies have shown that verbal communication is hampered already at fairly low noise levels (e.g., Sala et al., 2002; Bradley and Sato, 2008). An adult person perceives approximately 95% of running speech

produced at a distance of 1 m and 55 dBA background noise (ISO/TR 3352, 1974). Several studies have found that children are more impaired than adults by noisy listening conditions (Bradley and Sato, 2004; Klatte et al., 2010) and that their speech comprehension is more affected (Neuman et al., 2010). Thus, children require a better signal to noise ratio (SNR) than adults. Ratios between +6 dBA SNR and <0.5 s reverberation time (Crandell and Smaldino, 1995, 1996) to over +15 dBA SNR for the youngest children (Bradley and Sato, 2008) have been reported. Children with special needs may require even more favorable SNR and shorter reverberation times (American Speech-Language-Hearing Association, 2005). The group with special needs with regard to the needed SNR also include children with another first language (Tabri et al., 2011) since many studies have shown that non−native adult speakers have more difficulty perceiving speech in noise than native speakers (e.g., McAllister, 1990; Crandell and Smaldino, 1996; Rogers et al., 2006; Tabri et al., 2011). In a study including Swedish children learning English, Hurtig et al. (2016) reported fewer recalled words when presented in L2 compared to words presented in L1. Words presented with a high SNR (+12 dBA) improved recall compared to a low SNR (+3 dBA). Reverberation time interacted with SNR. At +12 dBA the shorter reverberation time improved recall, but at +3 dBA it impaired recall. Findings point to an increased cognitive load when perceiving L2 speech in noise. An increased cognitive load means that the listener needs to listen more attentively and that speech comprehension requires more effort. Pichora-Fuller et al. (2016) defines listening effort as "*the deliberate allocation of mental resources to overcome obstacles in goal pursuit when carrying out a task, with listening effort applying more specifically when tasks involve listening.*" Functional brain imaging reveals that the neural resources required to understand degraded speech extend beyond traditional language networks by including regions of the prefrontal cortex, premotor cortex, and the cingulo-opercular network (Peelle, 2018).

The noise in preschool is mainly activity noise. This means that the children, their speech and their activities constitute the main noise sources. According to preschool teachers, the noise levels are highest when children enter or leave the school/preschool, move from one place to another, eat lunch or play with hard toys (Jónsdóttir et al., 2015). Since children are the main noise source, they are also closer to the source compared to adults and naturally get a higher noise exposure. Adult height increases the distance to the floor and the noise source. The difference in height alone would correspond to approximately a 6 dB reduction in noise exposure, which corresponds well to reported mean noise levels at 82.6 dB LAeq in the study of child exposure (McAllister et al., 2009) and 76.1 dB LAeq based on recordings of preschool teachers (Södersten et al., 2002)

using the same individually worn equipment. Speech is a strong speech masker since it has a similar spectrum as the targeted speech (Lu and Cooke, 2008). In classrooms, other student's speech has been found to be the most disturbing noise (Boman and Enmarker, 2004).

Background noise may emanate from appliences in the building, from traffic outside the building or from the activities conducted in the building. Background noise levels caused by appliances in the school building should not exceed 28 dB LAeq or 33 dBLA$_{max}$ according to Finnish standard (SFS 5907, 2004) but most of the classrooms – 88% – fail this target (Sala and Rantala, 2016). Effects from activity noise is harder to monitor and varies more depending on noise type. The building material may dampen or amplify sound. A building with a lot of hard surfaces contribute to increased noise by reflecting sounds in a room, thus hampering speech perception and communication. The reflections of sounds in a room are measured in terms of reverberation time. In schools and preschools, favorable listening conditions are recommended and reverberation times should be between 0.4 and 0.6 s (Crandell and Smaldino, 2000) or 0.5–0.6 s according to the Finnish standard (SFS 5907, 2004). Since children's activities often are carried out closer to or on the floor reflections may be amplified.

Although unfavorable conditions for communication in schools and preschools are quite well documented, relatively little is known about how the children themselves perceive conditions in preschools in relation to noise, communication and voice. Interviews are frequently used to describe and explore a specific phenomenon (Malterud, 2009). Related to interviews with children fewer studies have been reported. The children's own thoughts on their daily environment could add potentially important information to teachers, other school personnel, and builders. During the last decade there has been a growing interest of capturing this information through interviews exploring children's own perception and reactions to road and aircraft noise (Haines et al., 2003), to noise in schools (Boman and Enmarker, 2004), noise, reactions to noise and communication in preschools (Dellve et al., 2013; Persson Waye et al., 2013) or speech disorders (Nyberg and Havstam, 2016) using individual or focus groups interviews.

The purpose of the present study was to interview preschool children from Finland, Sweden, and Iceland to increase our knowledge regarding children's own thoughts, perception and knowledge of noise, voice, and communication. We were also interested in investigating children's awareness of effects of noise and possible reactions to noise and to document if there were any differences between the three countries or depending on sex.

Ethical permission was obtained from the ethical board at Tampere University, Finland.

MATERIALS AND METHODS

A deductive research approach was used in the construction of the common interview guide (Williams et al., 2004). The guide included questions based on previous studies of adult voice ergonomic risk factors in learning spaces (e.g.,

Rantala et al., 2012), effects of noise and poor acoustics (e.g., Sala and Rantala, 2019) and also on the authors' collective clinical and research experiences involving preschool and school aged children (Jónsdóttir, 2002; McAllister et al., 2009; McAllister, 2019; McAllister and Simberg, 2019). The questions were open-ended and wording was adapted to match children's vocabulary and experiences (see **Supplementary Appendix 1**). When needed, follow-up questions and clarifications were added by the interviewer. The questions included the following main topics:

- Sounds and noise;
- Voice and perception of different voices and emotions;
- Difficulties hearing the teachers or peers and difficulties being heard by the teachers or their peers;
- Soundscape in different rooms indoors, outdoors;
- Communication related to rooms, activities and indoors vs. outdoors;
- Bodily reactions;
- Strategies used when it is noisy.

A letter about the project was sent out to the head of the preschools and when institutional participation was accepted, the teachers at the different preschools were informed. Eight preschools chose to participate, three in Finland and two each in Iceland and Sweden. An information letter was distributed to the preschools to be handed out to caregivers, and those who accepted gave a written informed consent for their child to participate. All children were age 5–6 years old, and had no known hearing, speech and language or other neurodevelopmental disorder. Eight children in the Swedish group had Swedish as their second language (L2), all other children were native speakers.

The number of children in the participating preschools were 65, 90, and 36 children in Finland (preschool 1–3, respectively), 57 and 63 (preschool 1 and 2, respectively) in Sweden and 108 and 148 (preschool 1 and 2, respectively) in Iceland. In all preschools children were divided into smaller groups of 16–21 children with three to four teachers/group. In one of the Finnish preschools, only 6-year-old children were enrolled. All other preschools had children varying from 1 to 6 years.

All preschools were situated in medium to large size university towns for the respective countries (Finland 230 000 inhabitants; Sweden 140 000, Iceland 18 000). All Finnish and Swedish preschools were runned by the city, in Iceland one was runned by the city the other was private. The preschool buildings all included at least one large gathering room and several smaller rooms for different play activities or doing arts and crafts. The outdoor play area had slides, swings, a sandbox, and a playhouse. Two Finnish preschools were on the first floor of apartment buildings and one Icelandic preschool was in a former church building. The other preschools were in buildings specifically designed for the purpose. No large roads were close to any of the preschools.

The socio-economical context of the Swedish preschools were middle class and low income, respectively with preschool 1 being in an area with below median income and preschool 2

TABLE 1 | Total number of participants divided into preschools, focus groups, and individual interviews related to country. Number of boys and girls and L2 speakers are also presented.

Country and preschools	Boys	Girls	n	No. of focus group interviews	No. of individual interviews	L2 speaker (Boys/Girls)
Finland total	5	9	14	3		0
Preschool 1	1	3				
Preschool 2*	3	3				
Preschool 3*	1	4				
Sweden total	5	11	16	3		8 (3/5)
Preschool 1	0	6				
Preschool 2	5	4				
Iceland total	9	9	18		18	0
Preschool 1	3	3				
Preschool 2*	6	6				
Total N	19	29	48	6	18	8

Preschools not built for the purpose.

in an area with somewhat above median income for region. The socio-economical context for the area of the Finnish and Icelandic preschools was middle class.

A total of 30 children were interviewed using focus groups with four to six children/group, 14 in Finland and 16 in Sweden. Focus group interviews of 18 children were also made in Iceland. However, the recordings were of poor quality and had to be discarded since large portions of the children's replies could not be transcribed. The focus groups were complemented by individual interviews of 18 children from two preschools in Iceland using the same interview guide, see **Table 1**. All interviews were done in a separate room at the preschool. The rooms were furnished with chairs around a table to facilitate eye contact during the interviews. The children in each group knew each other well which has been found to facilitate interaction (Gill et al., 2008) and we were aiming at collecting a broad description of children's experiences and thoughts on noise, voice and environment in the preschools. Following recommendations, especially for focus group interviews, each subject was discussed till no further comments or information were added by the children to ensure saturation in the subject (Charmaz, 2006).

The Finnish interviews were done with one moderator as part of a thesis project. One interviewer also carried out all the individual interviews with the children in Iceland. The Swedish focus group interviews were carried out by two moderators and speech-language pathology (SLP) students as part of their bachelor thesis. One moderator was active during the interview and one was the observer providing a second set of eyes and ears to increase the accumulation of information and to ensure validity of the analysis (Krueger and Casey, 2009). The observer handled the recording equipment and took notes during the interviews. All interviewers were certified SLPs or SLP students and all interviews were audio recorded. In Sweden a Tascam portable recorder, DR-40 and a Sennheiser microphone was used. The microphone was place on the table. In Finland a digital portable Zoom H2 recorder, with a built in microphone was used placed on the table. During the individual interviews in Iceland, an Olympus digital stereo

dictaphone with a built in microphone was used and placed in front of the child. Duration of the group interviews were between 30 and 45 min and between 20 and 35 for the individual interviews.

The interviews were transcribed verbatim by the students or the native speaking author and analyzed following recommendation for qualitative content analysis (Patton, 2002; Krippendorff, 2013). This means following a step-wise procedure starting with repeated readings of the transcripts to identify meaning units (Patton, 2002; Graneheim and Lundman, 2004). The meaning units were highlighted and commented in the document including first impressions and thoughts to obtain a multifacetted interpretation of the statement. Each meaning unit was condensed to reflect the main content. A single utterance could include several meaning units with different main content. In these cases utterances were split into several meaning units depending on content. The meaning units with a shared main content were grouped together. A thematic analysis of the content was made by the native speaking author and the themes were labeled to reflect included meaning units and utterances (Graneheim and Lundman, 2004). The meaning units where then further categorized into subcategories closer related to the specific topic addressed. During this process categories and themes were continuously discussed between the authors and reconsidered to ensure a trustworthy interpretation. Utterances produced by several children at the same time were excluded from the analysis since gender of the speaker could not be determined. Off topic utterances were counted but otherwise excluded since they did not contribute to the aim of the study.

All interviews were then translated to English in order to be re-analyzed by the other authors for inter-judge reliability of identified themes. All Swedish and Finnish transcriptions were re-analyzed and for the individual interviews 64% (374 utterances) were re-analyzed. Utterances categorized to a different theme were discussed between the authors to reach a final consensus. Inter-judge agreement of the thematic analyses across the three raters according to percentage absolute agreement was good to excellent varying between 92 and 98%.

Statistics

Number of utterances across countries and related to preschool buildings were analyzed using descriptive statistics. Non-parametric statistics was used throughout. Differences in total number of utterances for each theme was analyzed using the Friedman test. Wilcoxon Signed Rank test was used for a pairwise comparison of the three themes.

Kruskal–Wallis Test was used to analyze if the distributions of utterances across themes were different between the countries and the Mann Whitney U-test, for independent samples, to make a pairwise comparison of the number of utterances in each theme between countries. Differences in number of produced utterances between boys and girls was analyzed using Mann Whitney U-test.

In all statistical analysis, p values < 0.05 were considered indicating significant differences.

RESULTS

Mean number of utterances/child depended on how talkative a child was and also on interview method, with individual interviews generating more responses/child. The mean number of utterances in the focus groups was 11.4 (SD 5.4) for Finnish children and 19.4 (SD 19.4) for Swedish. There was a difference in mean number of utterances related to preschools in the focus groups, with the children in some preschools being somewhat more talkative [Swedish preschool 1 x̄ = 24 (SD 21.4), preschool 2 x̄ = 16.7 (SD 18.6); Finland preschool 1 x̄ = 15.7 (SD 7), preschool 2 x̄ = 9 (SD 2), and preschool 3 x̄ = 15.6 (SD 5.3)], however, the difference was not significant according to the Friedman test. For the individually interviewed Icelandic children mean number of utterances was 29.9 (SD 3.8). In the Finnish data, 10% of the children's answers were said in unison and 8% in the Swedish. These utterances were not included in the analysis since an individual speaker could not be identified.

Three themes were identified, *Experiences, Environment*, and *Strategies* with two to three subcategories each, see **Figure 1**. Inter-judge agreement of the thematic analyses across the three raters according to percentage absolute agreement was good to excellent varying between 92 and 98%.
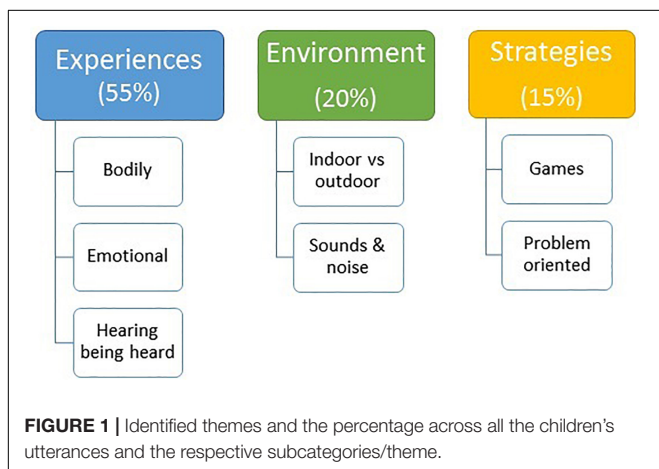
Themes and Subcategories

The most common theme was related to the children's own bodily and emotional experiences of noise, hearing and being heard themselves in the preschool. Thus *Experiences* made up a total of 55% of all utterances, followed by *Environment* and *Strategies* at 20 and 15%, respectively, see **Figure 1**. There was a significant difference regarding the number of utterances across themes (p = 0.000), with p-values varying from 0.004 to 0.000 in the pairwise comparisons according to the Wilcoxon Signed Rank test. Irrelevant utterances not related to the discussed topic were included in a category labeled *Other* that made up 10% of all utterances.

The number of utterances related to the preschool environment were also analyzed and compared related to the preschool buildings built for the purpose or not. In Finland there was a difference in number of utterances depending on the preschool building. Children in the preschools not original designed for the purpose commented more on the environment than other children. In Finland preschool 2 and 3 were in apartment buildings. Percentage utterances related to environment from children in preschool 2 and 3 were 19.1 and 34.3%, respectively (9 and 24 utterances) compared to only 6.7% (3 utterances) from children in preschool 1. In Iceland no such tendency was found with children in preschool 1 producing 20.9% of the utterances related to environment compared to 21.3% for children in the former church building (38 and 76 utterances, respectively). In Sweden both preschools were specifically built for the purpose. Utterances related to environment were 16.9 and 27.8%, respectively for preschool 1 and 2 (22 and 37).

No significant difference between countries were found for the distribution of utterances across themes except for the theme Strategies, where Swedish children produced significantly more utterances than the Finnish (p = 0.016; see **Figure 2**) according to the Kruskal–Wallis Test. Regardless of country, experiences

FIGURE 1 | Identified themes and the percentage across all the children's utterances and the respective subcategories/theme.

FIGURE 2 | Distribution of utterances in percent across the themes for the three countries. Swedish children produced significantly more utterances related to the theme strategies compared to the Finnish children, *p < 0.016.
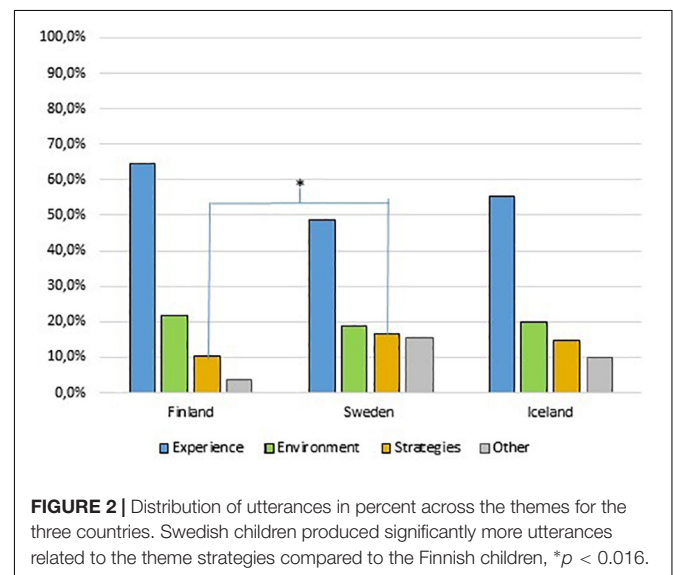
was the most common theme followed by environment and strategies, see **Figure 2**.

The themes could be further divided into two to three subcategories. For the theme experience they were bodily, emotional and experiences of hearing and being heard, for the theme environment subcategories were indoor vs. outdoor and sound and noise related to locality and for the theme strategies the subcategories were games and problem oriented. The distribution of responses across subcategories is shown individually for each country, see **Table 2**. Percentage of utterances/subcategory for each country is also presented.

No gender differences were found related to number of produced utterances in the different themes according to the Mann Whitney *U*-test, for independent samples, see **Figure 3**.

Below are some examples of utterances representing the main themes and subcategories.

Experiences

The children connected most of the bodily and emotional experiences and reactions of voice and noise to the throat and the ears. Most children said that voice comes from the mouth or the throat but some children suggested that voice is produced in the stomach.

Examples of utterances by the children related to their bodily and emotional experience of noise, hearing and being heard are given below. Several children expressed that noise was related to when you shout and play and make "a racket." Capital F (Finland), S (Sweden), and I (Iceland) indicates what country the child comes from. Information regarding the discussed topic is written within parenthesis. Sex of the child is indicated by G and B for girl and boy, respectively. Parentheses are added to visualize when the interviewer is asking a question, to specify the discussed topic or to provide a clarifying comment.

Bodily experiences

Bodily experiences and reactions were associated to noise and voice use. The children described how it felt in the body when speaking and shouting.

(F) G1: If I speak, then it kind of tickles a bit (in the mouth). . .and more often when using a loud voice.
(S) G2: (Noise is) Shouting. . . it hurts my ears.

Some children described having a sore throat after shouting or after a day at the preschool.
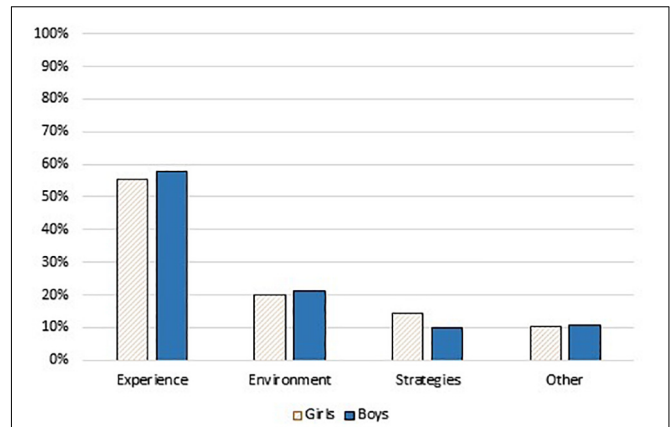
FIGURE 3 | Percentage utterances by sex across the themes.

(F) G1: If I shout and scream then my throat really hurts.
(S) G3: I . . . when I came home yesterday, I had to cough a lot, and then. . . then it felt like I had a lump in . . .
(S) G2: . . .the throat. (L2)

In a few cases the children also described bodily reactions related to other parts of the body like the "tummy." In some cases these utterances were a direct response to the question: *Can you feel it somewhere in your body that you have talked too much?*

(S) B1: eeh a tingling.in my tummy.
(F) G9: I feel it all the way to my legs.

Some children associated their bodily reactions to noise like these two children describing noise being painful to their ears.

(S) G2: (Noise is) shouting. . . it hurts my ears.
(I) B16: Noise and pain in my ears.

A few of children expressed a more detailed anatomical and physiological knowledge about voice production, like these two children.

(F) G4: (Voice comes) from muscles that kind of start to create that voice and then here in the throat it like starts to get ready to come here and then it goes to the vocal chords and from that spot it then goes to the mouth.

TABLE 2 | The distribution of responses across subcategories is shown individually for each country. Percentage of utterances/subcategory for each country is also presented.

		Finland	Sweden	Iceland
Experiences	Bodily	74 (46%)	67 (25%)	222 (42%)
	Emotional	15 (9%)	47 (18%)	22 (4%)
	Hear and being heard	18 (11%)	38 (14%)	78 (15%)
Environment	Indoor vs. outdoor	24 (15%)	32 (12%)	45 (9%)
	Sounds and noise	12 (7%)	27 (10%)	70 (13%)
Strategies	Games	1 (1%)	22 (8%)	18 (3%)
	Problem oriented	16 (10%)	30 (11%)	68 (13%)

(I) B16: (Voice comes) from. . . . . . .the tongue. . .the mouth. . . .not with the tongue as you could not talk then. From the mouth.

Seven of 18 Icelandic children (39%) said they had a sore throat sometimes when they came home from preschool and six of 16 Finnish (37%) children commented similarly. One Finnish boy (B2.2) told that he has lost his voice totally once or twice after having shouted or talked a lot. In the Swedish data four children (25%; see citations above) commented on a sore throat or a tingling in the tummy. However, the number of Swedish children having experienced a sore throat after preschool may be too low since several said they did not know what the term hoarseness or being hoarse meant. Then other children in the focus group helped explaining:

(S) B3: (it's) like this . . . sss ((sounds like a snake)).

Emotional experiences

The children's emotional experiences and reactions were often related to noise. Some felt bothered by the noise at the preschool and preferred the sound environment at home.

(F) B4: The noise bothers me a little bit.

The same child also mentioned that *It's nice when it's quiet*. This is similar to the opinion of a Swedish child (S) C5, that said *it feels good* (when it's quiet during gathering).

(F) G9: (I do) not really (like the sounds in the day care). I like it at home in the yard.

The replies in response to what the teacher does if s/he is being firm sometimes showed that some children interpreted this as being angry.

(S) G2: she shouts instead.
(S) G8: he is angry . . . a little.

Although several children said that noise bothered them and that they did not feel comfortable when it was noisy there were also children that said noise did not bother them.

(F) G4: The noise doesn't bother me at all.

Hearing and being heard

Several children mentioned that it sometimes was difficult both to hear others, including the teachers, and to make themselves heard at the preschool.

(I) G10: (Is it difficult for your teacher to hear you?) Yes the others are so noisy.

According to one child, a reason for the problem could be related to the teacher's voice use and how they raised their voices.

(I) G12: they (the teachers) shout sometimes so quietly.

In spite of the shared opinion of hearing difficulties due to high noise levels, children also connected the problem with other things such as "*bad hearing.*"

(F) G1: . . .the ear is stuck and there's really a lot of water in it. . . and then still if there's a lot of that dirt in the ear.

Fourteen of the 18 individually interviewed Icelandic children said it was difficult to hear what the other children said and 12/18 said it was difficult to hear the teacher in the preschool. Ten of 18 said they often had to repeat themselves to be heard.

(I) G9: Yes I just say again, "thank you" and "thank you."

Environment

Utterances related to different environments were mostly concerning indoors vs. outdoors and the different sound environments and noises depending on these settings. Some children also commented on specific activities or rooms and sounds in relation to that. A few utterances could be referred to both sub-categories under this theme.

Indoor vs. outdoor

(I) B5: Everyone talks so loud outside. Talk ordinary inside.
(S) G4: (there is more noise) inside.

The children also expressed that playing outside gave them the possibility to use their voices more freely.

(F) G9: I like that when we're outside we scream.
(S) B1: yeah we can shout outside but talk inside.
(I) B16: Outside then very loud. You are allowed to be noisy outside.

Sound and noise

The children sometimes had opposite opinions on where it was most noisy.

(I) G2: There is much noise outside.
(S) G4: (there is more noise) inside.
(S) G1: (during gathering) sometimes it's quiet like this ((just gestures)).

Several children associated noisiness with other children at the preschool. The typical comment was referring to others talking or shouting.

(F) B3: If others are talking then it's really noisy.
(I) G10: Yes, the others are so noisy.

In addition seven of the 18 Icelandic children said they found it better to talk inside. Most Icelandic children did not find the preschool environment noisy. Common answers were: G11: *Noisy but not high*; or B16: *Sometimes (noisy) just very little*.

Strategies

Utterances related to different strategies involved how the children described what they or the teachers did when it was noisy, including different games or actions directly aimed at reducing noise or improving verbal communication.

Games

In some preschools specific games were mentioned as ways of trying to control the noise.

(S) B4: sometimes we play the silent game.

In Swedish and Icelandic preschools children described a strategy to lower the noise level related to different activities involving whispering. From the childrens comments this strategy seemed to be used most often during lunchtime.

(S) B1: then you have to whisper And then we have to whisper...and the table that whispers the least wins and they they become whispermasters... today me and my best friend will be whispermasters.
(S) G5: during the silent game you have to be completely quiet but during whisper-lunch you have to whisper.
(S) G2: sometimes we sing doing signs.

In Iceland it was common for the teachers to ask the children to use their "*indoor voice*" when they were being too loud, according to some comments.

Problem oriented

Problem oriented strategies were actions directly aimed at an undesired behavior, improving communication or ways to avoid noise exposure. Typical strategies were to cover your ears or comments related to how you improve verbal communication.

(S) B3: if you do like this ((covers the ears)).
(S) B4: cause you cover your ears.
(F) G2: ...if someone is speaking, you don't talk over him/her and zip it (mouth).
(F) B5: If someone shouts then you must hush him/her.
(I) B7: By stopping talking so much.

One child had realized herself that she could rest her voice by taking a pause from speaking.

(F) G5: Often if I run really fast or speak really loud then I start to feel like I should stop for a while.

Another strategy could be to improve your own speech by talking loud and clear so others can hear you.

(I) G12: Just talk loud and clear.

In one Finnish preschool a traffic-light system also including a warning sound, had been implemented to alert everybody when noise levels were too high. The red light and the sound meant that you needed to lower your vocal loudness and be more quiet.

(F) G4: ...there're those kinds of traffic lights and you hear choo-choo-choo and then the red light turns on and it means that you gotta lower your voice.

In some cases the children expressed that the thing to do when it was noisy was to get a teacher that could calm the noisy children.

(S) G5: yah but we do it anyway (shout), I know one time, if, if someone wanted to tell the teachers that they should come and say something (to the noisy children).

Other

The category other was mostly made up by utterances that were regarded as irrelevant and off topic. However, in some cases,

especially among the Swedish L2 speakers, there were also clear misunderstandings. Below are two examples.

(S) G1: (Noise is) it's like eating (L2) (misunderstands buller as bullar, in English buns).
(S) G9: (Noise is) when you bake (L2).
(S) G1: (Voice is) that you should vote (L2) (misunderstands röst as rösta, in English vote).

Some Finnish children talked about "sound" instead of "voice" because the same word is used for both these concepts. In addition, two children talked about difficulties to understand foreign language when the topic of discussion was about difficulties hearing what the teachers' said.

In the individual interviews there were significantly more *I don't know* replies compared to the group interviews, in total 39 from the Icelandic children compared to two and one, respectively, in the Finnish and Swedish groups ($p = 0.004$ and 0.001, respectively) according to a Mann Whitney U-test, for independent samples. One Icelandic child contributed with 12 *I don't know* replies. A higher number can be expected in the individual interviews since all children needed to respond to all questions where in the focus groups those who felt sure about the concept replied. Twelve Icelandic children replied "*I don't know*" between 1 and 12 times. Mean was 2.2 times (SD 3.03).

DISCUSSION

In this study a total of 48 children were interviewed, in focus groups or individually in their preschools. The results show that the children's comments on sound and communication in the preschool was related to their own personal experiences of what they had seen, heard and felt. The results also revealed a budding awareness of high noise levels in the preschool and by describing effects on hearing and communication as well as strategies to avoid or decrease noise exposure. The children mostly described themselves or other children as the main noise source. Several children blamed the noise on other children playing and shouting. The children were less aware of effects of noise on voice but some had experienced a sore throat after preschool. Our findings are very similar to those of Dellve et al. (2013) also including focus group interviews of Swedish preschool children.

In everyday conversations, hearing is synonymous with understanding the content of what was said. This is also how the term to hear was interpreted in the interviews. However, when you hear a speaker talking a foreign language you can hear perfectly well but still not understand the content of the speech. This ambiguity of how the term is used among laymen was illustrated in some of the interviews, where the Icelandic children sometimes confused difficulties to hear what was said with difficulties understanding the content of the utterance and Finnish children who mentioned difficulties to understand foreign languages when discussing difficulties hearing what the teachers' said.

The most commonly identified theme in the children's utterances was related to the children's *Experiences* of noise and

difficulties hearing others and being heard by others as well as bodily and emotional experiences. For bodily reactions most children talked about ears hurting or a sore throat related to noise and loud voice use. However, also other body parts were mentioned like tingling in the mouth, tummy, or even legs. This is similar findings reported by Persson Waye et al. (2013). They interpreted the comments as a tendency for children to describe reactions to noise in a somatic way directly felt in the body (head, tummy) compared to adults. Another common response under this theme was that the children had problems hearing the teacher or other children. They also expressed that they sometimes had to shout and repeat what they said to make themselves heard and that the teachers shouted at times for the same reason. Most children in the individual interviews said they had experienced difficulties both hearing other children, the teachers and making themselves heard. Thus, the children have a potential awareness of effects of noise on communication. In a previous study on teachers use of amplification (WL 184 lapel condenser chest microphone combined with an amplifier and portable loudspeaker) over 95% of the participating children (6–9 years old) said that the use of amplification facilitated listening ("*I can hear better*"). They also asserted that "*the teacher does not shout as much*" and "*she is not so angry*" (Jónsdóttir, 2009). In the present study the children found it difficult to interpret the emotion of a teacher with a loud voice. In a noisy environment a loud voice is often necessary to make yourself heard. For the children a teacher's loud voice was often interpreted as angry. Brännström et al. (2015) reported similar findings that emotional content was more difficult both to convey and perceive in a noisy environment probably due to effects related to vocal loudness.

The children mostly blamed others than themselves for making noise and shouting. This shows an awareness of negative effects of noisy environments and that children, due to this awareness, they don't want to be blamed for being noise-makers. Like an Icelandic boy said, B18: *Yes sometimes I make noise just by accident*. This statement is well in line with what we all do automatically in a noisy setting, we increase vocal loudness (the Lombard effect; Lane and Tranel, 1971). Blaming other children for being noisy may also indicate that that listening conditions were generally unfavorable or that the high sound pressure levels impaired hearing and speech comprehension, possibly affecting the SNR required for good listening conditions for young children (Neuman et al., 2010). In the studied age-group children may need over +15 dBA SNR for good speech comprehension (Bradley and Sato, 2008) and children with special needs probably even more favorable conditions (American Speech-Language-Hearing Association, 2005). This includes children with another first language (Tabri et al., 2011). During the interviews two of the children with Swedish as their second language misunderstood phonetically similar words. Such misunderstandings were not found in the children speaking their first language. However, the larger number of *I don't know* replies in the individual interviews may also reflect a lack of understanding the questions or the terms used. The noisy conditions in the preschool could delay speech development and speech understanding due to the required SNR for good speech comprehension. Here also native speaking children may be at risk due to the frequent comment about

difficulties both hearing others and making themselves heard. In the long-term this may affect vocabulary and later reading and writing skills (Shield and Dockrell, 2008). In school children indoor noise and reverberation in classrooms were found to be associated with poorer performance in verbal tasks (Klatte et al., 2013). The findings point to the importance of good listening conditions for language learning and communication in children in general and especially for L2 speakers.

It was common to express a relation between noise and shouting but these utterances did not connect noise and shouting to having a sore throat. The habitual use of a loud voice during preschool hours may adapt children to this increased vocal loudness. On the basis of our clinical experience, parents often describe their child's speaking voice as being very loud when leaving preschool. This adaptation has been found also in teachers, where teachers working in loud background noise used louder voices already in the morning before work compared to those with classes with lower noise levels (Rantala et al., 2015). It seems reasonable to assume that vocal habits are established during childhood. Thus, undesirable and potentially straining speaking styles established during childhood may continue into adult life. The long term effects of maintaining a loud voice have not, as far as we are aware, been well documented scientifically even if several studies indicate a relationship between several vocal symptoms and vocally demanding professions (e.g., Fritzell, 1996; Roy et al., 2004). In the present study children sometimes commented on not being able to hear the teachers and, a few times, also on the teachers voices; (I) G12: *they* (the teachers) *shout sometimes so quietly*. The comment may indicate that this teacher had vocal fatigue and was unable to raise vocal loudness. Sala et al. (2001) found a strong association between the teaching profession and the 12 months prevalence of vocal fatigue. In a study on recovery after short term vocal loading in adults, patients with functional dysphonia were found to have slower recovery than the controls (Whitling et al., 2017). The recovery time for children after vocal loading, for example that during a day in a noisy preschool, has not been studied systematically. However, perceptual differences have been reported when comparing morning and afternoon recordings from children in preschool (McAllister et al., 2009) and in school children who had attended preschool and after school care compared with those who had not (Sederholm et al., 1995). This could imply a habituation to loud voice use as a long term effect of attending preschool.

The three themes *Experiences, Environment*, and *Strategies* were found in all interviews and no significant differences were found regarding the number of utterances between the three countries or depending on interview method except in the theme Strategies where the Swedish children had significantly more utterances than the Finnish. However, despite this the strategies to control noise by games and other actions that emerged between the two countries were very similar. Thus, the different rate of utterances regarding strategies was probably incidental. The off topic utterances were collected under the label other. Here utterances not connected to the discussion or misunderstandings were placed. The topics brought up included ballet dancing, clothes and traveling but has not been further analyzed.

The strategies the children describe were mostly related to different actions intended to lower the noise levels. The "silent game" and the "whisper lunch" both aim at this. The whisper lunch may do just that but to whisper is also quit a straining speaking technique that should be used with some caution. Other utterances included to *go get the teacher* when it was too noisy or describing that you need to *be quiet yourself* or *tell others to quiet down* to be able to listen properly. Some children described how to avoid noise by *covering their ears*. This is similar to findings by Dellve et al. (2013) also using focus group interviews. The children were also able to compare different environments with varying amounts of noise. These comparisons usually involved the preschool and the home environment *(I do) not really (like the sounds in the day care). I like it at home in the yard.* Several children felt bothered by the noise and preferred when it was quiet. Still, some said that noise did not bother them.

In a recently published study including children 9–13 years, findings suggested that noise conditions in crowded spaces are most challenging (Brännström et al., 2017). They also found that the extent of annoyance caused by the noise was task dependent, with tasks with high demands on verbal processing being more affected. Based on the noise mentioned most often in the present study, other children playing or shouting and teachers shouting seems to be the most disturbing noise for preschool children. This typically takes place in crowded spaces like the lunch room or the play hall. The reported annoyance related to verbal processing could also be linked the difficulties hearing others that was often described by the preschool children in the present study.

The children had more knowledge about noise and communication than we expected. They were aware that noise affects hearing and expressed difficulties both hearing others and making themselves heard. However, the connection between speaking in a noisy environment and having a soare throat was generally not made. This seems to reflect a knowledge gap regarding the potentially harmful effects of speaking in noise. Although the children were able to reflect on their preschool sound environment, some, especially the individually interviewed Icelandic children, responded "*I don't know*" quite often. Often these replies were following questions on voice or voice use. These replies might mirror a lack of knowledge or an insecurity about the topic. A higher number of such replies could be expected in the individual interviews since all children needed to respond to all questions even if they were not sure. In the focus groups the total number of these replies was one or two, respectively for Sweden and Finland.

Methodological Considerations

The interview guide was designed based on previous studies of adults (e.g., Rantala et al., 2012; Sala and Rantala, 2019) also including the authors' collective clinical and research experience of studies involving children (Jónsdóttir, 2002; McAllister et al., 2009; McAllister, 2019; McAllister and Simberg, 2019). The questions were adapted to children by using a simpler language and, if needed, terminology was explained further. Most questions were open ended in order to provide longer responses and start a discussion among the children but a few were direct dicotomic yes/no questions. However, children mostly

answered with longer utterances also to these questions showing an unexpected competence and ability to reflect.

The focus group interviews and the individual interviews provided an extensive material that allowed the children themselves to voice their opinions, perception, and knowledge on noise, voice and communication. Since they all knew each other well most children participated and contributed to the group interviews but the number of utterances varied with one or two children in each group contributing only a limited number. The individual interviews were included to amend the possible effect of more talkative and outgoing children's opinions that could dominate responses in the focus groups.

In the group interviews, the children sometimes had a tendency to repeat what another child had just uttered. This phenomenon called "other repetition" does not always mean that a child just imitates the friend. It has been shown that these repetitions have several functions in children's conversational discourses, such as affirming, agreeing with the other speaker, making matching and counter-claims (Keenan, 1975; Huang, 2010). Other repetition is very typical among preschool children while talking together (Karjalainen, 1996). In our study, some children's opinions undoubtedly were adopted from their peers'. Still, these utterances seemed to reflect also the repeating child's own perceptions and views.

The effects of the different interview methods on the results, if any, are difficult to assess. No differences were found related to interview method apart from significantly more "I don't know" replies given during the individual interviews. Nonetheless, the individual interviews confirmed observations in the focus groups and added information regarding how many children this applied to. One such example is how Icelandic children mentioned difficulties both hearing others (13/18), including the teachers (12/18), and being heard themselves (14/18). They also mentioned that they often had to repeat themselves (10/18). This was also found in the focus groups but how common it was could not be established.

There were no measurements made of the participating preschools regarding noise levels with children present or empty, nor regarding reverberation times. Both these measures would have added potentially important information on background noise and acoustic properties of the preschools but this was not the focus of the present study. Two included preschools were located in apartment buildings. There was a clear tendency for children in these preschools to talk about the environment more than other children. None of the included children had special needs known to the parents or preschool teachers at the time of the interviews. In the Swedish group, eight children with Swedish as a second language were included. More children with Swedish, Finnish or Icelandic as their second language could have added information regarding their specific difficulties in a noisy setting.

Practical Implications and Future Studies

Practical implications of the present results are the need of an increased awareness and knowledge regarding the effects of noise in preschools. The findings also point to a knowledge gap regarding how high noise levels affect voice use. Considering children's learning potential and curiosity, an adapted education

material for preschool children is surely needed. In future studies it would be interesting to study pedagogical effects on noise and communication and to include more L2 speakers to study possible effects of noise on language learning, vocabulary and comprehension using focus groups. It would also be interesting to interview other potentially vulnerable groups such as children with language disorders, attention deficits or cognitive impairment to study their thoughts, comprehension and reactions to noise and effects on communication. The differences in number of responses from children in preschools in apartment buildings may point to a need to study effects of environment on preschool children in more detail. Several factors may contribute since figures varied also between children in preschools built for the purpose.

CONCLUSION

Children are aware of high noise levels and blame other children for making noise and shouting. They describe reactions and strategies related to noise. They are aware of impaired communication in noise and effects on hearing but less aware of effects on voice. The experiences of children from three Nordic countries are quite similar possibly reflecting a shared cultural background. In addition, girls and boys describe their preschool sound environment and difficulties related to communication alike.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

Ethical permission was obtained from the ethical board at Tampere University, Finland. A letter about the project was sent

to the head of preschools and when institutional participation was accepted, the teachers at the different preschools were informed. Seven preschools chose to participate, three in Finland and two each in Iceland and Sweden. An information letter was distributed to the preschools to be handed out to caregivers, and those who accepted gave a written informed consent for their child to participate.

AUTHOR CONTRIBUTIONS

All authors contributed to the common interview guide and the data collection in collaboration with SLP-students, Erica Domej, Malin Eriksson, and Liisa Petäjistö, and with SLP Gudrun Sigurdardottir, analyzed the data, and discussed and provided the comments on the manuscript. Statistical analysis was mainly carried out by LR with the support of AM. Manuscript writing was mainly carried out by AM with the support of LR.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.01954/full#supplementary-material

REFERENCES

American Speech-Language-Hearing Association. (2005). *Guidelines for Addressing Acoustics in Educational Settings*. Rockville, MD: American Speech-Language-Hearing Association.

Basner, M., Babisch, W., Davis, A., Brink, M., Clark, C., Janssen, S., et al. (2014). Auditory and non-auditory effects of noise on health. *Lancet* 383, 1325–1332. doi: 10.1016/s0140-6736(13)61613-x

Boman, E., and Enmarker, I. (2004). Factors affecting pupils noise in schools: the building and testing models. *Environ. Behav.* 36, 207–228. doi: 10.1177/001391650325664

Bradley, J. S., and Sato, H. (2004). Speech recognition by grades 1, 3 and children in classrooms. *Can. Acoust.* 32, 26–27.

Bradley, J. S., and Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *J. Acoust. Soc. Am.* 123, 2078–2086. doi: 10.1121/1.2839285

Brännström, K. J., Holm, L., Lyberg-Åhlander, V., Haake, M., Kastberg, T., and Sahlén, B. (2015). Children's subjective ratings and opinions of typical and dysphonic voice after performing a language comprehension task in background noise. *J. Voice* 29, 624–630. doi: 10.1016/j.jvoice.2014.11.003

Brännström, K. J., Johansson, E., Vigertsson, D., Morris, D. J., Sahlén, B., and Lyberg-Åhlander, V. (2017). How children perceive the acoustic environment of their school. *Noise Health* 19, 84–94. doi: 10.4103/nah.NAH_33_16

Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Thousand Oaks, CA: Sage Publications.

Crandell, C. C., and Smaldino, J. (1995). The effects of room acoustics on normal-hearing children: implications for intervention. *J. Acoust. Soc. Am.* 97, 3262–3262. doi: 10.1121/1.411633

Crandell, C. C., and Smaldino, J. J. (1996). Speech perception in noise by children for whom english is a second language. *Am. J. Audiol.* 5, 47–51. doi: 10.1044/1059-0889.0503.47

Crandell, C. C., and Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Lang. Speech Hear. Serv. Sch.* 31, 362–370. doi: 10.1044/0161-1461.3104.362

Dellve, L., Samuelsson, L., and Persson Waye, K. (2013). Preschool children's experience and understanding of their soundscape. *Qual. Res. Psychol.* 10, 1–13. doi: 10.1080/14780887.2011.586099

European Parliament. (2003). Directive 2003/10/EC on the minimum health and safety requirements regarding exposure of workers to the risk arising from physical agents (noise). *OJEC* L 042, 38–44.

Fritzell, B. (1996). Voice disorders and occupations. *Logop. Phoniatr. Voco.* 2, 7–12. doi: 10.3109/14015439609099197

Gill, P., Stewart, K., Treasure, E., and Chadwick, B. (2008). Methods of data collection in qualitative research: interviews and focus groups. *Br. Dent. J.* 204, 291–295. doi: 10.1038/bdj.2008.192

Graneheim, U. H., and Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Educ. Today* 24, 105–112. doi: 10.1016/j.nedt.2003.10.001

Haines, M. M., Brentnall, S. L., Stansfeld, S. A., and Klineberg, E. (2003). Qualitative responses of children to environmental noise. *Noise Health* 5, 19–30.

Huang, C. C. (2010). Other-repetition in Mandarin child language: a discourse-pragmatic perspective. *J. Pragmat.* 42, 825–839. doi: 10.1016/j.pragma.2009.08.005

Hurtig, A., Keus van de Poll, M., Pekkola, E. P., Hygge, S., Ljung, R., and Sörqvist, P. (2016). Children's recall of words spoken in their first and second language: effects of signal-to-noise ratio and reverberation time. *Front. Psychol.* 6:2029. doi: 10.3389/fpsyg.2015.02029

ISO/TR 3352 (1974). *Acoustics – Assessment of Noise with Respect to its Effect on the Intelligibility of Speech*. Switzerland: International Organization for Standardization.

Jónsdóttir, V. (2009). *Is Amplification Necessary in a Classroom?*. Lissabon: Internoise.

Jónsdóttir, V., Rantala, L. M., Oskarsson, G. K., and Sala, E. (2015). Effects of pedagogical ideology on the perceived loudness and noise levels in preschools. *Noise Health* 17, 282–293. doi: 10.4103/1463-1741.165044

Jónsdóttir, V. I. (2002). Cordless amplifying system in classrooms. A descriptive study of teachers' and students' opinions. *Logop. Phoniatr. Voco.* 27, 29–36. doi: 10.1080/140154302760146952

Kallvik, E., Lindström, E., Holmqvist, S., Lindman, J., and Simberg, S. (2015). Prevalence of hoarseness in school-aged children. *J. Voice* 29, 260.e1–260.e19. doi: 10.1016/j.jvoice.2013.08.019

Karjalainen, M. (1996). *There is a Monster Here, in Play. Analysis of Three- and Four-Year-Old Preschool Children's Conversations in Terms of Acts, Turns, Topics and Special Features*. Dissertation, University of Oulu: Oulu.

Keenan, E. O. (1975). Making it last: repetition in children's discourse. *Annu. Meet. Berkeley Ling. Soc.* 1, 279–294. doi: 10.3765/bls.v1i0.2336

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4:578. doi: 10.3389/fpsyg.2013.00578

Klatte, M., Lachmann, T., and Meis, M. (2010). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise Health* 12, 270–282. doi: 10.4103/1463-1741.70506

Krippendorff, K. (2013). *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: SAGE Publications.

Kristiansen, J., Persson, R., Lund, S. P., Shibuya, H., and Nielsen, P. M. (2013). Effects of classroom acoustics and self-reported noise exposure on teachers' well-being. *Environ. Behav.* 45, 283–300. doi: 10.1177/0013916511429700

Krueger, R., and Casey, M. A. (2009). *Focus Groups. A Practical Guide for Applied Research*. Thousand Oaks, CA: Sage Publications.

Lane, H., and Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *J. Speech Hear. Res.* 14, 677–709. doi: 10.1044/jshr.1404.677

Lu, Y., and Cooke, M. (2008). Speech production modifications produced by competing talkers, babble, and stationary noise. *J. Acoust. Soc. Am.* 124, 3261–3275. doi: 10.1121/1.2990705

Malterud, K. (2009). *Qualitative Methods in Medical Research [in Swedish]*. Lund: Studentlitteratur.

McAllister, A. (2019). "Voice ergonomics for children. Children's voice ergonomics," in *Voice Ergonomics. Occupational and Professional Voice Care*, eds E. Sala and L. Rantala, (Newcastle upon Tyne: Cambridge Scholars Publishing), 130–153.

McAllister, A., Granqvist, S., Sjölander, P., and Sundberg, J. (2009). Child voice and noise: a pilot study of the effect of a day at the day-care on ten children's voice quality according to perceptual evaluation. *J. Voice* 23, 587–593. doi: 10.1016/j.jvoice.2007.10.017

McAllister, A., and Simberg, S. (2019). "Children's voice ergonomics," in *Voice Ergonomics. Occupational and Professional Voice Care*, eds E. Sala and L. Rantala, (Newcastle upon Tyne: Cambridge Scholars Publishing), 221–224.

McAllister, R. (1990). *Perceptual Foreign Accent: l2 Users' Comprehension Ability. New Sounds 90*. Amsterdam: University of Amsterdam.

McKellin, W. H., Shahin, K., Hodgson, M., Jamieson, J., and Pichora-Fuller, K. (2007). Pragmatics of conversation and communication in noisy settings. *J. Pragmat.* 39, 2159–2184. doi: 10.1016/j.pragma.2006.11.012

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1097/AUD.0b013e3181d3d514

Nyberg, J., and Havstam, C. (2016). Speech in 10-year-olds born with cleft lip and palate: what do peers say? *Cleft Palate Craniofac. J.* 53, 516–526. doi: 10.1597/15-140

Patton, M. (2002). *Qualitative Research & Evaluation Methods*, 3rd ed. Thousand Oaks, CA: SAGE Publication.

Peelle, J. E. (2018). Listening effort: how the cognitive consequences of acoustic challenge are reflected in brain and behavior. *Ear Hear.* 39, 204–214. doi: 10.1097/AUD.0000000000000494

Persson Waye, K., van Kamp, I., and Dellve, L. (2013). Validation of a questionnaire measuring preschool children's reactions to and coping with noise in a repeated measurement design. *BMJ Open* 3:e002408. doi: 10.1136/bmjopen-2012-002408

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., et al. (2016). Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear.* 37, 5S–27S. doi: 10.1097/aud.0000000000000312

Rantala, L. M., Hakala, S., Holmqvist, S., and Sala, E. (2015). Classroom noise and teachers' voice production. *J. Speech Lang. Hear. Res.* 58, 1397–1406. doi: 10.1044/2015_JSLHR-S-14-0248

Rantala, L. M., Hakala, S. J., Holmqvist, S., and Sala, E. (2012). Connections between voice ergonomic risk factors and voice symptoms, voice handicap, and respiratory tract diseases. *J. Voice* 26:819.e13-20. doi: 10.1016/j.jvoice.2012.06.001

Rogers, C. L., Lister, J. L., Febo, D. M., Besing, J. M., and Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Appl. Psycholinguist.* 27, 465–485. doi: 10.3766/jaaa.15125

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., and Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *J. Speech Lang. Hear. Res.* 47, 281–293.

Sala, E., Airo, E., Olkinuora, P., Simberg, S., Ström, U., Laine, A., et al. (2002). Vocal loading among day care center teachers. *Logop. Phoniatr. Voco.* 27, 21–28. doi: 10.1080/140154302760146943

Sala, E., Laine, A., Simberg, S., Pentti, J., and Suonpaa, J. (2001). The prevalence of voice disorders among day care center teachers compared with nurses: a questionnaire and clinical study. *J. Voice* 15, 413–423. doi: 10.1016/s0892-1997(01)00042-x

Sala, E., and Rantala, L. (2016). Acoustics and activity noise in school classrooms in Finland. *Appl. Acoust.* 114, 252–259. doi: 10.1016/j.apacoust.2016.08.009

Sala, E., and Rantala, L. (2019). "Noise and room acoustics," in *Voice Ergonomics. Occupational and Professional Voice Care*, eds E. Sala and L. Rantala, (Newcastle upon Tyne: Cambridge Scholars Publishing), 172–178.

Sederholm, E. (1995). Prevalence of hoarseness in ten-year-old children. *Scand. J. Logop. Phoniatr.* 20, 165–173. doi: 10.3109/14015439509098744

Sederholm, E., McAllister, A., Dalkvist, J., and Sundberg, J. (1995). Aetiologic factors associated with hoarseness in ten-year-old children. *Folia Phoniatr. Log.* 5, 262S–278S.

SFS 5907 (2004). *Acoustic Classification of Spaces of Buildings*. Helsinki: Finnish Standards Association.

Shield, B., and Dockrell, J. E. (2004). External and internal noise surveys of London primary schools. *J. Acoust. Soc. Am.* 115, 730–738. doi: 10.1121/1.1635837

Shield, B. M., and Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Södersten, M., Grankvist, S., Hammarberg, B., and Szabo, A. (2002). Vocal behavior and vocal loading factors for preschool teachers at work studied with binaural DAT recordings. *J. Voice* 16, 356–371. doi: 10.1016/s0892-1997(02)00107-8

Södersten, M., Ternström, S., and Bohman, M. (2005). Loud speech in realistic environmental noise: phonetogram data, perceptual voice quality, subjective ratings, and gender differences in healthy speakers. *J. Voice* 19, 29–46. doi: 10.1016/j.jvoice.2004.05.002

Szabo Portela, A., Granqvist, S., Ternström, S., and Södersten, M. (2018). Vocal behavior in environmental noise: comparisons between work and leisure conditions in women with work-related voice disorders and matched controls. *J. Voice* 32:126.e23-e38. doi: 10.1016/j.jvoice.2017.04.010

Tabri, D., Abou Chacra, K. M., and Pring, T. (2011). Speech perception in noise by monolingual, bilingual and trilingual listeners. *Int. J. Lang. Commun. Disord.* 46, 411–422. doi: 10.3109/13682822.2010.519372

Ternström, S., Bohman, M., and Södersten, M. (2006). Loud speech over noise: some spectral attributes, with gender differences. *J. Acoust. Soc. Am.* 119, 1648–1665. doi: 10.1121/1.2161435

Vilkman, E. (2004). Occupational safety and health aspects of voice and speech professions. *Folia Phoniatr. Logop.* 56, 220–253. doi: 10.1159/000078344

Whitling, S., Lyberg-Åhlander, V., and Rydell, R. (2017). Recovery from heavy vocal loading in women with different degrees of functional voice problems. *J. Voice* 31, .e1–.e645. doi: 10.1016/j.jvoice.2016.12.012

Williams, C., Bower, E. J., and Newton, J. T. (2004). Research in primary dental care part 6: data analysis. *Br. Dent. J.* 197, 67–73. doi: 10.1038/sj.bdj.4811467

Masked Speech Recognition in School-Age Children

Lori J. Leibold[1]* and Emily Buss[2]

[1] *Human Auditory Development Laboratory, Department of Research, Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE, United States, [2] Psychoacoustics Laboratories, Department of Otolaryngology/Head and Neck Surgery, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States*

Children who are typically developing often struggle to hear and understand speech in the presence of competing background sounds, particularly when the background sounds are also speech. For example, in many cases, young school-age children require an additional 5- to 10-dB signal-to-noise ratio relative to adults to achieve the same word or sentence recognition performance in the presence of two streams of competing speech. Moreover, adult-like performance is not observed until adolescence. Despite ample converging evidence that children are more susceptible to auditory masking than adults, the field lacks a comprehensive model that accounts for the development of masked speech recognition. This review provides a synthesis of the literature on the typical development of masked speech recognition. Age-related changes in the ability to recognize phonemes, words, or sentences in the presence of competing background sounds will be discussed by considering (1) how masking sounds influence the sensory encoding of target speech; (2) differences in the time course of development for speech-in-noise versus speech-in-speech recognition; and (3) the central auditory and cognitive processes required to separate and attend to target speech when multiple people are speaking at the same time.

Keywords: development, children, hearing, speech perception, masking

INTRODUCTION

Children must learn how to communicate in noisy environments such as classrooms (e.g., Knecht et al., 2002). Thus, it is not surprising that extensive research conducted over the past 30 years has focused on understanding children's masked speech recognition abilities (e.g., Elliott, 1979; Hall et al., 2002; Brown et al., 2010; McCreery et al., 2017; Dillon et al., 2018). Several consistent trends have emerged from this research. First, the detrimental effects of auditory masking on speech recognition are larger for children than for adults (reviewed by Erickson and Newman, 2017). Second, the ability to recognize speech in the presence of competing sounds develops throughout the school-age years and does not mature until adolescence (e.g., Cameron et al., 2009; Brown et al., 2010; Corbin et al., 2016). Finally, children's increased susceptibility to auditory masking relative to adults in the context of speech recognition is more pronounced and prolonged when the masker is also speech than when the masker is steady-state noise (e.g., Hall et al., 2002; Corbin et al., 2016). These results have collectively had significant impact on public health policy, leading to the establishment of classroom standards for noise levels (ANSI, 2010) as well as recommendations that speech-in-noise testing be included in the pediatric audiology test battery.

While children's considerable masked speech recognition difficulties have been well documented, a comprehensive model of the factors responsible for developmental effects has not been established. This review aims to characterize child/adult differences in the ubiquitous problem of recognizing speech in the presence of competing background sounds, with a specific goal of summarizing the literature pertaining to factors thought to be responsible for age-related changes in performance. The review begins with an overview of children's speech recognition abilities in steady-state noise. Historically, the development of speech-in-noise recognition has been a major focus for researchers in the field. This focus partly reflects an early emphasis on understanding bottom-up contributions to development, based on the premise that speech recognition in steady-state noise requires an accurate sensory representation of target speech. Findings from studies investigating the influence of top-down contributions of language knowledge and cognitive processing on children's recognition of speech that has been degraded by noise are then discussed. Building on this foundational work, the latter half of the review concentrates on age effects on the ability to recognize speech when several people are talking in the background. The research summarized in this section provides compelling evidence that central auditory and cognitive processing play a critical role in the development of speech-in-speech recognition. Finally, areas for future research are briefly highlighted.

SPEECH-IN-NOISE RECOGNITION

Children are poorer than adults are at recognizing phonemes, words, or sentences in a background of steady-state noise (e.g., Elliott, 1979; Nittrouer and Boothroyd, 1990; McCreery and Stelmachowicz, 2011; Dillon et al., 2018). For example, McCreery and Stelmachowicz (2011) evaluated syllable recognition in a speech-shaped noise masker. Participants were a large sample of 5- to 12-year-old children (n = 116) and young adults with normal hearing. Stimulus bandwidth was manipulated *via* filtering, and testing was completed at multiple signal-to-noise ratios (SNRs). Children consistently required more favorable SNRs than adults to achieve comparable performance. Similar child/adult differences have been reported using word and sentence stimuli (e.g., Buss et al., 2017), and findings from related studies indicate that children require greater spectral detail relative to adults in order to recognize filtered speech (Eisenberg et al., 2000; Mlot et al., 2010).

A closer examination of the literature reveals that speech-in-noise recognition improves gradually over the first decade of life; adult-like performance is not usually observed until 9–10 years when stimuli are presented diotically (e.g., Eisenberg et al., 2000; Corbin et al., 2016; Buss et al., 2017; but see Jacobi et al., 2017). Corbin et al. (2016) characterized the developmental trajectory for masked word recognition, including testing in the presence of speech-shaped noise. Participants were 5- to 16-year-old children and young adults with normal hearing. As a group, children needed an additional 2.3-dB SNR relative to adults to attain the same correct-response criterion. However, substantial age-related improvements in performance were observed across the age range of children tested. SRTs improved linearly with age until about 10 years of age, but SRTs for older children were indistinguishable from those observed for adults.

FACTORS RESPONSIBLE FOR DEVELOPMENTAL EFFECTS

Peripheral Encoding

Speech recognition relies on an accurate representation of incoming speech transmitted to the brain *via* the outer ear, middle ear, cochlea, and auditory nerve. Competing noise compromises this representation when the neural excitation produced by target speech and masking noise overlap on the basilar membrane (e.g., Miller, 1947). The term *energetic masking* is often used in the literature to describe the perceptual consequences of degraded peripheral encoding (reviewed by Brungart, 2005). These consequences include reduced audibility, which in turn limits access to acoustic speech features and exerts a negative influence on overall speech intelligibility (e.g., Fletcher and Galt, 1950; Miller and Nicely, 1955).

Extensive research conducted over the past 40 years has focused on understanding the limits of peripheral encoding in children (reviewed by Buss et al., 2012). Results of this work provide converging evidence that school-age children's speech-in-noise difficulties are not due to immaturity in the sensory representation of speech. Neural transmission through the brainstem auditory pathways appears to be somewhat sluggish during early infancy, but this immaturity appears to resolve by about 6 months of age (e.g., Gorga et al., 1989; Werner et al., 1994). While behavioral data indicate that auditory capabilities related to frequency, intensity, and temporal processing improve during infancy and the early school-age years (Buss et al., 2012), peripheral encoding of the basic properties of sound appears to reach adult-like precision by 6 months of age (reviewed by Eggermont and Moore, 2012). For example, findings from histological, anatomical, and physiological studies indicate mature cochlear function by at least term birth (e.g., Lavigne-Rebillard and Pujol, 1987; Abdala, 2001).

Listening Strategy

Children's pronounced speech-in-noise difficulties may be due in part to immature allocation of attention (e.g., Nittrouer et al., 1993; Choi et al., 2008; Youngdahl et al., 2018). Young children show a tendency to listen across a broad range of frequencies, rather than the mature strategy of focusing attention only on regions associated with relevant target speech (e.g., Polka et al., 2008; Youngdahl et al., 2018). In a recent study, Youngdahl et al. (2018) examined whether 5-year-olds, 7-years-olds, or young adults were susceptible to remote-frequency masking in the context of masked sentence recognition. Target sentences were presented in quiet or in noise. Importantly, target speech and masking noise were filtered to ensure no overlap in frequency. Adults and 7-year-olds performed similarly in quiet and masked conditions. In contrast, 5-year-olds performed more poorly in noise than in quiet. These remote-frequency masking effects

are in agreement with prior speech detection data reported for infants (Polka et al., 2008), as well as tone-in-noise detection data reported for infants and 4- to 6-year-old children (Bargones and Werner, 1994; Leibold and Neff, 2011).

Children may initially adopt a different listening strategy than adults in order to learn the important speech cues in their native language. This idea is supported by findings from a series of studies conducted by Nittrouer and colleagues investigating the perceptual attention that children and adults assign to the different acoustic components of phonemes (reviewed by Nittrouer, 2002). Whereas preschoolers attend more heavily to speech cues that are dynamic (e.g., formant transitions), adults and children as young as 7 years of age are more influenced by speech cues that are relatively stable across time (e.g., frication noise). This shift in perceptual attention, called the *perceptual weighting shift* (Nittrouer et al., 1993), is consistent with the idea that extensive listening experience is required before mature selective attention abilities emerge.

Linguistic Knowledge

It has been suggested that children's pronounced speech-in-noise difficulties reflect their inexperience with language. However, studies that have tested for associations between masked speech recognition and language abilities reveal mixed findings as some studies do not support this association (e.g., Garlock et al., 2001; McCreery and Stelmachowicz, 2011; Nittrouer et al., 2013; Klein et al., 2017; McCreery et al., 2017). Several studies have reported a correlation between children's speech-in-noise recognition scores and the size of their vocabulary (e.g., McCreery and Stelmachowicz, 2011; Vance and Martindale, 2012), but this relationship has not been observed in other studies (e.g., Eisenberg et al., 2000; Nittrouer et al., 2013).

Discrepancies observed between studies investigating the association between vocabulary knowledge and masked speech recognition may be due to differences in the stimuli used to evaluate this association. Investigators routinely select target speech that falls within the lexicon of the youngest children tested for a given experiment (e.g., Eisenberg et al., 2000; Nittrouer et al., 2013; McCreery et al., 2017). Findings from studies that included later acquired words provide important insight into the association between vocabulary size and masked speech recognition (e.g., Garlock et al., 2001; Klein et al., 2017). Klein et al. (2017) assessed masked word and non-word recognition in a group of 5- to 12-year-old children with hearing loss and an equal number of age-matched children with normal hearing. Vocabulary size for both groups of children was associated with speech-in-noise recognition performance when target stimuli were non-words or later acquired words. In contrast, no association between these two factors was observed when target stimuli were earlier acquired words.

Working Memory

There has been considerable recent interest in understanding how the cognitive process of working memory influences children's speech-in-noise recognition abilities. Working memory refers to the temporary storage and processing of incoming sensory information in a memory buffer, allowing for comparisons with stored representations (Baddeley, 2000; Cowan, 2004). Along with speech-in-noise recognition and language skills, working memory abilities improve with age during childhood (e.g., Camos and Barrouillet, 2015).

Data reported in the literature, albeit from a small number of studies, suggest that working memory may play an important role in the development of speech-in-noise recognition. Differences in working memory between children appear to be partly responsible for individual differences in performance on masked speech recognition tests, even when age effects are taken into account (e.g., Magimairaj and Montgomery, 2012; McCreery et al., 2017; but see Magimairaj et al., 2018). McCreery et al. (2017) measured speech-in-noise recognition and performance on four subtests of the Automated Working Memory Assessment (Alloway et al., 2008) in a group of 48 school-age children (5–12 years). Speech recognition was assessed in a speech-shaped noise masker for three types of targets: monosyllabic words, low-predictability sentences, and high-predictability sentences. Children with higher working memory scores showed better speech-in-noise recognition performance for all three types of target stimuli, after controlling for age and vocabulary size.

DEVELOPMENT OF SPEECH-IN-SPEECH RECOGNITION

Age effects for speech recognition in a masker composed of a small number of speech streams are pronounced relative to those observed in broadband noise with the same long-term average spectrum (e.g., Hall et al., 2002; Wightman and Kistler, 2005; Corbin et al., 2016). For example, Hall et al. (2002) used a forced-choice, picture-pointing task to assess recognition of spondaic words in the presence of speech-shaped noise or two-talker speech. Listeners were 5- to 10-year-old children and 19- to 48-year-old adults. On average, children required an additional 3 dB to perform as well as adults in the noise masker. In contrast, the magnitude of the child/adult difference was 8-dB SNR in the two-talker masker. Larger developmental effects for speech-in-speech relative to speech-in-noise recognition have also been reported using phonemes (Leibold and Buss, 2013), monosyllabic words (e.g., Corbin et al., 2016), and sentences (e.g., Wightman and Kistler, 2005).

Not only are child/adult differences more pronounced for speech-in-speech than for speech-in-noise recognition, mature performance is not reached until the teenage years (e.g., Wightman and Kistler, 2005; Brown et al., 2010; Leibold and Buss, 2013; Corbin et al., 2016). Corbin et al. (2016) assessed children's (5–16 years) and adults' word recognition in a two-talker speech masker as well as in a speech-shaped noise masker. Mature SRTs were observed by 10 years of age in the noise masker, but adult-like SRTs for the same children were not observed in the speech masker until after 13 years of age. These observations are consistent with the idea that the factors responsible for developmental effects in speech-in-speech recognition may differ from those responsible for speech-in-noise recognition, and may emerge at different stages of development.

FACTORS RESPONSIBLE FOR DEVELOPMENTAL EFFECTS

Perceptual Isolation of Target and Masker Speech

The ability to recognize speech produced by one talker when multiple people are talking at the same time relies on central auditory processing. This processing facilitates the grouping of sounds into separate auditory objects and is responsible for the selective allocation of attention (e.g., Bregman, 1990; Bronkhorst, 2000; Best et al., 2007). Collectively, this processing falls within the general framework of *auditory scene analysis* (Bregman, 1990). The perceptual consequences of a failure of grouping and/or selection are sometimes referred to as *perceptual* or *informational masking* (e.g., Carhart et al., 1969; Brungart, 2001). Regardless of terminology, immature grouping and/or selective attention abilities appear to limit the extent to which children perceptually isolate target and masker speech (reviewed by Leibold, 2017).

Auditory grouping refers to the segregation of simultaneous sounds as well as the linkage of sounds over time (e.g., Bregman, 1990; Bronkhorst, 2015). Acoustic differences between target and masker speech influence auditory grouping in adults (e.g., Bregman, 1990; Bronkhorst, 2000; Brungart, 2001; Darwin et al., 2003). For example, speech produced by different talkers tends to vary with respect to multiple acoustic vocal characteristics, including fundamental frequency (F0) and the distribution of formant frequencies (e.g., Fitch and Giedd, 1999). Adults capitalize on these acoustic differences in the context of speech-in-speech recognition, particularly when target and masker speech are produced by talkers that differ in sex (e.g., Festen and Plomp, 1990; Brungart, 2001). Other target/ masker acoustic differences that promote auditory grouping and have a positive impact on adults' speech-in-speech recognition performance include temporal onsets (e.g., Hukin and Darwin, 1995) and binaural cues associated with real or perceived spatial location (e.g., Freyman et al., 2001).

Children appear to take advantage of many of the same acoustic differences between target and masker speech that improve adults' speech-in-speech recognition performance (e.g., Litovsky, 2005; Cameron et al., 2009, 2011; Yuen and Yuan, 2014; Calandruccio et al., 2016). For example, Litovsky (2005) examined the effect of spatially separating target and masker speech on masked speech recognition performance. Listeners were 4- to 7-year-old children and adults. A forced-choice task with a picture-pointing response was used to estimate SRTs for words embedded in speech-shaped noise, competing sentences produced by one talker, or competing sentences produced by two talkers. Target stimuli were always delivered *via* a loudspeaker positioned directly in front of the listener at 0° azimuth. Maskers were presented from the same location as the target words (co-located) or from a loudspeaker positioned 90° to the side of the listener (separated). Spatial release from masking (SRM) was computed as the difference between the SRTs estimated in the co-located and spatially separated conditions. Children required a more advantageous SNR to achieve the same criterion level of performance as adults in all three masker conditions, but the magnitude of SRM was similar across age. Subsequent studies have confirmed that children benefit from target/masker differences in spatial location in the context of speech-in-speech recognition (e.g., Johnstone and Litovsky, 2006; Cameron et al., 2009; Murphy et al., 2011; Yuen and Yuan, 2014; Corbin et al., 2017). Note, however, that findings from more recent studies indicate that young children experience reduced SRM relative to older children and adults when the target stimuli and/or listening conditions are more challenging (e.g., Cameron et al., 2009; Brown et al., 2010; Yuen and Yuan, 2014; Corbin et al., 2016). For example, Brown et al. (2010) examined sentence recognition in a two-talker masker using the North American Listening in Spatialized Noise-Sentences test (NA LiSN-S). Listeners were a large sample of 12- to 19-year-old children ($n = 67$) and young adults ($n = 53$) with normal hearing. Testing included conditions in which the target and masker were perceived to have originated from the same location in space and conditions in which the target and masker were perceived to be spatially separated. The ability to benefit from perceived spatial separation remained immature until 14 years of age.

Prior studies investigating the extent to which children benefit from acoustic differences between target and masker speech have generally used stimuli that differ across multiple acoustic features (e.g., Litovsky, 2005; Calandruccio et al., 2016; Leibold et al., 2018). For example, Leibold et al. (2018) evaluated whether children and adults benefit from a mismatch in target/ masker sex when asked to recognize disyllabic words in a two-talker masker. SRTs for all listeners were higher (i.e., worse) when the target and masker speech were sex matched (e.g., male target speech presented in a male two-talker masker) relative to when target and masker speech were sex mismatched (e.g., male target speech presented in a female two-talker masker). Speech produced by males and females generally differs across multiple acoustic features, including F0, dispersion of formant frequencies, and phonation type (e.g., Fitch and Giedd, 1999). In a later study, Flaherty et al. (2019) observed a striking age effect in the ability to benefit from target/masker differences only in F0, holding other acoustic target/masker differences constant. Whereas adults and older children (>13 years) showed a robust benefit associated with target/ masker differences in mean F0, younger children (<7 years) did not. Flaherty et al. (2019) suggested that children might require additional acoustic cues (e.g., distribution of formant frequencies) in order to perceptually isolate target and masker speech. Additional evidence supporting this interpretation is provided by normative data for the LiSN-S clinical test (e.g., Cameron et al., 2009, 2011; Brown et al., 2010). That test battery includes conditions in which the target and masker speech are produced by the same female talker, as well as conditions in which the target and masker speech are produced by different female talkers. While children of all ages tend to show better performance when different talkers produced target and masker speech, adult-like benefit is not observed until 14 years of age.

In addition to auditory grouping, speech-in-speech recognition relies on the ability to selectively attend to the auditory object associated with target speech while disregarding other objects (e.g., Bronkhorst, 2000; Best et al., 2007). Results from several behavioral experiments indicate that children listen less selectively than adults (e.g., Doyle, 1973; Wightman and Kistler, 2005; Leibold and Buss, 2013). For example, Wightman and Kistler (2005) used a dichotic listening paradigm to investigate the influence of selective attention on children's increased susceptibility to speech-in-speech masking. Listeners were 4- to 16-year-olds and adults. In all conditions, a single target sentence and a single distractor sentence were simultaneously presented to the listener's right ear. In some conditions, an additional distractor sentence was presented to the listener's left ear. The task was to repeat back the target sentence while ignoring the distractor sentence(s). Children performed more poorly than adults in all conditions, with developmental improvements observed until about 13 years of age. While the addition of the contralateral distractor sentence negatively impacted performance for listeners of all ages, an analysis of listener error patterns revealed age effects in the ability to disregard speech presented to the contralateral ear. Most errors made by the youngest children tested (4–6 years) were intrusions from the distractor speech presented to the opposite ear as the target sentence. In contrast, errors made by older children and adults were generally intrusions from the distractor speech presented to the same ear as the target sentence.

Despite compelling evidence that selective auditory attention contributes to child/adult differences in masked speech recognition, this area of research remains under-studied. One complicating factor is that the relationship between selective attention and auditory grouping is bidirectional; the formation of auditory objects is influenced by selective attention and vice versa (e.g., Shamma et al., 2011). A related challenge is that we lack behavioral paradigms that can isolate effects of immature selective attention from failures in auditory object formation. Functionally, both processes impact speech-in-speech recognition. Results from electrophysiological studies have provided insight regarding the time course of development of these factors (e.g., Coch et al., 2005; Karns et al., 2015). For example, Karns et al. (2015) examined event-related potentials (ERPs) in the context of a dichotic listening experiment. Listeners were 3- to 5-year-olds, 10-year-olds, 13-year-olds, 16-year-olds, and young adults. Listeners were asked to attend to speech presented to a loudspeaker while ignoring speech presented to another loudspeaker at the same time, or they were asked to attend to speech presented by a male or female talker while ignoring speech produced by a talker that differed in sex. Age-related changes for both tasks were observed in both the latency and morphology of ERPs, with adult-like responses observed only for the oldest two groups of children tested (13 and 16 years).

Glimpsing

Adults take advantage of brief "glimpses" of target speech available during minima in the envelope of modulated noise (i.e., epochs in which SNR is relatively high), showing better speech recognition performance in modulated or interrupted noise than in nominally steady noise (e.g., Miller and Licklider, 1950; Howard-Jones and Rosen, 1993; Cooke, 2006). Speech maskers composed of a small number of speech streams likewise fluctuate over time. Thus, it has been suggested that children's increased susceptibility to speech-in-speech masking relative to adults may reflect immaturity in the ability to capitalize on glimpsing opportunities (e.g., Buss et al., 2017; Sobon et al., 2019).

Initial studies investigating children's speech recognition in temporally modulated noise yielded mixed results regarding child/adult differences in glimpsing (e.g., Stuart, 2008; Hall et al., 2014). More recent studies, however, indicate that school-age children derive less benefit from temporal glimpses in a one- or two-talker speech masker relative to adults (e.g., Buss et al., 2017; Sobon et al., 2019). Buss et al. (2017) evaluated word recognition in a one-talker or a two-talker masker. Listeners were 4- to 16-year-old children and young adults. SRTs were estimated adaptively in each masker, both with and without the addition of a speech-shaped noise. When present, the speech-shaped noise was 10 dB less intense in level than the corresponding speech masker. The rationale for assessing performance with the added noise was to examine the effect of masking the low-level speech cues that would otherwise be available during the envelope minima of the speech masker. The effect of adding noise was larger for older children and adults than for younger children. A follow-up experiment utilized a technique whereby time segments of the combined target and masker speech associated with poor SNRs were removed *via* digital signal processing. The goal of this technique is to approximate ideal segregation of target and masker speech by discarding the time/frequency segments of the stimulus dominated by the masker (e.g., Wang, 2005). Digital segregation reduced the child/adult difference. Nonetheless, young children continued to perform more poorly than older children and adults. Overall, the pattern of results observed across the two experiments reported by Buss et al. (2017) suggests young children are less adept than older children and adults at recognizing speech based on brief glimpses.

Results from Sobon et al. (2019) provide additional evidence that glimpsing abilities limit speech-in-speech recognition during childhood. Speech-in-noise and speech-in-speech recognition were evaluated in 8- to 10-year-olds and young adults. Data were collected using an adaptive sentence recognition task and subsequently fitted with psychometric functions. Similar psychometric slopes were observed for children and adults in the speech-shaped noise masker, but slopes were steeper for children than for adults in the two-talker masker. This result was interpreted as indicating that children were not able to benefit from transient improvements in SNR in the two-talker masker to the same extent as adults. This interpretation received additional support from an analysis using the extended speech intelligibility index (Rhebergen and Versfeld, 2005), to estimate the audibility of speech cues required for recognition. Children required more audibility overall than adults, but this difference was larger for the two-talker masker than the speech-shaped noise masker. These results are consistent with the idea that children's immature

speech-in-speech recognition is at least partly due to reduced glimpsing abilities. Immature segregation, selective attention, or a combination of these two effects may contribute to young children's reduced ability to recognize speech based on sparse cues.

SUMMARY AND FUTURE DIRECTIONS

Data summarized in this review provide compelling evidence that the ability to recognize masked speech follows a prolonged time course of development. Children have more difficulty recognizing speech in the presence of background sounds relative to adults, with age effects reported for a wide range of stimuli and listening conditions. Research on children's speech recognition in steady-state noise indicates that child/adult differences persist until about 9–10 years of age (e.g., McCreery and Stelmachowicz, 2011; Corbin et al., 2016). In contrast, child/adult differences appear to be larger and extend into adolescence when the masker is also speech (e.g., Hall et al., 2002; Brown et al., 2010; Corbin et al., 2016; Buss et al., 2017; but see Dillon et al., 2018). Masker-dependent differences in the time course of development highlight the importance of incorporating both listener and stimulus factors into models of masked speech recognition.

A focus for this review was to consolidate what is known about the factors responsible for developmental effects in masked speech recognition. Recognizing speech in the presence of background sounds depends upon on multiple stages of auditory, cognitive, and linguistic processing. It is important to highlight that immature processing within any stage of processing is likely to influence the extent to which children hear and understand speech in their everyday lives. It is well established that degradations in peripheral encoding negatively influence speech recognition (e.g., Miller and Nicely, 1955), but is perhaps less obvious to researchers outside the field that an immature ability to perceptually isolate target and masker speech can result in the same functional consequences. Efforts are needed to establish models that account for maturational effects, taking into account the specific contributions of the multiple factors and processes required to recognize masked speech.

There are a number of key challenges to address in future research. Efforts are underway to understand the many factors that affect children's masked speech recognition abilities, including age, audibility, masker complexity, working memory, and language skills (e.g., Lang et al., 2017). Another long-standing issue is the general dearth of behavioral paradigms and psychometric methods required to understand and quantify contributions of auditory grouping, selective attention, and/or more general cue requirements to children's speech-in-speech recognition abilities. As recent data by Sobon et al. (2019) indicate, factors such as the slope of the psychometric function and the SNR at which a criterion threshold is reached can provide more accurate and detailed estimates of child/adult differences than the conventional approach of considering threshold data alone. Finally, the studies discussed in this review involved children with typical development. Future research is needed to determine how listener factors such as peripheral hearing loss, neurological abnormalities, limited language experience, and cognitive impairment impact children's masked speech recognition abilities (e.g., Hillock-Dunn et al., 2015; Chermak et al., 2017).

AUTHOR CONTRIBUTIONS

LL and EB both contributed to the writing of the review.

FUNDING

REFERENCES

Abdala, C. (2001). Maturation of the human cochlear amplifier: distortion product otoacoustic emission suppression tuning curves recorded at low and high primary tone levels. *J. Acoust. Soc. Am.* 110, 1465–1476. doi: 10.1121/1.1388018

Alloway, T., Gathercole, S. E., Kirkwood, H., and Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Educ. Psychol.* 28, 725–734. doi: 10.1080/01443410802243828

ANSI (2010). *ANSI/ASA S12.60–2010 acoustical performance criteria, design requirements, and guidelines for schools, part 1: Permanent schools*. New York, NY: American National Standards Institute.

Baddeley, A. D. (2000). "Short-term and working memory" in *The Oxford handbook of memory*. eds. E. Tulving, and C. FIM (New York, NY: Oxford University Press), 77–92.

Bargones, J. Y., and Werner, L. A. W. (1994). Adults listen selectively; infants do not. *Psychol. Sci.* 5, 170–174.

Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (2007). Binaural interference and auditory grouping. *J. Acoust. Soc. Am.* 121, 1070–1076. doi: 10.1121/1.2407738

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.

Bronkhorst, A. W. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acta Acust.* 86, 117–128.

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: early processing and selection of multi-talker speech. *Atten. Percept. Psychophys.* 77, 1465–1487. doi: 10.3758/s13414-015-0882-9

Brown, D. K., Cameron, S., Martin, J. S., Watson, C., and Dillon, H. (2010). The North American Listening in Spatialized Noise-Sentences test (NA LiSN-S): normative data and test-retest reliability studies for adolescents and young adults. *J. Am. Acad. Audiol.* 21, 629–641. doi: 10.3766/jaaa.21.10.3

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Brungart, D. S. (2005). "Informational and energetic masking effects in multitalker speech perception" in *Speech separation by humans and machines*. ed. P. Divenyi (Boston, MA: Springer), 261–267.

Buss, E., Hall, J. W. III., and Grose, J. H. (2012). "Development of auditory coding as reflected in psychophysical performance" in *Human auditory development*. eds. L. A. Werner, R. R. Fey, and A. N. Popper (New York, NY: Springer), 107–136.

Buss, E., Leibold, L. J., Porter, H. L., and Grose, J. H. (2017). Speech recognition in one- and two-talker maskers in school-age children and adults: development of perceptual masking and glimpsing. *J. Acoust. Soc. Am.* 141, 2650–2660. doi: 10.1121/1.4979936

Calandruccio, L., Leibold, L. J., and Buss, E. (2016). Linguistic masking release in school-age children and adults. *Am. J. Audiol.* 25, 34–40. doi: 10.1044/2015_AJA-15-0053

Cameron, S., Brown, D., Keith, R., Martin, J., Watson, C., and Dillon, H. (2009). Development of the North American Listening in Spatialized Noise-Sentences test (NA LiSN-S): sentence equivalence, normative data, and test-retest reliability studies. *J. Am. Acad. Audiol.* 20, 128–146. doi: 10.3766/jaaa.20.2.6

Cameron, S., Glyde, H., and Dillon, H. (2011). Listening in Spatialized Noise-Sentences Test (LiSN-S): normative and retest reliability data for adolescents and adults up to 60 years of age. *J. Am. Acad. Audiol.* 22, 697–709. doi: 10.3766/jaaa.22.10.7

Camos, V., and Barrouillet, P. (2015). *Working memory: Loss and reconstruction*. New York, NY: Psychology Press.

Carhart, R., Tillman, T. W., and Greetis, E. S. (1969). Perceptual masking in multiple sound backgrounds. *J. Acoust. Soc. Am.* 45, 694–703.

Chermak, G. D., Musiek, F. E., and Weihing, J. (2017). Beyond controversies: the science behind central auditory processing disorder. *Hear. Rev.* 24, 20–24.

Choi, S., Lotto, A., Lewis, D., Hoover, B., and Stelmachowicz, P. (2008). Attentional modulation of word recognition by children in a dual-task paradigm. *J. Speech Lang. Hear. Res.* 51, 1042–1054. doi: 10.1044/1092-4388(2008/076)

Coch, D., Sanders, L. D., and Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *J. Cogn. Neurosci.* 17, 605–622. doi: 10.1162/0898929053467631

Cooke, M. (2006). A glimpsing model of speech perception in noise. *J. Acoust. Soc. Am.* 119, 1562–1573. doi: 10.1121/1.2166600

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (2016). Development of open-set word recognition in children: speech-shaped noise and two-talker speech maskers. *Ear Hear.* 37, 55–63. doi: 10.1097/AUD.0000000000000201

Corbin, N. E., Buss, E., and Leibold, L. J. (2017). Spatial release from masking in children: effects of simulated unilateral hearing loss. *Ear Hear.* 38, 223–235. doi: 10.1097/AUD.0000000000000376

Cowan, N. (2004). "On the psychophysics of memory" in *Psychophysics beyond sensation: Laws and invariants of human cognition. Scientific psychology series*. eds. C. Kaernbach, E. Schröger, and H. Müller (Mahwah, NJ: Erlbaum), 313–319.

Darwin, C. J., Brungart, D. S., and Simpson, B. D. (2003). Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. *J. Acoust. Soc. Am.* 114, 2913–2922. doi: 10.1121/1.1616924

Dillon, H., Mee, C., Moreno, J. C., and Seymour, J. (2018). Hearing tests are just child's play: the sound scouts game for children entering school. *Int. J. Audiol.* 57, 529–537. doi: 10.1080/14992027.2018.1463464

Doyle, A. B. (1973). Listening to distraction: A developmental study of selective attention. *J. Exp. Child Psychol.* 15, 100–115.

Eggermont, J. J., and Moore, J. K. (2012). "Morphological and functional development of the auditory nervous system" in *Human auditory development*. eds. L. A. Werner, R. R. Fey, and A. N. Popper (New York, NY: Springer), 61–105.

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., and Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *J. Acoust. Soc. Am.* 107, 2704–2710. doi: 10.1121/1.428656

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *J. Acoust. Soc. Am.* 66, 651–653.

Erickson, L. C., and Newman, R. S. (2017). Influences of background noise on infants and children. *Curr. Dir. Psychol. Sci.* 26, 451–457. doi: 10.1177/0963721417709087

Festen, J. M., and Plomp, R. (1990). Effects of fluctuating noise and interfering speech on the speech reception threshold for impaired and normal hearing. *J. Acoust. Soc. Am.* 88, 1725–1736.

Fitch, W. T., and Giedd, J. (1999). Morphology and development of the human vocal tract: a study using magnetic resonance imaging. *J. Acoust. Soc. Am.* 106, 1511–1522.

Flaherty, M. M., Buss, E., and Leibold, L. J. (2019). Developmental effects in children's ability to benefit from F0 differences between target and masker speech. *Ear Hear.* 40, 927–937. doi: 10.1097/AUD.0000000000000673

Fletcher, H., and Galt, R. H. (1950). The perception of speech and its relation to telephony. *J. Acoust. Soc. Am.* 22, 89–151.

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *J. Acoust. Soc. Am.* 109, 2112–2122. doi: 10.1121/1.1354984

Garlock, V. M., Walley, A. C., and Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *J. Mem. Lang.* 45, 468–492. doi: 10.1006/jmla.2000.2784

Gorga, M. P., Kaminski, J. R., Beauchaine, K. L., Jesteadt, W., and Neely, S. T. (1989). Auditory brainstem responses from children three months to three years of age: Normal patterns of response II. *J. Speech. Lang. Hear. Res.* 32, 281–288.

Hall, J. W. III, Buss, E., and Grose, J. H. (2014). Development of speech glimpsing in synchronously and asynchronously modulated noise. *J. Acoust. Soc. Am.* 135, 3594–3600.

Hall, J. W. III, Grose, J. H., Buss, E., and Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear Hear.* 23, 159–165. doi: 10.1097/00003446-200204000-00008

Hillock-Dunn, A., Taylor, C., Buss, E., and Leibold, L. J. (2015). Assessing speech perception in children with hearing loss: what conventional clinical tools may miss. *Ear Hear.* 36, e57–e60. doi: 10.1097/AUD.0000000000000110

Howard-Jones, P. A., and Rosen, S. (1993). Uncomodulated glimpsing in "checkerboard" noise. *J. Acoust. Soc. Am.* 93, 2915–2922.

Hukin, R. W., and Darwin, C. J. (1995). Comparison of the effect of onset asynchrony on auditory grouping in pitch matching and vowel identification. *Percept. Psychophys.* 57, 191–196. doi: 10.3758/BF03206505

Jacobi, I., Rashid, M. S., de Laat, J. A. P. M., and Dreschler, W. A. (2017). Age dependence of thresholds for speech in noise in normal-hearing adolescents. *Trends Hear.* 21:2331216517743641. doi: 10.1177/2331216517743641

Johnstone, P. M., and Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *J. Acoust. Soc. Am.* 120, 2177–2189. doi: 10.1121/1.2225416

Karns, C. M., Isbell, E., Giuliano, R. J., and Neville, H. J. (2015). Auditory attention in childhood and adolescence: an event-related potential study of spatial selective attention to one of two simultaneous stories. *Dev. Cogn. Neurosci.* 13, 3–67. doi: 10.1016/j.dcn.2015.03.001

Klein, K. E., Walker, E. A., Kirby, B., and McCreery, R. W. (2017). Vocabulary facilitates speech perception in children with hearing aids. *J. Speech Lang. Hear. Res.* 60, 2281–2296. doi: 10.1044/2017_JSLHR-16-0086

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., and Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms. *Am. J. Audiol.* 11, 65–71. doi: 10.1044/1059-0889(2002/009)

Lang, H., McCreery, R. W., Leibold, L. J., Buss, E., and Miller, M. K. (2017). "Effects of language and cognition on children's masked speech perception" in *Presented at the 44th annual scientific meeting of the American Auditory Society. Scottsdale, AZ, March 2–4, 2017, poster #171-SP12*. Available at: https://aas.memberclicks.net/assets/2017_posters.pdf (Accessed October 25, 2018).

Lavigne-Rebillard, M., and Pujol, R. (1987). Surface aspects of the developing human organ of Corti. *Acta Otolaryngol.* 104, 43–50.

Leibold, L. J. (2017). Speech perception in complex acoustic environments: developmental effects. *J. Speech Lang. Hear. Res.* 60, 3001–3008. doi: 10.1044/2017_JSLHR-H-17-0070

Leibold, L. J., and Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *J. Speech Lang. Hear. Res.* 56, 1144–1155. doi: 10.1044/1092-4388(2012/12-0011)

Leibold, L. J., Buss, E., and Calandruccio, L. (2018). Developmental effects in masking release for speech-in-speech perception due to a target/masker sex mismatch. *Ear Hear.* 39, 935–945. doi: 10.1097/AUD.0000000000000554

Leibold, L. J., and Neff, D. L. (2011). Masking by a remote-frequency noise band in children and adults. *Ear Hear.* 32, 663–666. doi: 10.1097/AUD.0b013e31820e5074

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *J. Acoust. Soc. Am.* 117, 3091–3099. doi: 10.1121/1.1873913

Magimairaj, B. M., and Montgomery, J. W. (2012). Children's verbal working memory: relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychol.* 140, 196–207. doi: 10.1016/j.actpsy.2012.05.004

Magimairaj, B. M., Nagaraj, N. K., and Benafield, N. J. (2018). Children's speech perception in noise: evidence for dissociation from language and working memory. *J. Speech Lang. Hear. Res.* 61, 1294–1305. doi: 10.1044/2018_JSLHR-H-17-0312

McCreery, R. W., Spratford, M., Kirby, B., and Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *Int. J. Audiol.* 56, 306–315. doi: 10.1080/14992027.2016.1266703

McCreery, R. W., and Stelmachowicz, P. G. (2011). Audibility-based predictions of speech recognition for children and adults with normal hearing. *J. Acoust. Soc. Am.* 130, 4070–4081. doi: 10.1121/1.3658476

Miller, G. A. (1947). The masking of speech. *Psychol. Bull.* 44, 105–129. doi: 10.1037/h0055960

Miller, G. A., and Licklider, J. C. R. (1950). The intelligibility of interrupted speech. *J. Acoust. Soc. Am.* 22, 167–173.

Miller, G. A., and Nicely, P. E. (1955). An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* 27, 338–352.

Mlot, S., Buss, E., and Hall, J. W. III. (2010). Spectral integration and bandwidth effects on speech recognition in school-aged children and adults. *Ear Hear.* 31, 56–62. doi: 10.1097/AUD.0b013e3181ba746b

Murphy, J., Summerfield, A. Q., O'Donoghue, G. M., and Moore, D. R. (2011). Spatial hearing of normally hearing and cochlear implanted children. *Int. J. Pediatr. Otorhinolaryngol.* 75, 489–494. doi: 10.1016/j.ijporl.2011.01.002

Nittrouer, S. (2002). Learning to perceive speech: how fricative perception changes, and how it stays the same. *J. Acoust. Soc. Am.* 112, 711–719. doi: 10.1121/1.1496082

Nittrouer, S., and Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *J. Acoust. Soc. Am.* 87, 2705–2715.

Nittrouer, S., Caldwell-Tarr, A., Tarr, E., Lowenstein, J. H., Rice, C., and Moberly, A. C. (2013). Improving speech-in-noise recognition for children with hearing loss: potential effects of language abilities, binaural summation, and head shadow. *Int. J. Audiol.* 52, 513–525. doi: 10.3109/14992027.2013.792957

Nittrouer, S., Manning, C., and Meyer, G. (1993). The perceptual weighting of acoustic cues changes with linguistic experience. *J. Acoust. Soc. Am.* 94, 1865–1865.

Polka, L., Rvachew, S., and Molnar, M. (2008). Speech perception by 6- to 8-month-olds in the presence of distracting sounds. *Infancy* 13, 421–439. doi: 10.1080/15250000802329297

Rhebergen, K. S., and Versfeld, N. J. (2005). A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. *J. Acoust. Soc. Am.* 117, 2181–2192. doi: 10.1121/1.1861713

Shamma, S. A., Elhilali, M., and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends Neurosci.* 34, 114–123.

Sobon, K. A., Taleb, N. M., Buss, E., Grose, J. H., and Calandruccio, L. (2019). Psychometric function slope for speech-in-noise and speech-in-speech: effects of development and aging. *J. Acoust. Soc. Am.* 145, EL284–EL290. doi: 10.1121/1.5097377

Stuart, A. (2008). Reception thresholds for sentences in quiet, continuous noise, and interrupted noise in school-age children. *J. Am. Acad. Audiol.* 19, 135–146. doi: 10.3766/jaaa.19.2.4

Vance, M., and Martindale, N. (2012). Assessing speech perception in children with language difficulties: effects of background noise and phonetic contrast. *Int. J. Speech Lang. Pathol.* 14, 48–58. doi: 10.3109/17549507.2011.616602

Wang, D. L. (2005). "On ideal binary mask as the computational goal of auditory scene analysis" in *Speech separation by humans and machines.* ed. P. Divenyi (New York, NY: Kluwer Academic Publishers), 181–197.

Werner, L. A., Folsom, R. C., and Mancl, L. R. (1994). The relationship between auditory brainstem response latencies and behavioral thresholds in normal hearing infants and adults. *Hear. Res.* 77, 88–98.

Wightman, F. L., and Kistler, D. J. (2005). Informational masking of speech in children: effects of ipsilateral and contralateral distracters. *J. Acoust. Soc. Am.* 118, 3164–3176. doi: 10.1121/1.2082567

Youngdahl, C. L., Healy, E. W., Yoho, S. E., Apoux, F., and Holt, R. F. (2018). The effect of remote masking on the reception of speech by young school-age children. *J. Speech Lang. Hear. Res.* 61, 420–427. doi: 10.1044/2017_JSLHR-H-17-0118

Yuen, K. C., and Yuan, M. (2014). Development of spatial release from masking in Mandarin-speaking children with normal hearing. *J. Speech Lang. Hear. Res.* 57, 2005–2023. doi: 10.1044/2014_JSLHR-H-13-0060

Rhyme Awareness in Children With Normal Hearing and Children With Cochlear Implants: An Exploratory Study

Linye Jing[1]*, Katrien Vermeire[2], Andrea Mangino[3] and Christina Reuterskiöld[1]

[1] Department of Communicative Sciences and Disorders, New York University, New York, NY, United States, [2] Department of Communication Sciences and Disorders, Long Island University, Brooklyn, NY, United States, [3] LIJ Hearing and Speech Center, Cohen Children's Medical Center, Northwell Health, New Hyde Park, NY, United States

Phonological awareness is a critical component of phonological processing that predicts children's literacy outcomes. Phonological awareness skills enable children to think about the sound structure of words and facilitates decoding and the analysis of words during spelling. Past research has shown that children's vocabulary knowledge and working memory capacity are associated with their phonological awareness skills. Linguistic characteristics of words, such as phonological neighborhood density and orthography congruency have also been found to influence children's performance in phonological awareness tasks. Literacy is a difficult area for deaf and hard of hearing children, who have poor phonological awareness skills. Although cochlear implantation (CI) has been found to improve these children's speech and language outcomes, limited research has investigated phonological awareness in children with CI. Rhyme awareness is the first level of phonological awareness to develop in children with normal hearing (NH). The current study investigates whether rhyme awareness in children with NH (n = 15, median age = 5; 5, IQR = 11 ms) and a small group of children with CI (n = 6, median age = 6; 11.5, IQR = 3.75 ms) is associated with individual differences in vocabulary and working memory. Using a rhyme oddity task, well-controlled for perceptual similarity, we also explored whether children's performance was associated with linguistic characteristics of the task items (e.g., rhyme neighborhood density, orthographic congruency). Results indicate that there is an association between vocabulary and working memory and performance in a rhyme awareness task in NH children. Only working memory was correlated with rhyme awareness performance in CI children. Linguistic characteristics of the task items, on the other hand, were not found to be associated with success. Implications of the results and future directions are discussed.

Keywords: rhyme awareness, neighborhood density, cochlear implants, vocabulary, working memory

INTRODUCTION

Successful literacy learning is the most important task for children to achieve in school. Seminal work as Liberman (1973), Lundberg et al. (1988) has shown that phonological awareness skills, a critical component of phonological processing, are closely linked to children's literacy outcomes. Phonological awareness enables children to actively analyze and reflect upon the sound structure of words. It facilitates the sound-to-letter knowledge required for decoding (i.e., reading) and encoding (i.e., spelling). To master reading and writing, children need to learn to decode written words. This decoding ability is highly dependent on phonological awareness skills, which enable children to break down speech into smaller phonological units such as words, syllables, onsets and rimes, and phonemes (see Torgesen et al., 1994; Adams, 1998).

Different tasks have been used to assess children's phonological awareness skills. In segmentation tasks, children break down a whole word into smaller phonological units by clapping out the number of syllables or sounds in a word. In identification tasks, children distinguish specific sounds within a word (e.g., Is there a/s/in "Mom"?). In manipulation tasks, children delete or substitute smaller units within a word (e.g., What is left if you take/um/away from "umbrella"?). Children are commonly asked to participate in such listening tasks during their early school years, and these tasks are included in phonological awareness tests. Strong performance in these tasks entail both sharp listening skills, as well as metalinguistic skills (i.e., making judgments about the linguistic structure of the items).

In this paper, we explore the potential relationship between different levels of hearing experience, vocabulary skills, and non-verbal working memory skills on success in a rhyme recognition task in a group of children, which includes a small group of children with cochlear implants (CI). All children with CI were congenitally deaf and implanted before the age of two. A carefully designed rhyme recognition task with a balanced rhyme density neighborhood, orthographic congruency, and the type of phoneme substitutions of the items, as well as a tight control for the perceptual saliency of phonemes, age of acquisition, and familiarity of the stimuli words, was used. This allowed us to explore how linguistic factors might be associated with accuracy in a task measuring rhyme awareness.

Development of Phonological Awareness in Children

There is a consensus that the grain size of phonological representation (i.e., syllable, onset/rime and phoneme) in typically developing (TD) children develops from larger to smaller units (Ziegler and Goswami, 2005). Onset-rime awareness is the first to appear at around age four, as shown in a seminal study by Bradley and Bryant (1983). Children were asked to identify the odd word from three to four single-syllable words with CVC (i.e., consonant – vowel – consonant) structure. The odd word differs from the rest by not sharing the same initial (e.g., *b*us, *b*un, *r*ug), medial (e.g., p*i*n, b*u*n, g*u*n) and final (e.g., do*ll*, ho*p*, to*p*) phonemes.

Results showed that the shared consonants in the initial positions (i.e., onset) as well as the combination of medial vowels and final consonants (i.e., rime) are the basis for making correct judgments in the oddity tasks. Four- and five-year-old children performed above chance level in both the onset and rime versions of the oddity task, suggesting proficiency in rhyme awareness (Bradley and Bryant, 1983; Kirtley et al., 1989). In other studies, children were asked to identify pairs of rhyming words instead of the odd word or the non-rhyming word (Carroll and Snowling, 2001). Since both paradigms assess children's ability to detect the rhyming phenomenon, some researchers also refer to this ability as "rhyme awareness."

Syllable segmentation skills also appear at around 4 years of age (Liberman et al., 1974), while phoneme awareness develops later and partly as a consequence of learning to read and write (Scarborough et al., 1998; Goswami, 2002). Liberman et al. (1974) used a tapping task to assess syllable and phoneme segmentation skills in children and found that 46% of four-year-old children could segment syllables but none could segment phonemes. In the study, 90% of six-year-old children were successful with syllable segmentation and 70% were able to segment phonemes. Taken together, these results support the notion of a large-to-small developmental trajectory of phonological awareness (i.e., from large units to small units).

As the first acquired phonological awareness skill, rhyme awareness serves as a stepping stone for the further development of a more fine-grained awareness of syllables and phonemes within a word. Extensive empirical evidence from rigorous longitudinal research has established a causal link between children's phonological awareness skills and literacy development (Stanovich, 1992; Wagner et al., 1997; Adams, 1998; Torgesen et al., 1999). Rhyme awareness was also found to be directly applied during reading in English. For example, a child knowing how to read the word *beak* finds it easier to read analogous words such as *peak*, *bean*, and *leak*. Such process is referred to as "orthographic analogy", during which children make a prediction about word pronunciation by using the shared spelling sequence between words (Goswami, 1998). Moreover, rime analogies (e.g., using *peak* to infer the pronunciation of *beak*) were found to be easier than onset analogies (e.g., using *beak* for *bean*) when children try to read unfamiliar words (Goswami, 1986). This evidence suggests that being able to identify words that rhyme is helpful to children who are learning to read.

Contributors to Phonological Awareness

Vocabulary knowledge is viewed as a support system for the development of phonological processing skills in young children. Phonological processing skills have been found to be related to vocabulary size (e.g., Edwards et al., 2004; Munson et al., 2005). Metsala and Walley (1998) have proposed the Lexical Restructuring Hypothesis, suggesting that the growth of vocabulary knowledge propels the holistic-to-segmental reorganization of phonological representation in young children. Under the pressure of a growing vocabulary, children need to differentiate between onsets, rimes, syllables, and eventually

phonemes to make more generalizations about the phonological structure of their language (Walley, 1993, 2008; Metsala, 1997).

One line of relevant research has focused on how phonological neighborhood influences children's phonological awareness performance. Phonological neighborhood is the total number of words differing from a target word by the addition, substitution or deletion of one phoneme in any position (Luce and Pisoni, 1998). For example, the neighbors of *rat* include *brat*, *rot*, and *at*. Targets from dense phonological neighborhoods have more similar words while targets from sparse neighborhoods have fewer similar words. Studies have arrived at different conclusions regarding the impact of phonological neighborhood density on phonological awareness skills. In Metsala (1999), children aged 3–4 years of age demonstrated better phoneme blending performance (e.g., select the pictures that match the word consisting of the sounds/b/.../ʊ/.../ʃ/) with words from dense neighborhoods, but this neighborhood density effect was not found in their onset-rime blending task (e.g., point to the picture with/d/.../ɪʃ/in it).

De Cara and Goswami (2003) argued that these inconsistent findings result from the one-phoneme-different definition of phonological neighborhood because young children do not have phoneme-level representations of words before literacy learning. Young children are more sensitive to the onset-rime level of phonological representations. The authors proposed that rhyme neighborhood density, which is the number of words that rhyme with each other (e.g., rat, cat, hat) would influence young children's rhyme awareness performance. They designed a rhyme oddity task that required children to listen to three words and verbally repeat the odd (i.e., non-rhyming) word (e.g., Which word is the odd one from "peak," "dot," "not"?). Words were selected from dense versus sparse rhyme neighborhoods in balanced numbers. Three types of odd words were created by altering the following phonemes in the rhyming words within a trial: a rime change (e.g., sock/rock/*win*), a vowel change (e.g., hat/rat/*neat*) and a coda change (e.g., feed/need/*deal*). Children's vocabulary sizes were measured by their raw score on the British Picture Vocabulary Scales. Results showed that four- to five-year-old children with larger vocabulary sizes were better at identifying the odd words from dense rhyme neighborhoods than words from sparse rhyme neighborhoods. This performance difference between dense versus sparse rhyme neighborhood was strongest for the coda change trials, followed by the rhyme change trials but absent for the vowel change trials. Children with weaker vocabulary skills did not show effects of either rhyme neighborhood density or its interaction with type of changes.

In a forced choice classification task, Storkel (2002) found that young children make decisions regarding which CVC word sounds alike based on the overlap in the rhyme of the word (dip – sip) for words from dense neighborhoods. For words from sparse neighborhoods, however, the manner feature of the final phoneme of the rhyme mattered in order for children to identify words as sounding alike (tug-mud). Children's segmental representation of words from dense neighborhoods is more fine-grained therefore, because they are organized by individual phonemes. Representations from sparse neighborhoods, however, are coarser since children perceive phonemes belonging to the same manner category as sounding the same.

Factors Influencing Phonological Awareness in Deaf and Hard of Hearing Children, and Those With Cochlear Implants

For deaf and hard of hearing (DHH) children, literacy is a difficult area and their average outcomes are below those of hearing children (Marschark and Spencer, 2010). One possible explanation for this poor outcome lies in the development of DHH children's phonological awareness. According to Locke's theory of neurolinguistic development (Locke, 1997), holistic utterances accrued between the fifth to seventh month of young children's lives form a foundation for analytical reconstruction and the acquisition of phonology, morphology and grammar from 20 to 37 months. Absent or degraded auditory input in DHH children compromises this process and may cause these children to treat the incoming speech signal in larger chucks, such as syllables rather than in phonemes (Briscoe et al., 2001). Indeed, DHH children have been found to have poor performance in tasks assessing rhyme awareness and phoneme awareness (Hanson and Fowler, 1987; Campbell and Wright, 1988, 1990; Harris and Beech, 1998; Sterne and Goswami, 2000).

Recent development in cochlear implant (CI) technology has offered a potential opportunity for profoundly deaf children to receive early auditory input, and achieve better literacy outcomes (Geers, 2003; Lyxell et al., 2008). Individual differences such as age of implantation and working memory have also been investigated in terms of their influence on CI children's literacy and pre-literacy skills. Yet only a limited number of studies have explored whether CI improve DHH children's phonological awareness.

A series of recent studies have been conducted by Nittrouer et al. (2012) and colleagues focusing on language and literacy outcomes in children with CI. In the first study, 50 children who had participated in an earlier study between the ages of 12 to 48 months participated at the end of their kindergarten year. The group consisted of children with CI, children with hearing loss wearing hearing aids, and children with normal hearing (NH). Outcome measure was a comprehensive measure combining language comprehension, expressive vocabulary, phonological awareness, literacy skills, narrative skills and speed of processing. Results showed that language comprehension before the age of 24 months was the best predictor for later success. Other strong predictors after the age of 36 months, were vocabulary skills and syntactic complexity (Nittrouer et al., 2012).

In a subsequent study (Nittrouer et al., 2014), the investigators used language samples collected from kindergarteners to investigate how children with CI and children with NH differ in terms of grammatical skills in spontaneous production during personal narratives. Measures of phonological awareness and lexical knowledge were also included. Results showed that children with CI performed at one standard deviation below the control group on language measures, including lexical skills, but two standard deviations below on measures

of phonological awareness. Lexical knowledge accounted for variance on three measures of language. One measure of phonological awareness, sensitivity to word-final phonemic structure, as well as number of bound morphemes accounted for variance above and beyond lexical knowledge. No factors related to hearing loss or intervention, except age at first implant, explained variance on language measures. The authors concluded by recommending intervention explicitly supporting grammatical skills for children with CI.

Morphosyntactic and phonological structure appeared to be mutually independent in second graders with NH, but not in children with CI according to results from Nittrouer et al. (2016). The authors found that the development of sensitivity to early predictors for phonological performance in children with CI included auditory comprehension and MLU. Predictors for morphosyntactic skills included MLU and expressive vocabulary. Children with CI were also followed up in 6th grade in Nittrouer et al. (2018). Phonological, lexical and morphosyntactic abilities were measured. It was found that compared to children with NH, deficits remained fairly consistent since earlier studies. The main area of concern was phonological skills, followed by lexical and morphosyntactic skills. Lexical skills and phonological awareness skills developed from second to sixth grade in both children with CI and NH. There were, however, no correlations between phonological awareness and expressive vocabulary at the later point in time, which can probably be explained by the fact that there was a strong correlation between word reading skills and phonological awareness. According to Hogan et al. (2005) phonological awareness and word reading are so strongly correlated at 2nd grade and after, that phonological awareness will not add additional information. It is clear from the studies cited above, however, that phonological awareness remains an area of vulnerability in children with CI.

In a longitudinal study, James et al. (2005) found that 5 to 10-year old children with CI initially had better syllable awareness than rhyme or phoneme awareness and they made significant improvement in their rhyme awareness over a period of 12 months. The authors claimed that the use of CI promotes the advancement of phonological awareness following the syllable – rhyme – phoneme developmental trajectory in TD children with NH. Additionally, the initial phonological awareness of children with CI were compared with a group of profoundly deaf children and another group of severely deaf children, both of which were using hearing aids (HAs) instead of CI. Children with CI were found to have the same level of syllable awareness as the less impaired group with better levels of residual hearing and using HA, but the same level of rhyme awareness skills as the profoundly deaf children using HA. The latter group had similar levels of residual hearing as the children with CI before implantation. The author concluded that cochlear implants benefited DHH children's syllable awareness, but not rhyme awareness.

In James et al. (2007), two groups of children with CI were recruited. The early group included children implanted between 2 and 3.6 years and the late group children included implanted between 5 and 7 years. Another group of younger reading-matched children with NH also participated. Z-scores

were calculated for the performance of children with NH performance in a number of phonological awareness tasks. Phonological awareness performance of the early group fell on the lower end of NH children's z-score distribution, while late-implanted children's scores fell mostly below the distribution. The early group also achieved greater progress over time than the late group overall. Notably, some late-implanted children demonstrated the most improvement. In Johnson and Goswami (2010), early-implanted children (i.e., before the age of three) were also found to have equivalent rhyme awareness performance compared to reading-level matched peers with NH, while late-implanted children (i.e., later than 43 months) had significantly lower performance. When they combined children with CI who performed above chance level from both the early and late groups, they found that these children's performances were not significantly different from that of their reading matched peers. This suggests that time of implantation is not the only decisive factor. The fact that age of implantation is not the only factor that matters for positive outcomes has also been illustrated in a study by Willstedt-Svensson et al. (2004). These authors found that the best predictor of lexical and grammatical development in children with CI was the percentage correctly imitated vowels in a non-word repetition task, instead of age of implantation. Other factors that are important for a positive outcome are length and quality of intervention, as well as interaction style of parents (Nittrouer, 2010). Overall, these studies suggest that a CI does offer a better chance for DHH children to acquire typical phonological awareness skills. Early implantation is generally more beneficial, but individual outcomes are highly variable.

Another line of research, has investigated the association between verbal working memory, short-term phonological memory (STPM), and the development of language skills in children (Gathercole and Baddeley, 1993). Typically, working memory (WM) tasks are thought to involve both maintenance of information and some type of manipulation simultaneously, which is also the case in phonological awareness tasks. STPM on the contrary, is considered a subskill of WM and only involves rote memory span, such as in a forward digit span task (Kronenberger et al., 2013). It has been shown in a multitude of studies of children with CI that verbal working memory skills, typically measured by digit span tests, is an area of vulnerability (Pisoni and Cleary, 2003; Pisoni et al., 2011; Kronenberger et al., 2013). AuBuchon et al. (2015) showed that even when digit spans are presented visually, WM performance in CI users is lower than that of individuals with typical hearing. The authors suggested that this population experience WM weaknesses that go beyond issues related to audibility and speech production. They provided an explanation that stresses the importance of auditory input for the development of phonological representations in long-term memory, which supports reactivation and recovery in a short-term memory task.

Researchers have used a non-word repetition task and a non-word discrimination task as an index of STPM in children with CI. Non-word repetition is traditionally used to assess the function of the phonological loop in the Baddeley and Gathercole model of working memory (Baddeley, 1986; Gathercole and Baddeley, 1990a). There is a large body of research demonstrating

a link between non-word repetition skills and language abilities in children (e.g., Gathercole and Baddeley, 1990a,b; Montgomery, 1995; Sahlén et al., 1999a,b). Some researchers have also used the Competing Language Processing Task (CLPT, Gaulin and Campbell, 1994) to assess WM skills in a dual-processing task. Ibertsson et al. (2009) found that children and adolescents aged 11 to 19, who were CI users, performed poorer on non-word repetition and non-word discrimination compared to the results of NH children aged 5, 7, and 10 pulled from other studies. The CI group's performance was similar to that of the 14- to 15-year-olds with NH on the WM task, which includes dual processing. Willstedt-Svensson et al. (2004) used non-word repetition, non-word discrimination and an adapted version of the CLPT (Towse et al., 1998) to study STPM, WM, as well as novel word learning in fifteen children 5 to 11 years old with CI devices. Children were congenitally deaf and had received their implants between the age of 2 and 6 years of age. Findings indicated that age of implantation was linked to performance in a novel word learning task. There was also a correlation between performance in the non-word repetition task and the WM task with novel word learning ability. In a paper presenting an overview of studies focusing on cognitive development and communication skills in Swedish-speaking children with CI, Lyxell et al. (2008) found that in tasks requiring phonological processing, CI users typically perform at lower levels than individuals with NH. In other WM tasks, however, the difference between groups is not as prominent, and sometimes even absent. CI user performance on non-verbal WM tasks was investigated by Cleary et al. (2001). These investigators created a WM task requiring memory for sequences of visual-spatial cues or the same cues paired with auditory signals. Children with CI and NH were asked to reproduce each sequence by pressing buttons on a response box. Results showed that the CI users obtained shorter spans on both tasks than the NH children. The children with CI also showed a smaller gain with the addition of auditory cues compared to the NH group. The authors concluded that the results indicate atypical WM development regardless of input modality. This study indicates that auditory deprivation during the first years of life may affect areas above and beyond language, such as WM.

Orthographic information is yet another factor influencing children's performance in phonological awareness tasks. "Orthographic congruency" describes whether or not the phonological information and the orthographic information of words lead to the same phonological judgment. For example, Campbell and Wright (1988) compared rhyme awareness in DHH children and children with NH. Children were shown pictures of "d*og*/fr*og*" (i.e., congruent) and "h*air*/b*ear*" (i.e., incongruent). In congruent trials, the rimes of the words were spelled and pronounced the same while in incongruent trials, they were spelled differently. Results showed that both children with NH had higher accuracy with congruent trials while DHH children only made correct rhyme judgments with the congruent trials. Research on syllable awareness (Sterne and Goswami, 2000) and phoneme awareness (Miller, 1997) have also found a similar effect of orthographic congruency. Taken together, these studies show that children rely on orthographic information in

phonological awareness tasks, but DHH children rely on such information to a larger degree.

The relationship between vocabulary, phonological neighborhood density and phonological awareness in children with CI is less studied. Dillon et al. (2012) found a possible relationship between larger vocabulary size and more robust phonological representations in children with CI. It is unknown if rhyme awareness in children who were implanted early is subject to a rhyme neighborhood density effect and if performance is linked to vocabulary. Children with CI do not tend to reach the same level of vocabulary development as children with NH (Yoshinaga-Itano et al., 2010). Some research has shown that children implanted by the age of 2 have a better chance of achieving receptive vocabulary skills within normal range, however (Hayes et al., 2009). Kirk et al. (1995) found that children with CI are sensitive to phonological neighborhood density in speech recognition the same way as children with NH are. Therefore, it is possible that CI children have the same sensitivity to rhyme neighborhood density as NH children in phonological awareness tasks. However, weaker vocabulary skills may take a toll on CI children's development of phonological awareness skills.

Assessments of phonological awareness in children with CI could be skewed for three reasons. First, assessment tools fail to recognize that some English phonemes are harder to identify than others, even for people with NH (Cutler et al., 2004). This fact denies fair assessment for children with CI, who may receive auditory input with poorer quality than children with NH. Carroll and Snowling (2001) found that phonologically similar non-rhyming words were the most difficult for children with NH to reject in a rhyme matching task. It is reasonable to assume that children with CI would be even more confused with phonologically similar items. Secondly, when making phonological judgments, DDH children rely more on orthographic transparency (e.g., Sterne and Goswami, 2000), but assessment tools typically do not take this into account. Finally, most assessment tools do not include words from balanced phonological neighborhoods. Meanwhile children with normal NH were found to perform better with words from dense phonological neighborhoods in a phoneme blending task (Metsala, 1999) and in a rhyme oddity task (De Cara and Goswami, 2003).

Aims of the Current Study and Hypotheses

It is known that general oral language skills matter for the development of phonological awareness skills (Cooper et al., 2002), but in this study we focused on the importance of vocabulary skills for success in a rhyme recognition task. We use a rhyme recognition task (i.e., oddity task), with items created to only contain sound changes with maximal differences in terms of perceptual saliency (Cutler et al., 2004), from dense and sparse rhyme neighborhoods and controlled for orthographic congruency. The study was guided by the following questions:

1. Do *individual differences* in vocabulary knowledge and working memory capacity predict children's performance on a rhyme recognition task?

We predict that children with better vocabulary knowledge and stronger working memory capacity will perform better in a rhyme recognition task. This prediction is based on past evidence of positive correlations between children's rhyme awareness skills and vocabulary size or working memory capacity.

2. How do *linguistic characteristics* of words (i.e., rhyme neighborhood density, orthographic congruency and type of sound changes) influence children's performance in a rhyme oddity task?

Based on De Cara and Goswami (2003), we anticipate that only children with larger vocabulary size will be influenced by rhyme neighborhood density, such that their accuracy will be higher for words from dense rhyme neighborhoods. We also predict that the performance of children with larger vocabulary size will be mediated by the trial types. In coda change trials, children's accuracy for words from dense rhyme neighborhoods would be significantly higher than words from sparse rhyme neighborhoods. Such differences will not be as prominent in vowel change or rhyme change trials. Children with smaller vocabulary sizes will not show effects of either rhyme neighborhood density or its interaction with type of changes.

Additionally, we expect that children will perform better on orthographically congruent trials than incongruent trials. This prediction is based on past findings that both children with NH and CI rely on orthographic information when making rhyme judgments.

MATERIALS AND METHODS

Participants

Fifteen children with NH (mean age = 5; 2, SD = 10 months) and six congenitally deaf children (mean age = 6; 10, SD = 6 months) with cochlear implants participated in the study. Participants were recruited through distribution of flyers at medical centers, university clinics and public spaces (e.g., libraries, cafés, etc.). Written informed consent was obtained from the parents of all participating children in the study. All the children's primary language was English. Two children with CI were bilaterally implanted and the other four were unilaterally implanted and used a hearing aid on the contralateral ear. All children with CI were implanted before the age of two. Demographic information of all children is listed in **Tables 1**, **2**.

Procedure

Children completed four standardized tests and a rhyme oddity task in a random order to avoid an effect of fatigue on results. Vocabulary was assessed by the Peabody Picture Vocabulary Test – 4 (PPVT-4, Dunn and Dunn, 2007). Children were asked to point to a picture, from a selection of four, that represented the word the experimenter spoke. Non-verbal intelligence was assessed by the Primary Test of Non-verbal Intelligence (PTONI,

TABLE 1 | Demographic information for all participating children.

	NH	CI
	median (IQR)	median (IQR)
Chronological age (year; month)	5; 5 (11 months)	6; 11.5 (3.75 months)
PTONI (standard score)	121 (28)	119.5 (20.5)
PPVT (standard score)	121.5 (11.75)	84.5 (4)
PPVT (raw score)	114.5 (34.25)	89 (7.5)
Block recall[1]	4 (1)	4 (0.75)

[1]*Span scores on the block recall in Working Memory Test Battery for Children (WMTB-C).*

Ehrler and McGhee, 2008). This task required children to select a picture that did not belong to a set, in terms of visual patterns, by pointing. General language ability was assessed with the Test of Early Language Development, fourth edition (TELD-4, Hresko et al., 2017) for all except one child, who was given the Clinical Evaluation of Language Fundamentals – Preschool 2 (CELF-Preschool 2, Semel et al., 2004). Working memory was measured by the block recall subtest in the Working Memory Test Battery for Children (WMTB-C, Pickering and Gathercole, 2001), which is a non-verbal task where the child points to series of blocks following the sequence modeled by the experimenter. Children with CI completed the experimental procedure in the same way as children with NH, without any adaptation.

The Rhyme Oddity Task

To assess rhyme awareness, a rhyme oddity task adapted by De Cara and Goswami (2003) was used. The task consisted of 36 trials of three words: two words rhyming with each other, and one word not rhyming with the other two. The non-rhyming word's position in each trial was semi-randomized, which resulted in six different semi-randomized versions of the task. Each child received one version of the task, with the 36 trials presented in a fully randomized order.

Children saw a picture of a boy looking and listening attentively, which prompted the beginning of each trial. Then an icon of a loudspeaker appeared on the computer screen, while the audio of the first word was played simultaneously. This was then followed by a second speaker icon and the second word; and the third speaker icon and the final word with previous speakers remaining on the screen. Children were instructed to point to the loudspeaker that played the "non-rhyming" word at the end of each trial.

Prior to the experimental trials, a training session was provided. The children first played a rhyming game where the experimenter presented three printed pictures of objects (e.g., star, *egg*, car). Children were asked to point to the non-rhyming picture after the experimenter named the three pictures. After demonstrating an understanding of the task, children moved on to "play this game on the computer." The computerized task began with six practice trials. In the first two practice trials, the experimenter paused and explained the procedure in a step-by-step manner (e.g., "Do you see the little boy? We need to really listen now! First you will see a speaker and it will play a word. . . . Can you point to the word that does not rhyme with

TABLE 2 | Demographic information for children with CI.

ID	Age	Sex	SES[1]	Age at identification (ms)	Age at HA fitting (ms)	intervention	Communication mode	Etiology	School environment	Age at implant (Type of CI processor)		Initial stimulation (ms)	Aided PTA[4]
										I[2]	II[3]		
1	6;10	F	MA	1.5	6	EI	Auditory oral	Unknown, genetics negative	Mainstreamed in Kindergarten, prior School for deaf	0;9 (Cochlear N5)		unknown	16 dB
2	6;10	F	MA	1.5	6	EI	Auditory oral	Unknown, genetics negative	Mainstreamed in Kindergarten, prior School for deaf	0;9 (Cochlear N5)		unknown	16 dB
3	7;2	F	BA	1.5	3	EI	Auditory oral	Unknown	Mainstreamed school	1;0 (Cochlear N5)	4;11 (Cochlear N6)	14	15 dB
4	7;4	F	BA	18	11	Speech-language	Auditory oral	Parents reported Kawasaki disease	School for Deaf	1;9 (Cochlear N5)		unknown	10 dB
5	5;10	M	MA	3	5.5	EI	Total communication	Genetics	School for Deaf	0;10 (Cochlear N5)		18	20 dB
6	7;1	M	N/A	4	4	Speech-language	Auditory oral	Unknown	School for Deaf	0;9 (Cochlear N6)		unknown	26 dB

[1]SES (maternal education). [2,3]"I" and "II" indicate first and second implantation (if applicable). [4]Aided PTA (most recently measured prior to the test session).

the other two?"). Children who understood and followed the first two practice trials completed the next four practice trials independently and proceeded to the experiment. Children who had problems with the rhyming game or practice trials were able to repeat any part of the training until they fully understood.

The stimuli words were recorded by a native female speaker of American English using a professional digital recorder (i.e., Fostex FR-2LE). The sound file was edited and normalized in the Audacity software for computer presentation. The stimuli were presented to the children from a laptop computer (i.e., Thinkpad X230) and through a loudspeaker (i.e., Mackie MR mk3) with a Behringer U-control UCA222 soundcard. The stimuli were presented at a 22.05 kHz sampling rate and 65 dB SPL. The speaker was positioned approximately 1 m in front of the children at 0° azimuth.

Stimuli

Words in the rhyme oddity task were well-controlled for phonological similarity. The stimuli in the rhyme oddity task were single-syllable words with an initial consonant (i.e., onset), a middle vowel (i.e., vowel) and a final consonant (i.e., coda). The vowel and the coda form the rime of words. The perceptual qualities of the vowels, codas, or the rimes in the non-rhyming words were created to be maximally different from their counterparts in the rhyming words by using confusion matrices in Cutler et al. (2004). The confusion matrices provide information about the likelihood of mistaking an English vowel or consonant for another one in background noise by listeners with typical hearing (e.g., confusing /p/ for /b/). In the current study, the vowels, codas, and the rimes in the non-rhyming words were the least likely to be confused with those in the rhyming words. In past research, none of the rhyme oddity tasks or rhyme matching tasks using auditory stimuli have taken into consideration the perceptual similarities between speech sounds. It is possible that any performance differences between words from dense versus sparse rhyme neighborhoods may have been affected by the lack of control of perceptual similarities in the rhyming items. The current study circumvents this problem by including stimuli that are as perceptually different as possible.

Three linguistic characteristics of the stimuli words were manipulated in the rhyme oddity task. First, words were selected from both dense and sparse rhyme neighborhoods using the auditory database reported in De Cara and Goswami (2002). Eighteen trials have words from dense rhyme neighborhoods (hereafter dense trials) and the other 18 words from sparse neighborhoods (hereafter sparse trials). A t-test validated that the dense versus sparse manipulation was significant. The mean rhyme neighborhood density for the dense stimuli was 25.3 (SD = 4.0) and the mean rhyme neighborhood density for the sparse stimuli was 7.7 (SD = 2.9), $t(53) = 25.89$, $p < 0.001$.

Additionally, three types of non-rhyming words were created by altering the following phonemes in the rhyming words within a trial: a "rime change" (e.g., sock/rock/*win*), a "vowel change" (e.g., hat/rat/*neat*) and a "coda change" (e.g., feed/need/*deal*).

Finally, orthographic congruency of the stimuli was also controlled by having the rimes (VC2) in half of the rhyming words spelled congruently (e.g., feed/need) and the other half

spelled incongruently (e.g., date/wait). Children did not see the spellings of the stimuli, rather, they needed to listen and select the non-rhyming word based on auditory input. These manipulations were made to reveal if children with CI and NH are influenced by orthographic information when making rhyme judgments in an auditory mode.

Word familiarity and age of acquisition were also controlled for in the stimuli. The familiarity ratings of all words were above 6.75 on a 1 to 7 scale as reported in Luce and Pisoni (1998). The age of acquisition ratings are below age 4; 22 using a 1–7 scale (Ages 0–2 = 1, 2–4 = 2, above 13 = 7) (Cortese and Khanna, 2008). Stimuli words and summary statistics for the variables of interest are shown in the **Supplementary Appendix A**.

Statistical Analysis

We first investigated whether group differences existed between children's age, hearing experience, language and cognitive abilities. One NH child did not return for their second session, resulting in missing data in the PPVT and block recall tests. Therefore, this child's data was not included in the group comparison tests for these two scores. For children with CI, their hearing experience was quantified by their length of amplification use with CI.[1] For children with NH, experience receiving postnatal auditory input, equals their chronological age. PPVT raw score was used as a proxy for children's "absolute vocabulary size," which is common practice in past literature investigating the relationship between phonological processing and vocabulary development (e.g., Gathercole et al., 1991; Metsala, 1999). Wilcoxon rank sum tests were used for group comparisons on children's chronological age, hearing experience, PPVT raw score and standard score, general language standard score, PTONI standard score and block recall raw score.

Participants received binary scoring for the rhyme oddity task. To answer the first research question concerning the relationship between individual differences and rhyme awareness, a generalized mixed-effect logistic regression was fitted to this binary outcome variable using the lme4 package (Bates et al., 2015) in RStudio Version 1.0.136 (R Development Core Team, 2017) and following Harel and McAllister (2019). The fixed effect structure included the following predictor variables: PPVT raw score, Block recall span score, Group (NH versus CI), and interactions between Group and all the other variables. All predictor variables except for group were transformed into z-scores to facilitate model convergence. The Group variable was sum-coded to allow for interpretation of other variables as overall predictors of accuracy performance. The random effects included test items and participant.

To answer the second research question concerning the association between linguistic characteristics and children's rhyme awareness, a second mixed-effect logistic regression was fitted to participants' binary accuracy data. The fixed effect structure included the following predictor variables: Group,

PPVT_r, RND, Ortho, Change, two-way interactions between Group and PPVT_r, PPVT_r and Change, PPVT_r and RND, as well as a three-way interaction term between PPVT_r, RND and Change. Again, Group, RND, Ortho and Change were sum-coded to allow for interpretation of other variables as overall predictors of accuracy performance. The random effects included test items and participant.

RESULTS

Group Comparison

Results from the Wilcoxon rank sum tests (**Table 3**) revealed that NH children's chronological age was significantly lower than that of the children with CI ($z = -2.56$, $p = 0.01$), but that the group of CI children's time with CI amplification was similar to the chronological age of the NH children ($z = -1.14$, $p = 0.13$). NH children had significantly higher language scores ($z = -3.04$, $p < 0.001$) and vocabulary scores (PPVT raw scores $z = -1.85$, $p = 0.03$) compared with children with CI. However, there were no group differences on any of the non-language related measures including non-verbal intelligence (PTONI, $z = -0.46$ $p = 0.32$) or working memory capacity (Block Recall, $z = 0.24$, $p = 0.59$).

Individual Differences

Spearman's correlations of predictor variables are summarized in **Table 4**. Correlations are shown without a Bonferroni correction, since this procedure is overly conservative according to Perneger (1998). Results from our first model (**Table 5**) showed significant effects of group ($\beta = -0.36$, $p < 0.001$) suggesting that children with CI had lower average performance than children with NH at the rhyme awareness task. The association between PPVT_r and rhyme awareness was significant ($\beta = 0.05$, $p < 0.001$), with a positive slope indicating that, on average,

TABLE 3 | Wilcoxon rank sum tests results comparing NH and CI on their age, hearing age and standardized tests scores.

		n	*z*-value	*p*
Age	NH	15	−2.56	*p* = 0.01*
	CI	6		
Hearing experience	NH	15	−1.14	*p* = 0.13
	CI	6		
PPVT_s[1]	NH	14	−3.24	*p* < 0.001*
	CI	6		
PPVT_r[2]	NH	14	−1.85	*p* = 0.03*
	CI	6		
Language[3] (standard score)	NH	15	−3.04	*p* < 0.001*
	CI	6		
PTONI (standard score)	NH	15	−0.46	*p* = 0.32
	CI	6		
Block recall	NH	14	0.24	*p* = 0.59
	CI	6		

[1,2]*PPVT_s: PPVT standard score derived from chronological age; PPVT_r: PPVT raw score.* [3]*One child completed CELF for language assessment while all other children completed TELD.*

---

[1]Some CI children received amplification through hearing aids before their CI implantation, yet the auditory benefit of their hearing aids was deemed inadequate which is why they qualified to be eligible for CIs.

TABLE 4 | Spearman's correlation matrix for independent variables.

	1	2	3	4	5	6
Chron. Age ($n = 21$)	–					
PPVT_s ($n = 20$)	−0.39***	–				
PPVT_r ($n = 20$)	0.05	0.83***	–			
General language ($n = 21$)	−0.35***	0.81***	0.59***	–		
Block recall span ($n = 20$)	0.46***	0.12**	0.31***	0.19***	–	
PTONI ($n = 21$)	0.16***	0.26***	0.35***	0.58***	0.26***	–

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 5 | Regression results for individual differences.

	Estimate	Std. error	Statistic	p value
(Intercept)	0.77	0.00	393.42	$p < 0.001$
Group[1]	−0.36	0.00	−189.02	$p < 0.001$
PPVT_r[2]	0.05	0.00	25.38	$p < 0.001$
WM[3]	0.82	0.00	429.41	$p < 0.001$
Group × PPVT_r	−0.47	0.00	−244.78	$p < 0.001$
Group × WM	−0.06	0.00	−32.53	$p < 0.001$
(Random effect) item	0.59			
(Random effect) subject	0.45			

[1]Group: CI (cochlear implant) or NH (normal hearing). [2]PPVT_r, Peabody Picture Vocabulary Test raw score. [3]WM, working memory.

children with larger vocabulary size were more successful at the task. The interaction between group and vocabulary was significant ($\beta = -0.47$, $p < 0.001$), suggesting that the slopes for vocabulary were different between children with NH and CI, as can be seen in **Figure 1A**. The association between WM and rhyme awareness was significant ($\beta = 0.82$, $p < 0.001$) with a positive slope suggesting that children with better WM skills had better rhyme awareness performance. The interaction between group and WM was also significant ($\beta = -0.06$, $p < 0.001$), suggesting that the slopes for WM were different between children with NH and CI, as can be seen in **Figure 1B**.

To probe these two interactions, we divided the children based on groups (NH versus CI) and performed two additional models on these two groups, respectively. In **Table 6**, results for NH children show that the association between PPVT_r and rhyme awareness was significant ($\beta = 0.53$, $p < 0.05$) with a positive slope suggesting that NH children with larger vocabulary size were more successful at the task. The association between WM and rhyme awareness was also significant ($\beta = 0.90$, $p < 0.001$), with a positive slop suggesting that NH children with better working memory skills had better rhyme awareness performance. In **Table 7**, results showed that the association between PPVT_r and rhyme awareness was not significant in the CI group. The association between WM and rhyme awareness was significant in the CI group ($\beta = 0.68$, $p < 0.05$) with a positive slope suggesting

that CI children with better WM skills had better rhyme awareness performance.

Characteristics of Items in the Rhyme Recognition Task

As illustrated in **Table 8**, results from our second mixed-effects logistic model did not show a significant effect for Group, PPVT_r, Change or Ortho. There was no significant interaction between Group and PPVT_r, PPVT_r and RND, PPVT_r and Change and no significant three-way interaction between PPVT_r, RND and Change.

Qualitative Analyses of Vocabulary Size, Rhyme Awareness and Linguistic Characteristics

We conducted two additional descriptive analyses to qualitatively explore the relationship between vocabulary size, rhyme awareness and linguistic characteristics. In the first analysis, we plotted bivariate relationships between three pairs of variables: PPVT raw score and chronological age; PPVT standard score and chronological age; rhyme awareness performance and chronological age (**Figures 2A–C**). **Figure 2A** shows a pattern of increasing PPVT raw scores in NH children with increasing chronological age. This pattern was still present for the NH children when PPVT scores were reported as standard scores (**Figure 2B**). There are only six children with CI and therefore no clear conclusions can be made, but the same pattern does not seem to be present in this small group during visual inspection (**Figures 2A,B**). Both CI and NH children appeared to perform better in the rhyme awareness task with increasing age based on visual inspection of the graphs (**Figure 2C**).

The second analysis was a qualitative exploration of which type of non-rhyming words were the most challenging for children with NH and CI, respectively. NH and CI children performance on the trials containing non-rhyming words with a C2, V, and VC changes were plotted in **Figure 3**. Visual qualitative inspection revealed that children with NH performed similarly with the three types of non-rhyming words. Children with CI seemed to be slightly more challenged when the non-rhyming word differed from the rhyming word by a change in the middle vowel (V-change).

DISCUSSION

In this study we explored how vocabulary skills and working memory matter for phonological awareness skills in children. We included a small group of six congenitally deaf children with CI, who had been implanted before the age of two. Compared to many previous studies, which have included children with a wide range of age of implantation, our group all children had been implanted early. The children with CI were older than the NH children, but had similar hearing experience and non-verbal intelligence. In the rhyme recognition task, we intentionally maximized the difference of perceptual saliency of words within each trial to grant fair assessment of rhyme awareness in children

FIGURE 1 | Rhyme oddity task performance as a function of vocabulary **(A)** and working memory **(B)**.
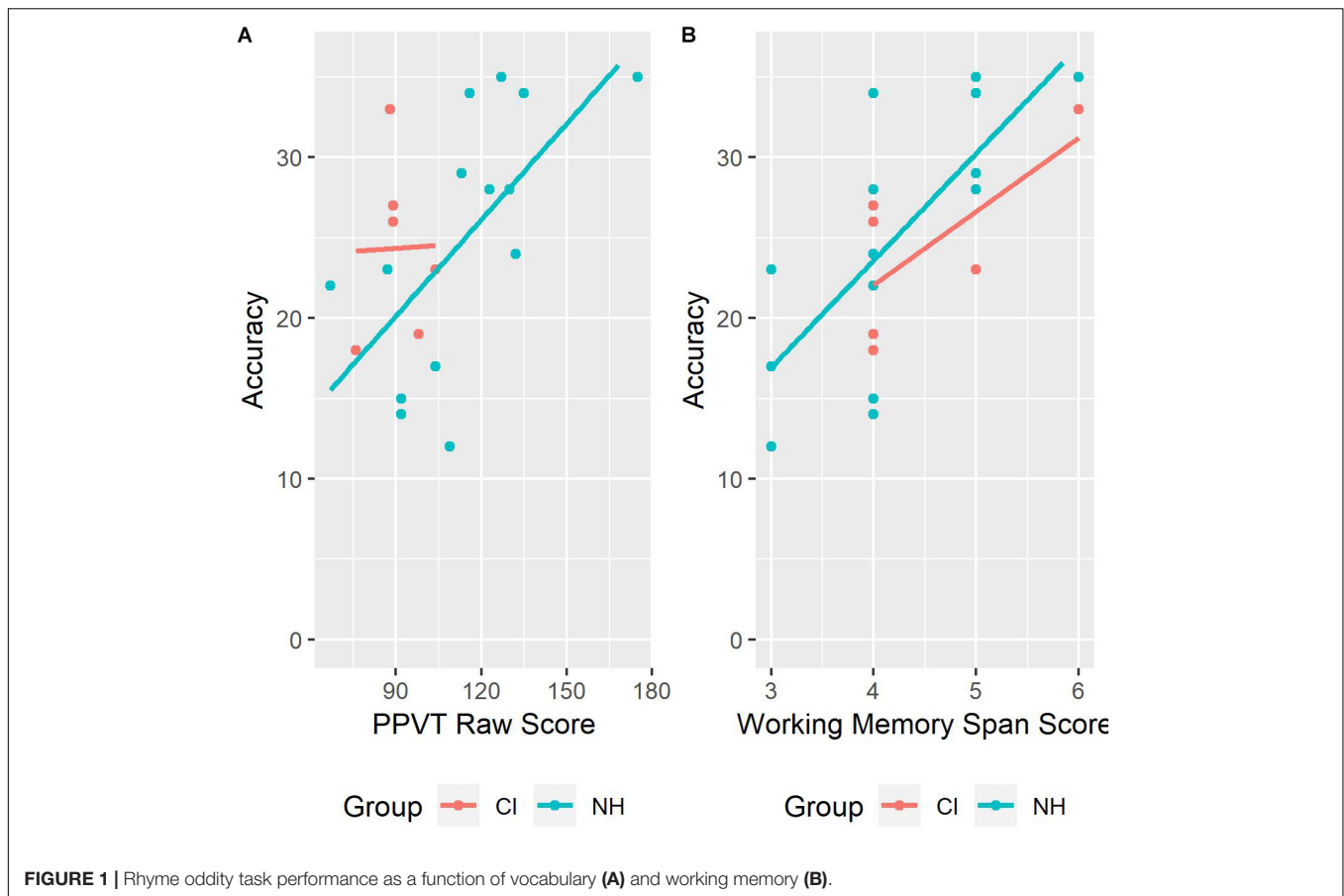
TABLE 6 | Regression results for individual difference in the NH group.

	Estimate	Std. error	Statistic	*p* value
(Intercept)	1.19	0.20	5.91	$p < 0.001$
PPVT_r	0.53	0.23	2.27	$p < 0.05$
WM	0.90	0.23	3.96	$p < 0.001$
(Random effect) item	0.54			
(Random effect) Subject	0.48			

TABLE 7 | Regression results for individual differences in the CI group.

	Estimate	Std. error	Statistic	*p* value
(Intercept)	0.94	0.27	3.48	$p < 0.001$
PPVT_r	−0.11	0.24	−0.44	$p = 0.66$
WM	0.68	0.29	2.35	$p < 0.05$
(Random effect) item	0.77			
(Random effect) Subject	0.41			

with CI. Making sure that the non-rhyming word in each trial has a minimal probability of perceptual confusion with the rhyming words is of particular importance when assessing phonological processing in children with hearing impairments. Poorer success rates compared to children with NH may otherwise not be a function of poorer phonological processing skills but may be secondary to less optimal auditory input.

Our results show that vocabulary size measured by PPVT raw scores, predicted success in the rhyme awareness task among children with NH. Other studies have found that phonological processing skills are related to vocabulary size (e.g., Edwards et al., 2004; Munson et al., 2005). In Metsala (1999), performance on phonological awareness tasks was related to overall vocabulary size, age of acquisition of words, and neighborhood density.

Researchers have shown that vocabulary skills are important for the development of phonological awareness skills and have suggested that the holistic to segmental development of phonological awareness skills is a secondary effect of vocabulary acquisition. As a child learns more words, there is a need to make distinctions between increasingly smaller segments because many words have dense phonological neighborhoods (Metsala and Walley, 1998). The children with CI in our study had poorer vocabulary skills compared with the NH children, which is consistent with previous research showing that vocabulary skills develop slower in children with CI (Yoshinaga-Itano et al., 2010). We did not find a positive correlation between vocabulary size and rhyme awareness in our children with CI. This finding is in contrast with the

TABLE 8 | Regression results for linguistic characteristics.

	Estimate	Std. error	Statistic	p value
(Intercept)	0.89	0.43	2.07	p = 0.04
Group	−0.03	0.43	−0.06	p = 0.95
PPVT_r	0.50	0.54	0.92	p = 0.36
RND[1]	−0.12	0.11	−1.06	p = 0.29
Change1[2]	0.03	0.16	0.22	p = 0.83
Change2	−0.25	0.15	−1.66	p = 0.10
Ortho[3]	0.00	0.10	−0.03	p = 0.97
Group × PPVT_r	−0.51	0.54	−0.95	p = 0.34
PPVT_r × RND	0.00	0.12	−0.04	p = 0.97
PPVT_r × Change1	0.18	0.17	1.04	p = 0.30
PPVT_r × Change2	−0.20	0.16	−1.27	p = 0.20
PPVT_r × RND × Change1	0.06	0.16	0.39	p = 0.70
PPVT_r × RND × Change2	−0.14	0.15	−0.92	p = 0.36
(Random effect) item	0.54			
(Random effect) subject	0.85			

[1] RND, rhyme neighborhood density. [2] Change: type of changes in the rime-ending of the non-rhyming words. [3] Ortho, orthographic congruency.

results from Dillon et al. (2012) who found that in children with CI vocabulary size was a mediating factor between reading skills and phonological awareness skills. In their study, there was a weaker correlation between phonological awareness and reading when vocabulary was controlled. **Figures 2A,B** in our study, show that one CI child was slightly younger than

the remaining five, and had a lower PPVT raw score. In the older five children with CI, the PPVT standard score had a negative slope, indicating that the vocabulary skills of these children might not have developed following a predicted pattern over time. There was a positive correlation between accuracy rates in the rhyme awareness task and chronological age, however, which might indicate that other factors were more important in supporting these children in developing their phonological awareness skills. Since our study has a small sample size of children with CI, we remain cautious in interpreting these results.

Contrary to our expectation, we did not find a significant interaction between rhyme neighborhood density and vocabulary size, as measured by PPVT raw score. Children with larger vocabulary sizes performed comparably with words from dense versus sparse neighborhoods and so did children with smaller vocabulary sizes. One explanation may be that our version of the rhyme oddity task is less taxing compared to the earlier version in De Cara and Goswami (2003), since we intentionally minimized the perceptual similarity between trial words. Storkel (2002) found that children had more detailed segmental representation of words from dense neighborhoods than words from sparse neighborhoods. Consequently, children found it more difficult to judge whether words sound the same when these words were from sparse neighborhoods. In words from sparse neighborhoods, children perceived words ending with sounds from the same category in terms of manner of articulation as the same (tug-mud).
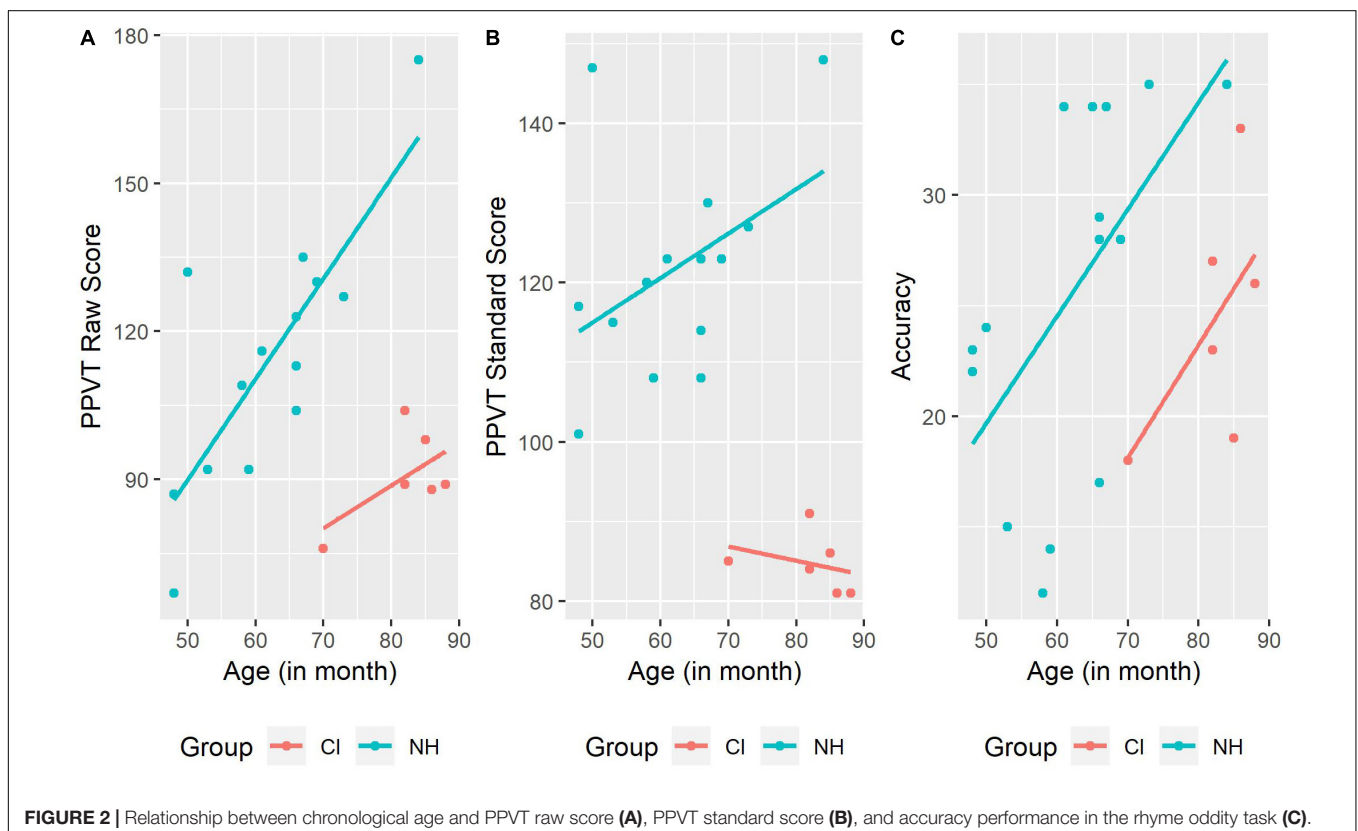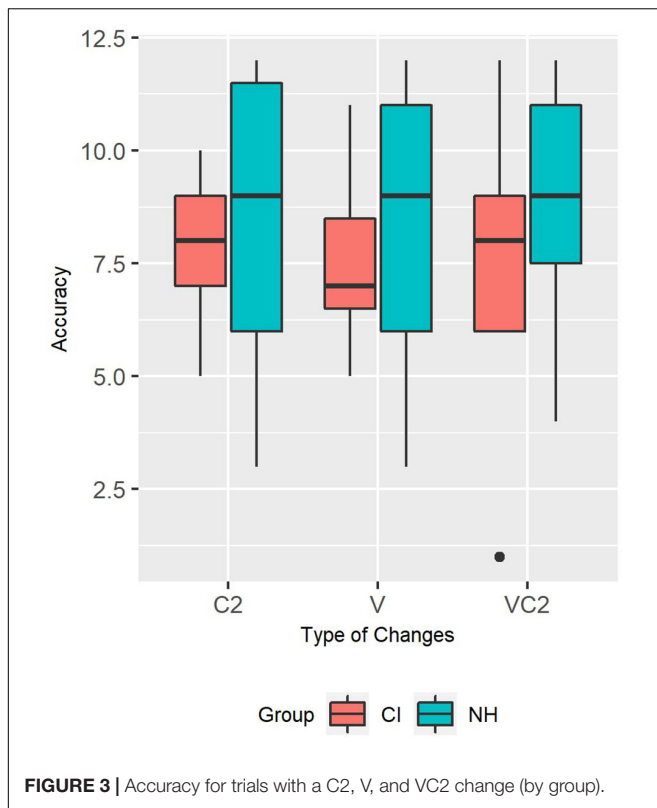

FIGURE 2 | Relationship between chronological age and PPVT raw score **(A)**, PPVT standard score **(B)**, and accuracy performance in the rhyme oddity task **(C)**.

FIGURE 3 | Accuracy for trials with a C2, V, and VC2 change (by group).

However, since our stimuli from sparse neighborhoods were made to be maximally different from each other, this might have reduced the level of difficulty while children made decisions about rhyming. This may be a reason why children showed similar performance with words from dense versus sparse neighborhoods.

We were not able to replicate the three-way interaction between vocabulary size, rhyme neighborhood density and type of changes reported in De Cara and Goswami (2003). Our results indicated that children's performance was equally accurate in the coda change, vowel change, and rime change trials and no rhyme neighborhood effects were shown in any type of changes. This null finding is, however, consistent with some earlier studies, in which no performance differences were found between coda change conditions and vowel change conditions (Bradley and Bryant, 1983; Kirtley et al., 1989; Bryant et al., 1990). One explanation provided by De Cara and Goswami (2003) for their novel finding is that their rhyme oddity task with 5-year-olds used pre-recorded speech stimuli. The auditory nature of the stimuli did not provide lip cues. Therefore, children could only rely on linguistic cues to make rhyme judgments. Since a coda change trial provide the least number of linguistic cues (i.e., a consonant) compared to the vowel and the rhyme change trials, it is the most linguistically demanding condition and might be the most discriminative condition for detecting an effect of rhyme neighborhood density. Our rhyme oddity task was reduced in terms of perceptual similarity between trial words, however. This might have caused a loss of discriminating power in the coda trials, and thus suppressed rhyme neighborhood density

effects. As can be seen during visual inspection of **Figure 3**, our children with CI seemed to be most challenged by rhyme changes including a vowel change. Perhaps CI children tend to rely on acoustic information carried in the vowel when processing speech, which made this sound change particularly difficult in spite of the fact that we had made changes as salient as possible.

Many of the participating children were old enough to have been exposed to orthographic forms in reading and may have stored not only phonological forms of words, but also orthographic forms. It is not well known how orthographic representations support individuals in phonological processing tasks, although we know that orthographic support facilitates word learning in children with developmental language disorders (Ricketts et al., 2015). Our results revealed no significant effects for orthographic congruency, however. Past studies that have identified such effect have either used written tasks, or a picture identification task without any auditory stimuli (e.g., Campbell and Wright, 1988; Miller, 1997; Sterne and Goswami, 2000). In written tasks, readily available information of orthographic congruency would have a direct impact on children's rhyme judgments. In picture identification tasks, children must access the phonological information of the words through lexical retrieval, which may activate of the words' orthographies. Children in our study only heard the pronunciation of the stimulus words and might have processed and analyzed the phonological components of these words without activating their orthographic representation. As a result, orthographic congruency did not show an influence on children's performance in the rhyme oddity task.

Non-verbal working memory skills were not different between children with NH and children with CI with similar hearing experience. On the surface level, this result contradicts the results from Cleary et al. (2001), where children with CI performed worse than children with NH on tasks assessing non-verbal working memory. However, a closer look revealed that the CI children in their study had shorter hearing experience than the chronologically age-matched children with NH. Correlation coefficients in the current study (**Table 4**) also showed that working memory scores had a stronger correlation with hearing experience than with chronological age. Together, this suggests that hearing experience contributes to working memory skills in children with CI. Our finding that non-verbal working memory predicts children's rhyme awareness is consistent with previous findings that phonological processing skills are linked to children's short-term memory skills regardless of hearing status (Pisoni and Geers, 2000; Pisoni and Cleary, 2003; Willstedt-Svensson et al., 2004).

To summarize, we found that both vocabulary size and non-verbal working memory skills are important factors for rhyme recognition skills in children with NH. In children with CI, only working memory was found to be significant. However, vocabulary learning is still important for children with CI. The children with CI in our study had poorer vocabulary skills than children with NH. Past research (Dillon et al., 2012) has found a positive relationship between vocabulary and children's phonological awareness skills. Nittrouer et al. (2018) did not find a strong correlation between expressive vocabulary and

phonological awareness in 6th grade children with NH or with CI, however. Our study has a very limited sample size of children with CI, and therefore results are difficult to generalize. For our NH children, the results indicate a positive relationship between vocabulary skills and rhyme awareness, which is consistent with earlier studies on children with NH (e.g., Metsala and Walley, 1998; Edwards et al., 2004; Munson et al., 2005). Finally, working memory skills are important for phonological awareness tasks regardless of hearing status. This finding is expected based on previous literature, and also suggests that mentally comparing items in a phonological awareness task involves a memory component.

The current study is a first attempt to use a rhyme recognition task with a stringent control of perceptual similarity of distinguishing phonemes, which might have reduced the level of difficulty in task. Increasing the level of saliency of the distinguishing phonemes in the task may have had an effect on how rhyme neighborhood density or type of rhyme changes in our task played a role. This may also be a reason why we did not find an effect of orthographic congruency. Future studies might examine whether different levels of perceptual similarities of stimuli would have an effect on children's performance in rhyme awareness tasks. Such studies may also lead to the development of balanced stimuli to be included in standardized rhyme awareness tests. Task administration was randomized. Randomization may, however, have affected the robustness of the correlations. The most important limitation of the current study is the small number of children in the CI group. The small sample size also makes it difficult to investigate the impact of background characteristics and other factors, such as parental engagement on children's rhyme awareness skills. In future studies the goal will be to include a more balanced number of participants in the groups to study phonological processing skills in this population.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the New York University Committee on Activities Involving Human Subjects, the Institutional Review Board of Northwell Health, and the NYU Langone Medical School Office of Science and Research Institutional Review Board with written informed consent from all subjects' caregivers. All subjects' caregivers were given written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the New York University Committee on Activities Involving Human Subjects, the Institutional Review Board of Northwell Health, and the NYU Langone Medical School Office of Science and Research Institutional Review Board.

AUTHOR CONTRIBUTIONS

LJ designed the study, collected the data, analyzed the results, and wrote the manuscript. KV contributed to the design of the study through her expertise in audiology, and provided input regarding the analyses. AM collected the data from children with CI and scored standardized tests. CR participated in all aspects of the project except the data collection.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg. 2019.02072/full#supplementary-material

REFERENCES

Adams, M. J. (1998). *Beginning to Read: Thinking and Learning About Print (10. Print)*. Cambridge, MA: MIT Press.

AuBuchon, A. M., Pisoni, D. B., and Kronenberger, W. G. (2015). Short-term and working memory impairments in early-implanted, long-term cochlear implant users are independent of audibility and speech production. *Ear Hear.* 36, 733–737. doi: 10.1097/AUD.0000000000000189

Baddeley, A. D. (1986). *Working Memory. Oxford Psychology Series, 11*. Oxford: Clarendon Press.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Bradley, L., and Bryant, P. E. (1983). Categorizing sounds and learning to read-a causal connection. *Nature* 301, 419–421. doi: 10.1038/301 419a0

Briscoe, J., Bishop, D. V. M., and Frazier Norbury, C. (2001). Phonological processing, language, and literacy: a comparison of children with mild-to-moderate sensorineural hearing loss and those with specific language impairment. *J. Child Psychol. Psychiatr.* 42, 329–340. doi: 10.1017/S0021963001007041

Bryant, P. E., MacLean, M., Bradley, L. L., and Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Dev. Psychol.* 26, 429–438. doi: 10.1037/0012-1649.26.3.429

Campbell, R., and Wright, H. (1988). Deafness, spelling and rhyme: how spelling supports written word and picture rhyming skills in deaf subjects. *Q. J. Exp. Psychol. A* 40, 771–788. doi: 10.1080/14640748808402298

Campbell, R., and Wright, H. (1990). Deafness and immediate memory for pictures: dissociations between "inner speech" and the "inner ear"? *J. Exp. Child Psychol.* 50, 259–286. doi: 10.1016/0022-0965(90)90042-7

Carroll, J. M., and Snowling, M. J. (2001). The effects of global similarity between stimuli on children's judgment of rime and alliteration. *Appl. Psycholinguist.* 22, 327–342. doi: 10.1017/S0142716401003034

Cleary, M., Pisoni, D., and Geers, A. (2001). Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants. *Ear Hear.* 22, 395–411. doi: 10.1097/00003446-200110000-00004

Cooper, D. H., Roth, F. P., Speece, D. L., and Schatschneider, C. (2002). The contribution of oral language skills to the development of phonological awareness. *Appl. Psycholinguist.* 23, 399–416. doi: 10.1017/S0142716402003053

Cortese, M. J., and Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behav. Res. Methods* 40, 791–794. doi: 10.3758/BRM.40.3.791

Cutler, A., Weber, A., Smits, R., and Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *J. Acoust. Soc. Am.* 116, 3668–3678. doi: 10.1121/1.1810292

De Cara, B., and Goswami, U. (2002). Similarity relations among spoken words: the special status of rimes in English. *Behav. Res. Methods Instrum. Comput.* 34, 416–423. doi: 10.3758/bf03195470

De Cara, B., and Goswami, U. (2003). Phonological neighbourhood density: effects in a rhyme awareness task in five-year-old children. *J. Child Lang.* 30, 695–710. doi: 10.1017/S0305000903005725

Dillon, C. M., de Jong, K., and Pisoni, D. B. (2012). Phonological awareness, reading skills, and vocabulary knowledge in children who use cochlear implants. *J. Deaf Stud. Deaf Educ.* 17, 205–226. doi: 10.1093/deafed/enr043

Dunn, L. M., and Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test.* Upper Saddle River, NJ: Pearson.

Edwards, J., Beckman, M., and Benjamin, M. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *J. Speech Lang. Hear. Res.* 47, 421–436. doi: 10.1044/1092-4388(2004/034)

Ehrler, D. J., and McGhee, R. L. (2008). *PTONI: Primary Test of Nonverbal Intelligence.* Austin, TX: Pro-Ed.

Gathercole, S. E., and Baddeley, A. D. (1990a). Phonological memory deficits in language disordered children: is there a causal connection? *J. Mem. Lang.* 29, 336–360. doi: 10.1016/0749-596X(90)90004-J

Gathercole, S. E., and Baddeley, A. D. (1990b). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *Br. J. Psychol.* 81, 439–454. doi: 10.1111/j.2044-8295.1990.tb02371.x

Gathercole, S. E., and Baddeley, A. D. (1993). Phonological working memory: a critical building block for reading development and vocabulary acquisition? *Eur. J. Psychol. Educ.* 8:259. doi: 10.1007/BF03174081

Gathercole, S. E., Willis, C., and Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: reading and vocabulary development in children. *Br. J. Psychol.* 82, 387–406. doi: 10.1111/j.2044-8295.1991.tb02407.x

Gaulin, C. A., and Campbell, T. F. (1994). Procedure for assessing verbal working memory in normal school-age children: some preliminary data. *Percept. Mot. Skills* 79, 55–64. doi: 10.2466/pms.1994.79.1.55

Geers, A. E. (2003). Predictors of reading skill development in children with early cochlear implantation. *Ear Hear.* 24(Suppl.), 59S–68S. doi: 10.1097/01.AUD.0000051690.43589.5D

Goswami, U. (1986). Children's use of analogy in learning to read: a developmental study. *J. Exp. Child Psychol.* 42, 73–83. doi: 10.1016/0022-0965(86)90016-0

Goswami, U. (1998). "The role of analogies in the development of word recognition," in *Word Recognition in Beginning Literacy*, eds J. L. Metsala and L. C. Ehri (Mahwah, NJ: Lawrence Erlbaum Associates Publishers).

Goswami, U. (2002). Phonology, reading development, and dyslexia: a cross-linguistic perspective. *Ann. Dyslexia* 52, 139–163. doi: 10.1007/s11881-002-0010-0

Hanson, V. L., and Fowler, C. A. (1987). Phonological coding in word reading: evidence from hearing and deaf readers. *Mem. Cognit.* 15, 199–207. doi: 10.3758/BF03197717

Harel, D., and McAllister, T. (2019). Multilevel models for communication sciences and disorders. *J. Speech Lang. Hear. Res.* 62, 783–801. doi: 10.1044/2018_JSLHR-S-18-0075

Harris, M., and Beech, J. R. (1998). Implicit phonological awareness and early reading development in prelingually deaf children. *J. Deaf Stud. Deaf Educ.* 3, 205–216. doi: 10.1093/oxfordjournals.deafed.a014351

Hayes, H., Geers, A. E., Treiman, R., and Moog, J. S. (2009). Receptive vocabulary development in deaf children with cochlear implants: achievement in an intensive auditory-oral educational setting. *Ear Hear.* 30, 128–135. doi: 10.1097/AUD.0b013e3181926524

Hogan, T. P., Catts, H. W., and Little, T. D. (2005). The relationship between phonological awareness and reading. *Lang. Speech Hear. Serv. Sch.* 36, 285–293.

Hresko, W. P., Reid, D. K., and Hammill, D. D. (2017). *Test of early language development: TELD-4.* Austin, TX: Pro-ed, Inc.

Ibertsson, T., Hansson, K., Asker-Àrnason, L., and Sahlén, B. (2009). Speech recognition, working memory and conversation in children with cochlear implants. *Deafness Educ. Int.* 11, 132–151. doi: 10.1002/dei.261

James, D., Rajput, K., Brinton, J., and Goswami, U. (2007). Phonological awareness, vocabulary, and word reading in children who use cochlear implants: does age of implantation explain individual variability in performance outcomes and growth? *J. Deaf Stud. Deaf Educ.* 13, 117–137. doi: 10.1093/deafed/enm042

James, D., Rajput, K., Brown, T., Sirimanna, T., Brinton, J., and Goswami, U. (2005). Phonological awareness in deaf children who use cochlear implants. *J. Speech Lang. Hear. Res.* 48, 1511–1528. doi: 10.1044/1092-4388(2005/105)

Johnson, C., and Goswami, U. (2010). Phonological awareness, vocabulary, and reading in deaf children with cochlear implants. *J. Speech Lang. Hear. Res.* 53, 237–261. doi: 10.1044/1092-4388(2009/08-0139)

Kirk, K. I., Pisoni, D. B., and Osberger, M. J. (1995). Lexical effects on spoken word recognition by pediatric cochlear implant users. *Ear Hear.* 16, 470–481. doi: 10.1097/00003446-199510000-00004

Kirtley, C., Bryant, P., MacLean, M., and Bradley, L. (1989). Rhyme, rime, and the onset of reading. *J. Exp. Child Psychol.* 48, 224–245. doi: 10.1016/0022-0965(89)90004-0

Kronenberger, W. G., Pisoni, D. B., Harris, M. S., Hoen, H. M., Xu, H., and Miyamoto, R. T. (2013). Profiles of verbal working memory growth predict speech and language development in children with cochlear implants. *J. Speech Lang. Hear. Res.* 56, 805–825. doi: 10.1044/1092-4388(2012/11-0356)

Liberman, I. Y. (1973). 1. Segmentation of the spoken word and reading acquisition. *Bull. Orton Soc.* 23, 64–77. doi: 10.1007/BF02653842

Liberman, I. Y., Shankweiler, D., Fischer, F. W., and Carter, B. (1974). Explicit syllable and phoneme segmentation in the young child. *J. Exp. Child Psychol.* 18, 201–212. doi: 10.1016/0022-0965(74)90101-5

Locke, J. L. (1997). A theory of neurolinguistic development. *Brain Lang.* 58, 265–326. doi: 10.1006/brln.1997.1791

Luce, P. A., and Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear Hear.* 19, 1–36. doi: 10.1097/00003446-199802000-00001

Lundberg, I., Frost, J., and Petersen, O.-P. (1988). Effects of an extensive program for stimulating phonological awareness in preschool children. *Read. Res. Q.* 23, 263–284. doi: 10.1598/rrq.23.3.1

Lyxell, B., Sahlén, B., Wass, M., Ibertsson, T., Larsby, B., Hällgren, M., et al. (2008). Cognitive development in children with cochlear implants: relations to reading and communication. *Int. J. Audiol.* 47(Suppl. 2), S47–S52. doi: 10.1080/14992020802307370

Marschark, M., and Spencer, P. E. (2010). Promises (?) of deaf education: from research to practice and back again. *Oxford Handbook Deaf Stud. Lang. Educ.* 2, 1–4.

Metsala, J. L. (1997). An examination of word frequency and neighborhood density in the development of spoken-word recognition. *Mem. Cogn.* 25, 47–56. doi: 10.3758/bf03197284

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *J. Educ. Psychol.* 91, 3–19. doi: 10.1037/0022-0663.91.1.3

Metsala, J. L., and Walley, A. C. (1998). "Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability," in *Word Recognition in Beginning Literacy*, eds J. L. Metsala and L. C. Ehri (Mahwah, NJ: Erlbaum).

Miller, P. (1997). The effect of communication mode on the development of phonemic awareness in prelingually deaf students. *J. Speech Lang. Hear. Res.* 40, 1151–1163. doi: 10.1044/jslhr.4005.1151

Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: the role of phonological working memory. *J. Speech Hear. Res.* 38, 187–199. doi: 10.1044/jshr.3801.187

Munson, B., Swenson, C., and Manthei, S. (2005). Lexical and phonological organization in children. *J. Speech Lang. Hear. Res.* 48, 108–124. doi: 10.1044/1092-4388(2005/009)

Nittrouer, S. (2010). *Early Development of Children with Hearing Loss*. San Diego, CA: Plural Publishing.

Nittrouer, S., Caldwell, A., and Holloman, C. (2012). Measuring what matters: effectively predicting language and literacy in children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 76, 1148–1158. doi: 10.1016/j.ijporl.2012.04.024

Nittrouer, S., Lowenstein, J. H., and Holloman, C. (2016). Early predictors of phonological and morphosyntactic skills in second graders with cochlear implants. *Res. Dev. Disabil.* 55, 143–160. doi: 10.1016/j.ridd.2016.03.020

Nittrouer, S., Muir, M., Tietgens, K., Moberly, A. C., and Lowenstein, J. H. (2018). Development of phonological, lexical, and syntactic abilities in children with cochlear implants across the elementary grades. *J. Speech Lang. Hear. Res.* 61, 2561–2577. doi: 10.1044/2018_JSLHR-H-18-0047

Nittrouer, S., Sansom, E., Low, K., Rice, C., and Caldwell-Tarr, A. (2014). Language structures used by kindergartners with cochlear implants: relationship to phonological awareness, lexical knowledge and hearing loss. *Ear Hear.* 35, 506–518. doi: 10.1097/AUD.0000000000000051

Perneger, T. V. (1998). What's wrong with bonferroni adjustments. *BMJ* 316, 1236–1238. doi: 10.1136/bmj.316.7139.1236

Pickering, S., and Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB-C): Manual*. London: Pearson.

Pisoni, D., and Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear Hear.* 24, 106S–120S. doi: 10.1097/01.AUD.0000051692.05140.8E

Pisoni, D., and Geers, A. (2000). Working memory in deaf children with cochlear implants: correlations between digit span and measures of spoken language processing. *Ann. Otol., Rhinol. Laryngol.* 185, 92–93. doi: 10.1177/0003489400109s1240

Pisoni, D., Kronenberger, W., Roman, A., and Geers, A. (2011). Article 7: measures of digit span and verbal rehearsal speed in deaf children following more than 10 years of cochlear implantation. *Ear Hear.* 32, 60S–74S. doi: 10.1097/AUD.0b013e3181ffd58e

R Development Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ricketts, J., Dockrell, J. E., Patel, N., Charman, T., and Lindsay, G. (2015). Do children with specific language impairment and autism spectrum disorders benefit from the presence of orthography when learning new spoken words? *J. Exp. Child Psychol.* 134, 43–61. doi: 10.1016/j.jecp.2015.01.015

Sahlén, B., Reuterskiold-Wagner, C., Nettelbladt, U., and Radeborg, K. (1999a). Non-word repetition in children with language impairment: pitfalls and possibilities. *Int. J. Lang. Commun. Disord.* 34, 337–352. doi: 10.1080/136828299247441

Sahlén, B., Wagner, C. R., Nettelbladt, U., and Radeborg, K. (1999b). Language comprehension and non-word repetition in children with language impairment. *Clin. Linguist. Phon.* 13, 369–380. doi: 10.1080/026992099299031

Scarborough, H. S., Ehri, L. C., Olson, R. K., and Fowler, A. E. (1998). The fate of phonemic awareness beyond the elementary school years. *Sci. Stud. Read.* 2, 115–142. doi: 10.1207/s1532799xssr0202_2

Semel, E., Wiig, E. H., and Secord, W. A. (2004). *Clinical Evaluation of Language Fundamentals-Preschool-2 (CELF-Preschool-2)*. San Antonio, TX: Psychological Corp.

Stanovich, K. E. (1992). "Speculations on the causes and consequences of individual differences in early reading acquisition," in *Reading Acquisition*, eds P. B. Gough, L. C. Ehri, and R. Treiman (Hillsdale, NJ: Lawrence Erlbaum Associates, Inc).

Sterne, A., and Goswami, U. (2000). Phonological awareness of syllables, rhymes, and phonemes in deaf children. *J. Child Psychol. Psychiatr.* 41, 609–625. doi: 10.1111/1469-7610.00648

Storkel, H. L. (2002). Restructuring of similarity neighbourhoods in the developing mental lexicon. *J. Child Lang.* 29, 251–274. doi: 10.1017/s0305000902005032

Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (1994). Longitudinal studies of phonological processing and reading. *J. Learn. Disabil.* 27, 276–286. doi: 10.1177/002221949402700503

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: group and individual responses to instruction. *J. Educ. Psychol.* 91, 579–593. doi: 10.1037/0022-0663.91.4.579

Towse, J. N., Hitch, G. J., and Hutton, U. (1998). A reevaluation of working memory capacity in children. *J. Mem. Lang.* 39, 195–217. doi: 10.1006/jmla.1998.2574

Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R., et al. (1997). Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Dev. Psychol.* 33:468. doi: 10.1037/0012-1649.33.3.468

Walley, A. C. (1993). The role of vocabulary development in childrens spoken word recognition and segmentation ability. *Dev. Rev.* 13, 286–350. doi: 10.1006/drev.1993.1015

Walley, A. C. (2008). *In Blackwell Handbooks in Linguistics: Vol. The Handbook of Speech Percpetion*. Malden: Blackwell Publishing.

Willstedt-Svensson, U., Löfqvist, A., Almqvist, B., and Sahlén, B. (2004). Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *Int. J. Audiol.* 43, 506–515. doi: 10.1080/14992020400050065

Yoshinaga-Itano, C., Baca, R. L., and Sedey, A. L. (2010). Describing the trajectory of language development in the presence of severe-to-profound hearing loss: a closer look at children with cochlear implants versus hearing aids. *Otol. Neurotol.* 31, 1268–1274. doi: 10.1097/MAO.0b013e3181f1ce07

Ziegler, J. C., and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychol. Bull.* 131, 3–29. doi: 10.1037/0033-2909.131.1.3

Predictors of Reading Comprehension in Children With Cochlear Implants

Malin Wass[1]*, Lena Anmyr[2,3], Björn Lyxell[4,5], Elisabet Östlund[6†], Eva Karltorp[2,6] and Ulrika Löfkvist[2,4]

[1] Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden, [2] Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Stockholm, Sweden, [3] Department of Social Work in Health, Karolinska University Hospital, Stockholm, Sweden, [4] Department of Special Needs Education, University of Oslo, Oslo, Norway, [5] Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden, [6] Department of Otorhinolaryngology, Karolinska University Hospital, Stockholm, Sweden

Children with a profound hearing loss who have been implanted with cochlear implants (CI), vary in terms of their language and reading skills. Some of these children have strong language skills and are proficient readers whereas others struggle with language and both the decoding and comprehension aspects of reading. Reading comprehension is dependent on a number of skills where decoding, spoken language comprehension and receptive vocabulary have been found to be the strongest predictors of performance. Children with CI have generally been found to perform more poorly than typically hearing peers on most predictors of reading comprehension including word decoding, vocabulary and spoken language comprehension, as well as working memory. The purpose of the current study was to investigate the relationships between reading comprehension and a number of predictor variables in a sample of twenty-nine 11–12-year-old children with profound hearing loss, fitted with CI. We were particularly interested in the extent to which reading comprehension in children with CI at this age is dependent on decoding and receptive vocabulary. The predictor variables that we set out to study were word decoding, receptive vocabulary, phonological skills, and working memory. A second purpose was to explore the relationships between reading comprehension and demographic factors, i.e., parental education, speech perception and age of implantation. The results from these 29 children indicate that receptive vocabulary is the most influential predictor of reading comprehension in this group of children although phonological decoding is, of course, fundamental.

Keywords: reading comprehension, children with CI, vocabulary, word decoding, cochlear implants, simple view of reading, lexical quality hypothesis

INTRODUCTION

Children with a profound hearing loss who have been implanted with cochlear implants (CI), show substantial variation in reading skills. Some children have been reported to read well within the normal range of hearing peers on measures of word decoding and reading comprehension (e.g., Dillon et al., 2011). Many others, however, struggle with both the decoding and comprehension aspects of reading (Geers, 2003; Kyle and Harris, 2006; Harris and Terlektsi, 2010; Geers and Hayes, 2011).

Some previous research has been focused on the causes of variation in reading ability within this group of children (e.g., Connor and Zwolan, 2004; Dillon et al., 2011;

Von Muenster and Baker, 2014) but the use of different predictor variables between studies as well as the heterogeneity of the children included in the research makes it difficult to draw general conclusions. Examples of the variation that can be seen across studies, are the children's age range, main communication mode, and the predictors that have been measured.

This study set out to investigate the cognitive and linguistic predictors that are known to be most relevant for reading comprehension in children with typical hearing, in a group of 11–12-year-old children with profound hearing loss who use CI. This is in contrast to previous studies on reading comprehension in children with CI which have mostly been focused on demographic factors (c.f. Connor and Zwolan, 2004; Dillon et al., 2011) and/or have included children with broad age ranges. The children included in this research used mainly oral communication and the majority of them were bilaterally implanted.

The theoretical background of reading comprehension and its main cognitive and linguistic predictors, as documented in typically hearing children, is reviewed below, followed by a summary of findings from previous research on children with CI.

Reading Comprehension in Children With Typical Hearing

One of the most fundamental prerequisites for reading comprehension is the ability to efficiently decode written words. Early reading typically involves the effortful grapheme-phoneme conversion by which children sound words out by adding and blending letter sounds (Coltheart et al., 2001). Word decoding then gradually becomes more automatized and effortless as whole words are recognized instantly by sight, so called *orthographic word recognition* (Ehri, 2005, 2014). Thereby, more cognitive resources can be used for comprehension and the acquisition of new information from the text (Perfetti, 2007).

In addition to word decoding, the reader further needs language skills that enable him or her to understand what is being read. The relative importance of decoding and language skills for reading comprehension has been found to vary depending on the children's age (e.g., Ouellette and Beers, 2010; Melby-Lervåg and Lervåg, 2014). That is, decoding is relatively more important in the early stages of reading development whereas language and vocabulary generally plays a greater role for children who have learned to master basic word reading skills (e.g., Lervåg and Aukrust, 2010; Melby-Lervåg and Lervåg, 2014). The nature of the language skills that are most relevant for reading comprehension is explained differently in two models of reading comprehension; *the Simple View of Reading* (Gough and Tunmer, 1986; Tunmer and Chapman, 2012) and the *Lexical Quality Hypothesis* (Perfetti, 2007).

The Simple View of Reading (SVR) suggests that reading comprehension is constituted by two components: word decoding and comprehension of oral language and that both components are equally important (Gough and Tunmer, 1986; Tunmer and Chapman, 2012). According to Tunmer and Chapman (2012), language comprehension is a hypothetical construct, which can be split up into component processes such

as the retrieval of individual words in lexical memory (receptive vocabulary) and the knowledge about how words and syntactic structures should be used. The broad definition of language comprehension in the SVR makes it difficult to measure with precision (e.g., Ouellette and Beers, 2010) and an increasing number of correlational studies suggest that larger proportions of reading comprehension are explained by variance in vocabulary than by listening comprehension (Braze et al., 2007; Protopapas et al., 2007; Verhoeven and Van Leeuwe, 2008; Ouellette and Beers, 2010; Olson et al., 2011), in particular for children beyond the earliest stages of reading development.

The lexical quality hypothesis (LQH, Perfetti, 2007; Perfetti and Stafura, 2014), on the other hand, stresses the importance of word knowledge and assumes that it, together with decoding ability, is the most central component of reading comprehension. According to the LQH, the quality of word representations within any reader's vocabulary, varies depending on how familiar the reader is with the word in terms of several aspects including lexical meaning, pragmatic use, and orthographic and phonological characteristics (Perfetti, 2007). The LQH assumes that knowledge of word meaning affects reading comprehension not only indirectly via its effect on listening comprehension but also directly. This view is supported by results from hierarchical regression analyses which show that vocabulary significantly contributes to reading comprehension beyond the effects of language comprehension (Ouellette and Beers, 2010; Perfetti and Stafura, 2014).

Irrespective of the theoretical framework applied in research, there has been some confusion regarding the definition of the decoding component in reading comprehension, whether it refers to phonological decoding, orthographic word recognition or both (Ouellette and Beers, 2010; Tunmer and Chapman, 2012). According to Tunmer and Greaney (2010) the most sensitive measures of decoding should be expected to vary depending on children's level of reading development. That is, for beginning readers, phonological decoding is the most frequently used decoding strategy, which should be used as the main measure of decoding whereas word recognition or even speed of word recognition should be used as more sensitive measures of decoding for advanced readers. According to Tunmer and Chapman (2012), a composite measure of both phonological decoding and orthographic word recognition is suitable for assessment of decoding skill for a broad range of readers.

Other cognitive skills that predict additional variance in children's reading comprehension include working memory (e.g., Currie and Cain, 2015), a variable which is relatively more important for longer passages of text, and phonological skills (Melby-Lervåg and Lervåg, 2014) which is generally more important in early stages of reading.

Reading Comprehension in Children With CI

When it comes to the general cognitive and linguistic predictors of reading comprehension, children with CI have typically been

found to perform more poorly than hearing peers on both decoding (Geers, 2003; Geers and Hayes, 2011; Nakeva von Mentzer et al., 2014), vocabulary (Geers et al., 2009; Fagan and Pisoni, 2010; Dillon et al., 2011; Coppens et al., 2013; Walker et al., 2019) and spoken language comprehension (e.g., Geers et al., 2009), as well as phonological and complex working memory (e.g., Wass et al., 2008). This would, in turn, suggest generally poorer preconditions for reading comprehension in this group of children.

A few studies have specifically investigated the relationships between reading comprehension and various predictor variables in children with CI (e.g., Connor and Zwolan, 2004; Asker-Árnason et al., 2007; Vermeulen et al., 2007; Von Muenster and Baker, 2014). The age range of the participating children is, however, typically relatively broad and the measures used to assess reading and predictors of reading vary substantially between studies.

Connor and Zwolan (2004) explored a number of demographic, cognitive and linguistic predictors of reading comprehension in ninety-one 11 year-old children with CI. They found age at implantation to have strong effects on reading comprehension (the younger the better) both directly and through its positive effects on vocabulary growth. It should be noted here that the children included in their study were implanted at 6.7 years of age on average and thus got access to oral language relatively late. This is because prelingually deaf children, who have been implanted later than 3.5 years of age, have been shown to benefit less from cochlear implantation and typically show poorer development of speech and comprehension of oral language (Kral and Sharma, 2012). The study by Connor and Zwolan did not include a measure of word decoding and thus the relative effects of decoding and oral language cannot be compared.

The children studied by Dillon et al. (2011) were implanted relatively earlier, at 2.5 years of age on average, but the age range was broader (6–14 years). Twenty-seven English-speaking children with CI were included in their study. Although there was a substantial individual variation within the group, the children performed on average within the typical range for hearing children on measures of decoding and reading comprehension whereas their receptive vocabulary was below this range. Reading comprehension, as measured by the PIAT- R (Markwardt, 1998) was further found to be strongly associated with receptive vocabulary and phonological awareness. The strength of these correlations were, however, not compared to the correlation between reading comprehension and decoding. The authors note that age at implantation was moderately correlated with non-word reading ($r = 0.56$) and reading comprehension ($r = 0.43$), and duration of implant use was strongly correlated with measures of phonological awareness and reading ($r = 0.86$).

An Australian study by Von Muenster and Baker (2014) on 47 children with unilateral CI aged 5;4–12;6 years, reported strong correlations between reading comprehension, as measured by the Neale Analysis of Reading Ability and each of the following skills: word and non-word decoding, $r \approx 0.8$–0.9, expressive and receptive language ($r \approx 0.8$). There was also a strong correlation between reading comprehension and receptive vocabulary as measured by PPVT (Dunn and Dunn, 2007), $r \approx 0.7$. Notably, none of the measures of reading (decoding and comprehension) used in their study was significantly related to measures of auditory perception, age at implantation or duration of implant use.

Results from a sample of fifty Dutch children with CI in a similar age range (7–16 years) was reported by Vermeulen et al. (2007). The authors found strong correlations between reading comprehension and measures of both word recognition and receptive vocabulary. The latter was, however, a relatively stronger predictor, explaining 29% of the variance in reading comprehension after age and educational factors had been taken into account.

To sum up, the few studies on children with CI which have investigated cognitive and linguistic factors associated with reading comprehension, have typically included children in broad age ranges. Since the predictors of reading comprehension are known to vary with age and level of reading development, it is therefore important to study the theoretically most relevant predictors in children with CI at more narrow age ranges in order to find the most important predictors of successful reading at every particular stage in development. Based on findings from typically hearing children, decoding should be expected to play a greater role for reading comprehension in younger children who may not yet read fluently whereas vocabulary should be relatively more important as children become fluent readers (c.f. Lervåg and Aukrust, 2010). Furthermore, the extent to which age at implantation, decoding and language and vocabulary factors contribute to reading comprehension should be expected to vary depending on the characteristics of the sample studied. For example, age at implantation may be more important for reading (and language) for children who have been implanted relatively late (c.f. Kral and Sharma, 2012).

The purpose of the current study was to investigate the relationships between reading comprehension and a number of cognitive and linguistic predictor variables in a sample of twenty-nine 11–12-year-old children with profound hearing loss, fitted with CI. We were particularly interested in the extent to which reading comprehension in children with CI at this age was dependent on decoding and receptive vocabulary. The predictor variables that we set out to study were word decoding, receptive vocabulary, phonological skills, and working memory.

A second purpose was to explore the relationships between reading comprehension and demographic factors, i.e., parental education, speech perception and age of implantation.

MATERIALS AND METHODS

Participants

Twenty-nine children (14 girls) participated in this study as part of a longitudinal research project on reading development and language in children with CI. Results from an earlier measurement have been reported previously

for most of the children included in the current sample (Wass et al., 2019). The inclusion criteria were that all children should be able to follow the regular national school curriculum and perform at or above the 25th percentile on Raven's Colored Progressive Matrices (Raven et al., 2003).

Written informed consent was obtained from the children and from their parents. The children were, on average, 11;8 (years; months) of age at the time of testing (range: 11;0–12;8).

The mean age at implantation of the first CI was 24 months (range 7–69 months). Twenty-six of the children (90%) had bilateral implants and were implanted with their second CI at 29 months of age on average (range 8–105 months). Three children had bimodal hearing (CI and hearing aid).

Twenty-three of the children had used oral communication only for their whole lives, 3 children had used oral communication in combination with sign support until they started to speak themselves but had exclusively used oral communication since then. Three children were reported to have used oral communication in combination with sign language from the time they were diagnosed with their hearing loss and that they still used both communication modes.

All children were tested at the hearing implant clinic, Karolinska University Hospital at their annual follow-up appointment. They also attended regular speech and listening rehabilitation at their local hospitals during the rest of the year (Wass et al., 2019).

The sample was heterogeneous in terms of cause of deafness and age of implantation of first and second CI. Etiology and age at implantation for the sample are summarized in **Table 1**.

Speech perception in quiet as measured by phonetically balanced lists was, on average, 81.1% (SD: 15.9). One child had missing speech perception data. Raven Colored Progressive Matrices test (Raven et al., 2003) was used to measure non-verbal cognitive ability and the participants' percentile scores ranged between 25 and 95.

TABLE 1 | Age at implantation and etiology.

	Mean	SD (range)
Age at CI1 (months)	24.0	17.8 (7–69)
Age at CI2 (months)	28.4	21.0 (8–105)
Etiology	#of children	Proportion
Acquired		9/29
Congenital CMV	4	
Meningitis	5	
Genetic		12/29
Unspecific heredity*	2	
Connexin 26	3	
Usher type 1	2	
Jervell-Lange Nielsen syndr.	3	
Pendred's syndr.	2	
Unknown	8	8/29

*Close family members also have a hearing impairment.

Test Measures

The Swedish reading test LäSt (Elwér et al., 2009) was used to measure decoding of words and non-words, respectively. Reading comprehension was assessed with a Swedish version of the Woodcock Reading Mastery Test (Byrne et al., 2009).

Receptive Vocabulary was measured with Peabody Picture Vocabulary Test (PPVT-III; Dunn and Dunn, 2007).

A Sentence Completion and Recall task (Wass et al., 2008) was used to measure complex working memory. In this task, the children are asked to fill in missing words in sets of sentences e.g., "Crocodiles are green. Tomatoes are . . .". After every set of sentences, the child should also repeat back the words that she/he had filled in. The sentence sets consisted of two, three, and four sentences and the total number of correctly stored and reproduced words was recorded by the test leader, with a maximum score of 18.

A phoneme deletion task (Magnusson and Nauclér, 1993) was used to assess phonological skills. In this test, the children are asked to remove phoneme segments of spoken words, e.g., "Say summer without an 's'. The maximum score is 12.

Procedure

The children were individually tested by an experienced speech language pathologist and an audiologist working at the hearing implant clinic. All language and cognitive tests were presented in random order and administered during two consecutive days, in two 1-hour sessions. The audiologist tested the children's hearing ability during the first day. The test instructions were given in oral language.

Analyses

Relationships between the various skills were analyzed in correlation and hierarchical regression analyses. We only had comparison data from typically hearing children for the reading comprehension test, for which we had results from 21 children with typical hearing who were 10–11 years of age. The average performance on the reading comprehension test was 31.6 (SD: 6.2) for these children and the reading comprehension performance of the children with CI was compared to this comparison data.

RESULTS

Means and standard deviations for all test measures are displayed in **Table 2** and correlations are displayed in **Table 3**.

Twenty-six out of the 29 children with CI (almost 90%) performed within 1 SD from the mean of the comparison group on the reading comprehension measure.

Significant bivariate correlations were found between reading comprehension and all of the cognitive/linguistic predictor variables *with r*s in the moderate-strong range.

Neither age at implantation of first or second CI nor speech perception in quiet at the time of the follow-up visit were significantly correlated with any of the measures of reading.

TABLE 2 | Tests administered; means, standard deviations, and range.

Ability	Test	Quantification	Mean (SD; range)	N
Measures of language and cognitive skills:				
Reading comprehension	Woodcock reading mastery test	Number of correctly completed sentences (maximum score = 68)	35 (6.74; 22–51)	29
Non-word decoding fluency	LäSt	Number of correctly read non-words in 2*45 s (maximum score = 126)	79.9 (20.10; 41–114)	28
Word decoding fluency	LäSt	Number of correctly read non-words in 2*45 s (maximum score = 200)	139.9 (22.1; 105–176)	28
Receptive vocabulary	PPVT-III	Number of correctly identified pictures (maximum score = 228)	152.0 (28.6; 80–194)	29
Complex working memory	Sentence completion and recall	Number of words correctly recalled (maximum score = 18)	13.62 (2.57; 6–17.5)	29
Phonological skills	Phoneme deletion	Number of correct answers (maximum score = 12)	11.72 (0.70; 9–12)	29

Missing data for one participant on non-word decoding fluency and word decoding fluency due to fatigue.

Parental education was coded as a dichotomous variable, that is children whose parents' highest education was high-school or a shorter education constituted one group ($N = 10$) and the other group had parents with a university degree ($N = 19$). The effects of parental education on children's reading ability was explored in a Mann-Whitney group comparison. There were no significant differences between the groups on chronological age or age of implantation of the first CI. The two groups, however, differed significantly on reading comprehension ($U = 35$, $p < 0.01$) and also on non-word decoding ($U = 46.5$, $p < 0.05$). The group difference on the measure of word decoding approached significance ($U = 50.0$, $p = 0.055$). Point-biseral correlations further showed that parental education was significantly correlated with receptive vocabulary,

$r_{pb} = 0.593$ ($p < 0.001$), non-word decoding, $r_{pb} = 0.435$, $p < 0.05$, and word decoding, $r_{pb} = 0.381$, $p < 0.05$, and with reading comprehension, $r_{pb} = 0.552$, $p < 0.01$.

Subsequently, a set of hierarchical regression analyses were conducted with reading comprehension as the dependent variable. The results of these analyses are presented in **Table 4** and **Table 5**. Age at testing was not included as a predictor in any of the regression analyses as it was not correlated with reading comprehension ($r = 0.044$).

In the first analysis (displayed in **Table 4**), Raven's CPM, a composite measure of decoding (LäSt words + LäSt non-words) and receptive vocabulary (PPVT-III) were used as independent variables. Raven's percentile was entered at the first step and decoding was entered at the second step. Together these two variables accounted for 38.6 percent of the variance in reading comprehension but the decoding measure did not significantly improve model fit.

When receptive vocabulary was entered at the third step, the model predicted 64.9 percent of the variance in reading comprehension. Neither non-verbal IQ nor decoding contributed significantly to reading comprehension once receptive vocabulary had been entered into the model.

A second set of hierarchical regression analyses were then conducted (**Table 5**) in which the measure of phonological decoding (LäSt non-words) was used instead of the composite measure of decoding. In this analysis, all three variables, non-verbal IQ, non-word decoding and receptive vocabulary, significantly improved model fit and together they accounted for 65.2 percent of performance on the reading comprehension test although only the beta weight for receptive vocabulary was significant.

In a third set of regression analyses we wanted to explore the effects of phonological awareness, complex working memory and parental education. The contribution of these variables was explored in three separate analyses in which Raven's CPM was entered at the first step, non-word decoding at the second step and phoneme deletion, sentence completion and recall and parental education, respectively, were entered at the third step. Neither of these variables significantly predicted reading comprehension.

TABLE 3 | Bivariate correlations.

	1	2	3	4	5	6	7	8	9	10
1. Age CI 1	1									
2. Speech perception	−0.305	1								
3. Parental education[a]	−0.132	−0.003	1							
4. Receptive vocabulary	−0.204	0.218	0.593**	1						
5. Phoneme deletion	0.032	0.265	0.341	0.375*	1					
6. Complex WM	−0.226	0.417*	0.264	0.639***	0.484**	1				
7. nw-decoding	−0.195	0.225	0.435*	0.458*	0.469*	0.393*	1			
8. wd-decoding	−0.134	0.188	0.381*	0.357	0.431*	0.396*	0.815***	1		
9. Decoding composite average	−0.171	0.216	0.427*	0.425*	0.472*	0.414*	0.948***	0.957***	1	
10. Reading comprehension	−0.152	0.069	0.552**	0.783***	0.501**	0.447*	0.512**	0.417*	0.485**	1

*Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed). ***p < 0.001. [a]Correlation coefficients denote point-biseral correlations, r_{pb}.

TABLE 4 | Significant predictors of reading comprehension.

Models	Variables	R	R^2	Adj. R^2	ΔR^2	B	SE	β	t
1	Step 1	0.540	0.292	0.265					
	NVIQ					0.174	0.053	0.540**	3.27
2	Step 2	0.621	0.386	0.336	0.094				
	NVIQ					0.134	0.054	0.417*	2.47
	Decoding					0.111	0.057	0.330	1.95
3	Step 3	0.806	0.649	0.605	0.264				
	NVIQ					−036	0.048	0.111	0.744
	Decoding					0.056	0.045	0.168	1.24
	PPVT					0.153	0.036	0.649***	4.25

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

TABLE 5 | Significant predictors of reading comprehension.

Models	Variables	R	R^2	Adj. R^2	ΔR^2	B	SE	β	t
1	Step 1	0.540	0.292	0.265					
	NVIQ					0.174	0.053	0.540**	3.27
2	Step 2	0.640	0.410	0.363	0.118				
	NVIQ					0.132	0.053	0.411*	2.50
	Non-word decoding					0.123	0.055	0.368*	2.24
3	Step 3	0.807	0.652	0.608	0.242				
	NVIQ					0.038	0.047	0.119	0.81
	Non-word decoding					0.060	0.046	0.180	1.32
	PPVT					0.149	0.037	0.633**	4.08

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

DISCUSSION

The aim of this research was to explore the predictors of reading comprehension in 11–12-year-old Swedish children with profound hearing loss using CI. The results from these 29 children indicate that receptive vocabulary is the most influential predictor of reading comprehension in this group of children although decoding is still, of course, important.

Although the focus of this research was not to compare the performance of children with CI to typically hearing peers, it should be mentioned that most of the children in our sample (26/29) performed within 1 SD below NH mean or above, on the measure of reading comprehension. These children thus had relatively high performance compared to the approximately 50% of children with CI who have previously been reported to perform within this range (e.g., Geers, 2003; Asker-Árnason et al., 2007). The difference in results may, of course, be due to the size and representativeness of our sample, at least compared to the 181 children included in the study by Geers (2003). The results may also be due to the fact that the children who participated in the study by Geers (2003) were implanted between 1990-1996 whereas the children included in the current sample were all implanted approximately a decade later and the technological advances in implant technology may thus have improved the auditory preconditions for reading development in this population of children. On the other hand, neither age at

implantation of first and second CI nor speech perception in quiet at the time of testing were significantly correlated with any of the measures of reading in our sample of children. These findings are in line with results reported by Von Muenster and Baker (2014) who did not find significant relationships between reading and hearing measures. It should be noted that the children included in the current study and the children who participated in the sample by Von Muenster and Baker were implanted at approximately the same age, i.e., slightly above 3 years. It is possible that effects of age at implantation can only be seen for children implanted at a relatively later age (c.f. Connor and Zwolan, 2004).

The results from the hierarchical regression analyses showed that receptive vocabulary was the main predictor of reading comprehension in our sample of children with CI. Interestingly, when a composite measure of word decoding and non-word decoding was used in the analysis, this composite measure of decoding failed to predict a significant proportion of variance in reading comprehension even before receptive vocabulary was added to the model. When the composite measure of decoding was replaced by the non-word decoding measure in the analysis, all three variables significantly predicted model fit although only the beta-weights for receptive vocabulary turned out to be significant. The regression model in which non-word decoding was used as a decoding measure further explained more variance in reading comprehension than the composite measure of word and non-word decoding. This

difference was not significant but it is interesting in light of the discussion about what aspects of decoding are most important for children of different ages. There may be a tendency for phonological decoding to be relatively more important for reading comprehension than composite measures of phonological and orthographic decoding for children with CI at age 11–12 years. According to Tunmer and Greaney (2010), phonological decoding is the most frequently used decoding strategy for beginning readers, and according to the current results it seems that phonological decoding is still the most influential aspect of decoding for children with CI at age 11–12.

These findings suggest that, for children with CI at this age, vocabulary is relatively more important for reading comprehension than measures of word decoding. In comparison, recent results from Bell et al. (2019) suggest that decoding is relatively more important than language measures at age 8 whereas the opposite pattern was found in an age-matched comparison group of typically hearing children. It thus seems that at age 11–12, the decoding skills of children with CI has reached the level of decoding skill at which differences in reading comprehension are, similar to typically hearing children, more dependent on vocabulary. The vocabulary knowledge of children with CI should further be expected to vary in part depending on the length of auditory deprivation before cochlear implantation (Fagan and Pisoni, 2010) and in general as children have been both diagnosed and implanted at a gradually earlier age over the last two decades. However, in the large scale study by Geers et al. (2009) on 151 children with CI who were fitted with CI:s before 24 months of age, almost 50% of the children did not have vocabulary skills within the expected range for NH children at age 5–6. Thus vocabulary skills is still an important area of linguistic development for children with CI as it is fundamental both for language abilities in general and for the development of skilled reading.

The educational implications would thus be that the focus of support and teaching in this age group should be both on broadening and deepening of the children's vocabularies and comprehension of oral language. Of course, early education needs to focus on the decoding aspects of reading but it is also important to consider vocabulary development at an early age. This may be of particular interest as new findings suggest that vocabulary depth may be hard to catch up at later ages (c.f. Walker et al., 2019). The findings from the current research are of clinical importance as delays in spoken vocabulary in children with CI have been reported in a number of studies (Geers et al., 2009; Fagan and Pisoni, 2010; Stiles et al., 2012; Coppens et al., 2013).

Group comparisons and correlation analyses demonstrated that parental education had a significant effect on both word reading and reading comprehension in our sample. The children whose parents had a university degree had significantly higher scores on reading comprehension than children whose parents had high school level education or less. When entered as a predictor variable in the hierarchical regression analysis, parental education did not contribute significantly to reading comprehension. The strong correlation between parental education and receptive vocabulary may, however, suggest that parental education has an indirect effect on reading comprehension through its effect on receptive vocabulary.

Effects of maternal education on children's language and reading ability have indeed been found in a number of studies on both children with typical hearing (Dollaghan et al., 1999; Magnusson, 2007) and on children with hearing loss (Lieu et al., 2010; Yoshinaga-Itano et al., 2010). The results regarding parental education in the current study should, however, be interpreted with caution as the number of participants in the two groups differed substantially.

Neither complex working memory nor phonological skills contributed significantly to reading comprehension in our group of children with CI. This is not surprising considering the fact that not even decoding was a strong predictor of reading comprehension and that our sample was relatively small.

Regarding the representativeness of the current sample, as noted in Wass et al. (2019), the participants of this study were all recruited from Karolinska University Hospital in Sweden which has a catchment area of around 5 million people and the majority of children implanted in Sweden receive their implants and are followed up regularly by the CI-team at this hospital. The inclusion criterion was that the children should follow the national school curriculum. We thus consider the sample to be relatively representative of children with CI in Sweden who have no additional disabilities that prevent them from attending general education.

Six of the children were reported to use or have used some combination of oral language and sign as support or sign language. This language exposure may thus potentially have had a negative effect on their development of vocabulary and reading comprehension (c.f. Fitzpatrick et al., 2016; Geers et al., 2017). However, as most of the children in the current sample mainly used oral communication, we believe that early exposure to sign as support or sign language are unlikely to affect the current results at a group level.

In summary, it seems that receptive vocabulary is a strong predictor of reading comprehension in 11–12-year-old children with CI. These results support and extend the findings from other studies (Verhoeven and Van Leeuwe, 2008; Ouellette and Beers, 2010; Olson et al., 2011; Melby-Lervåg and Lervåg, 2014; Perfetti and Stafura, 2014) by suggesting that vocabulary is a main predictor of reading comprehension also in children with CI at this age.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of Etikprövningslagen, the Research Ethics Review Committee at Linköping University with written

informed consent from all subjects in accordance with the Declaration of Helsinki. The protocol was approved by the Research Ethics Review Committee at Linköping University (Dnr 2011/295-31).

AUTHOR CONTRIBUTIONS

FUNDING

REFERENCES

Asker-Árnason, L., Wass, M., Ibertsson, I., Lyxell, B., and Sahlén, B. (2007). The relationship between reading comprehension, working memory and language in children with cochlear implants. *Acta Neuropsychol.* 5, 163–187.

Bell, N., Angwin, A. J., Wilson, W. J., and Arnott, W. L. (2019). Reading development in children with cochlear implants who communicate via spoken language: a psycholinguistic investigation. *J. Speech Lang. Hear. Res.* 62, 456–469. doi: 10.1044/2018_JSLHR-H-17-0469

Braze, D., Tabor, W., Shankweiler, D. P., and Mencel, W. E. (2007). Speaking up for vocabulary: reading skill differences in young adults. *J. Learn. Disabil.* 40, 226–243. doi: 10.1177/00222194070400030401

Byrne, B., Coventry, W. L., Olson, R. K., Samuelsson, S., Corley, R., Willcutt, E. G., et al. (2009). Genetic and environmental influences on aspects of literacy and language in early childhood: continuity and change from preschool to Grade 2. *J. Neurolinguistics* 22, 219–236. doi: 10.1016/j.jneuroling.2008.09.003

Coltheart, M., Rastle, K., Perry, C., Langdon, R., and Ziegler, J. (2001). A dual route cascaded model of word recognition and reading aloud. *Psychol. Rev.* 108, 204–256. doi: 10.1037//0033-295x.108.1.204

Connor, C. M., and Zwolan, T. A. (2004). Examining multiple sources of influence on the reading comprehension skills of children who use cochlear implants. *J. Speech Lang. Hear. Res.* 47, 509–527.

Coppens, K. M., Tellings, A., Verhoeven, L., and Schreuder, R. (2013). Reading vocabulary in children with and without hearing loss: the roles of task and word type. *J. Speech Lang. Hear. Res.* 56, 654–666. doi: 10.1044/1092-4388(2012/11-0138)

Currie, N. K., and Cain, K. (2015). Children's inference generation: the role of vocabulary and working memory. *J. Exp. Child Psychol.* 137, 57–75. doi: 10.1016/j.jecp.2015.03.005

Dillon, C. M., De Jong, K., and Pisoni, D. B. (2011). Phonological awareness, reading skills and knowledge in children who use cochlear implants. *J. Deaf Stud. Deaf Educ.* 17, 205–226. doi: 10.1093/deafed/enr043

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N., et al. (1999). Maternal education and measures of early speech and language. *J. Speech Lang. Hear. Res.* 42, 1432–1443. doi: 10.1044/jslhr.4206.1432

Dunn, L. M., and Dunn, D. M. (2007). *Peabody Picture Vocabulary Test*. 2nd Edn. Minneapolis, MN: NCS Pearson INC.

Ehri, L. C. (2005). Learning to read words: theory, findings, and issues. *Sci. Stud. Read.* 9, 167–188. doi: 10.1207/s1532799xssr0902_4

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Sci. Stud. Read.* 18, 5–21. doi: 10.1080/10888438.2013.819356

Elwér, Å, Fridolfsson, I., Samuelsson, S., and Wiklund, C. (2009). *LäSt. Hogrefe Psykologiförlaget AB*. Stockholm: Hogrefe Psykologiförlaget AB.

Fagan, M. K., and Pisoni, D. B. (2010). Hearing experience and receptive vocabulary development in deaf children with cochlear implants. *J. Deaf Stud. Deaf Educ.* 15, 149–161. doi: 10.1093/deafed/enq001

Fitzpatrick, E. M., Hamel, C., Stevens, A., Pratt, M., Moher, D., Doucet, S. P., et al. (2016). Sign language and spoken language for children with hearing loss: a systematic review. *Pediatrics* 37, 1–19.

Geers, A. E. (2003). Predictors of reading skill development in children with early cochlear implantation. *Ear Hear.* 24, 59S–68S. doi: 10.1097/01.aud.0000051690.43989.5d

Geers, A. E., and Hayes, H. (2011). Reading, writing, and phonological processing skills of adolescents with 10 or more years of cochlear implant experience. *Ear Hear.* 32, 49S–59S. doi: 10.1097/AUD.0b013e3181fa41fa

Geers, A. E., Mitchell, C. M., Warner-Czyz, A., Wang, N.-Y., and Eisenberg, L. S. (2017). Early sign language exposure and cochlear implantation benefits. *Pediatrics* 140:e20163489. doi: 10.1542/peds.2016-3489

Geers, A. E., Moog, J. S., Biedenstein, J., Brenner, C., and Hayes, H. (2009). Spoken language scores of children using cochlear implants compared to hearing age-mates at school entry. *J. Deaf Stud. Deaf Educ.* 14, 371–385. doi: 10.1093/deafed/enn046

Gough, P. B., and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial Spec. Educ.* 7, 6–10. doi: 10.1177/074193258600700104

Harris, M., and Terlektsi, E. (2010). Reading and spelling abilities of deaf adolescents with cochlear implants and hearing aids. *J. Deaf Stud. Deaf Educ.* 16, 24–34. doi: 10.1093/deafed/enq031

Kral, A., and Sharma, A. (2012). Developmental neuroplasticity after cochlear implantation. *Trends Neurosci.* 35, 111–122. doi: 10.1016/j.tins.2011.09.004

Kyle, F. E., and Harris, M. (2006). Concurrent correlates and predictors of reading and spelling achievement in deaf and hearing school children. *J. Deaf Stud. Deaf Educ.* 11, 273–288. doi: 10.1093/deafed/enj037

Lervåg, A., and Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *J. Child Psychol. Psychiatry* 51, 612–620. doi: 10.1111/j.1469-7610.2009.02185.x

Lieu, J. E. C., Tye-Murray, N., Karzon, R. K., and Piccirillo, J. F. (2010). Unilateral hearing loss is associated with worse speech-language scores in children. *Pediatrics* 125, e1348–e1355. doi: 10.1542/peds.2009-2448

Magnusson, E., and Nauclér, K. (1993). *Bedömning av Språklig Medvetenhet Hos Förskolebarn Och Skolbarn*. Löddeköpinge: Pedagogisk Design.

Magnusson, K. (2007). Maternal education and children's academic achievement during middle childhood. *Dev. Psychol.* 43, 1497–1512. doi: 10.1037/0012-1649.43.6.1497

Markwardt, F. C. (1998). *Peabody Individual Achievement Test – Revised*. Circle Pines, MN: American Guidance Service.

Melby-Lervåg, M., and Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: a meta-analysis of studies comparing first- and second-language learners. *Psychol. Bull.* 140, 409–433. doi: 10.1037/a0033890

Nakeva von Mentzer, C., Lyxell, B., Sahlén, B., Dahlström, Ö, Lindgren, M., Ors, M., et al. (2014). Computer-assisted reading intervention with a phonics approach for children using cochlear implants or hearing aids. *Scand. J. Psychol.* 55, 448–455. doi: 10.1111/sjop.12149

Olson, R. K., Keenan, J. M., Byrne, B., Samuelsson, S., Coventry, W. L., Corley, R., et al. (2011). Genetic and environmental influences on vocabulary and reading development. *Sci. Stud. Read.* 15, 26–46. doi: 10.1080/10888438.2011.536128

Ouellette, G., and Beers, A. (2010). A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Read. Writ.* 23, 189–208. doi: 10.1007/s11145-008-9159-1

Perfetti, C., and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Sci. Stud. Read.* 18, 22–37. doi: 10.1080/10888438.2013.827687

Perfetti, C. A. (2007). Reading ability: lexical quality to comprehension. *Sci. Stud. Read.* 11, 357–383. doi: 10.1080/10888430701530730

Protopapas, A., Sideridis, G. D., Simos, P. G., and Mouzaki, A. (2007). Development of lexical mediation in the relation between reading

comprehension and word reading skills in Greek. *Sci. Stud. Read.* 11, 165–197. doi: 10.1080/10888430701344322

Raven, J., Raven, J. C., and Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Pearson Assessment.

Stiles, D. J., McGregor, K. K., and Bentler, R. A. (2012). Vocabulary and working memory in children fit with hearing aids. *J. Speech Lang. Hear. Res.* 55, 154–167. doi: 10.1044/1092-4388(2011/11-0021)

Tunmer, W. E., and Chapman, J. W. (2012). The simple view of reading redux: vocabulary knowledge and the independent components hypothesis. *J. Learn. Disabil.* 45, 453–466. doi: 10.1177/0022219411432685

Tunmer, W. E., and Greaney, K. T. (2010). Defining dyslexia. *J. Learn. Disabil.* 43, 229–243. doi: 10.1177/0022219409345009

Verhoeven, L., and Van Leeuwe, J. (2008). Prediction of the development of reading comprehension: a longitudinal study. *Appl. Cogn. Psychol.* 22, 407–423. doi: 10.1002/acp.1414

Vermeulen, A. M., van Bon, W., Schreuder, R., Knoors, H., and Snik, A. (2007). Reading comprehension of deaf children with cochlear implants. *J. Deaf Stud. Deaf Educ.* 12, 283–302. doi: 10.1093/deafed/enm017

Von Muenster, K., and Baker, E. (2014). Oral communicating children using a cochlear implant: good reading outcomes are linked to better language and phonologicall processing abilities. *Int. J. Paediatr. Otorhinolaryngol.* 78, 433–444. doi: 10.1016/j.ijporl.2013.12.009

Walker, E. A., Redfern, A., and Oleson, J. J. (2019). Linear mixed-model analysis to examine longitudinal trajectories in vocabulary depth and breadth in children who arew hard of hearing. *J. Speech Lang. Hear. Res.* 25, 525–542. doi: 10.1044/2018_JSLHR-L-ASTM-18-0250

Wass, M., Ibertsson, T., Lyxell, B., Sahlén, B., Hällgren, M., Larsby, B., et al. (2008). Cognitive and linguistic skills in Swedish children with cochlear implants – measures of accuracy and latency as indicators of development. *Scand. J. Psychol.* 49, 559–576. doi: 10.1111/j.1467-9450.2008.00680.x

Wass, M., Löfkvist, U., Anmyr, L., Karltorp, E., Östlund, E., and Lyxell, B. (2019). Correlates of orthographic learning in swedish children with cochlear implants. *Front. Psychol.* 10:143. doi: 10.3389/fpsyg.2019.00143

Yoshinaga-Itano, C., Baca, R. L., and Sedey, A. L. (2010). Describing the trajectory of language development in the presence of severe to profound hearing loss: a closer look at children with cochlear implants versus hearing aids. *Otol. Neurotol.* 31, 1268–1274. doi: 10.1097/MAO.0b013e3181f1ce07

frontiers
in Psychology

Noise, Age, and Gender Effects on Speech Intelligibility and Sentence Comprehension for 11- to 13-Year-Old Children in Real Classrooms

Nicola Prodi[1]*, Chiara Visentin[1], Erika Borella[2], Irene C. Mammarella[3] and Alberto Di Domenico[4]

[1] Department of Engineering, University of Ferrara, Ferrara, Italy, [2] Department of General Psychology, University of Padova, Padua, Italy, [3] Department of Developmental and Social Psychology, University of Padova, Padua, Italy, [4] Department of Psychological, Health and Territorial Sciences, University of Chieti, Chieti, Italy

The present study aimed to investigate the effects of type of noise, age, and gender on children's speech intelligibility (SI) and sentence comprehension (SC). The experiment was conducted with 171 children between 11 and 13 years old in ecologically-valid conditions (collective presentation in real, reverberating classrooms). Two standardized tests were used to assess SI and SC. The two tasks were presented in three listening conditions: quiet; traffic noise; and classroom noise (non-intelligible noise with the same spectrum and temporal envelope of speech, plus typical classroom sound events). Both task performance accuracy and listening effort were considered in the analyses, the latter tracked by recording the response time (RT) using a single-task paradigm. Classroom noise was found to have the worst effect on both tasks (worsening task performance accuracy and slowing RTs), due to its spectro-temporal characteristics. A developmental effect was seen in the range of ages (11–13 years), which depended on the task and listening condition. Gender effects were also seen in both tasks, girls being more accurate and quicker to respond in most listening conditions. A significant interaction emerged between type of noise, age and task, indicating that classroom noise had a greater impact on RTs for SI than for SC. Overall, these results indicate that, for 11- to 13-year-old children, performance in SI and SC tasks is influenced by aspects relating to both the sound environment and the listener (age, gender). The presence of significant interactions between these factors and the type of task suggests that the acoustic conditions that guarantee optimal SI might not be equally adequate for SC. Our findings have implications for the development of standard requirements for the acoustic design of classrooms.

Keywords: classroom acoustics, intelligibility, sentence comprehension, listening effort, noise, children, gender, response times

INTRODUCTION

Oral communication in classrooms is a complex phenomenon involving different types of speech material (from simple commands to complex lectures) and speaker-listener interactions (e.g., teacher to class, one-to-one during group work, one to small group, etc.). While these two factors may combine in various ways, giving rise to different communication scenarios, all of the currently-used standards for classroom acoustics are only conceived to guarantee speech intelligibility (SI). The standards provide for limits in terms of acoustic indicators, which are designed to account for the separate and/or joint effects of background noise and reverberation on speech reception (e.g., the Speech Transmission Index of the International Electrotechnical Commission, 2011). Unfortunately, SI is on the surface of the levels of representation involved in verbal processing (Hustad, 2008), and it mainly provides information about the correct reception of the acoustic-phonetic cues in a message. Differently, communication during lessons requires a higher level of language processing. It relies on messages with variable syntactic forms, and on lexical, semantic and contextual information, and listeners are expected not only to understand the content, but also to integrate it with previously acquired experience and knowledge.

The testing of listening comprehension in adult and pediatric populations has been the object of several publications. Specific tests have been developed, based on listening to text passages and answering content questions (Valente et al., 2012; Sullivan et al., 2015; Rudner et al., 2018; von Lochow et al., 2018), or on implementing oral instructions (Klatte et al., 2010a). The tasks presented in such studies are similar to tasks that students perform in their everyday life, and are consequently ecologically valid, but their inherent complexity can make them difficult to administer routinely for the assessment of classroom acoustics.

To improve on assessments based on SI alone, a viable alternative to listening comprehension is to consider sentence comprehension (SC). This approach provides information on levels of language processing beyond speech reception because auditory, syntactic, contextual, and semantic information can be manipulated in a simple and scalable manner. For instance, Uslar et al. (2013) described how linguistic complexity could be modulated to improve the audiological matrix sentence test for adults (Wagener et al., 1999), and gain information on the usage of their cognitive capacity while listening in noise. It is generally assumed that the more the extraction of meaning from the speech signal is elaborate, the greater the burden on the top–down cognitive resources of the listener (Downs and Crum, 1978), leaving less cognitive capacity left over for higher-level speech processing (Rudner and Lunner, 2014). Increasing the linguistic difficulty of sentences, or chaining the sentences together would thus help to clarify the speech processing needs in classrooms, adding to the information provided by the basic SI results. Comparisons between the two tasks (SI and SC) have not been conducted systematically, whereas some results are available for comparisons between SI and certain more complex listening comprehension tasks. For instance, Fontan et al. (2015) tested young adults and, using a task that involved commands to move objects, they retrieved transcripts of instructions for SI and also monitored subsequent actions. When the authors compared the scores for SI and comprehension, they found a modest correlation between the two tasks ($r = 0.35$), and concluded that SI was a poor predictor of comprehension in real communication settings. Klatte et al. (2010a) compared SI (word-to-picture matching) and comprehension (execution of oral instructions) in 7- and 9-year-old children, using classroom noise (typical classroom sounds without speech) and background speech as maskers. They found that classroom noise had a stronger effect on SI, but background speech was more harmful for comprehension.

Overall, the literature points to a weak relationship between task performance accuracy in SI and comprehension tasks for normally-hearing listeners. Fontan et al. (2017) points out that intelligibility and comprehension measures might be considered as complementary, providing information on different aspects of speech communication. Exploring the effects of noise and reverberation on both tasks could therefore facilitate the development of effective tools for controlling the sound environment in the classroom, considering at once speech signal transmission and communicative performance.

Several explanations have been advanced for the specific impact of noise and reverberation on verbal task outcomes in classrooms. In particular, the way noise interferes with speech depends not only on the level of noise, but also on its spectro-temporal characteristics. The adverse effect of a background noise may originate from either energetic or informational masking (Mattys et al., 2012). In the former case, speech and masker overlap in time and frequency in such a way that portions of the signal are no longer audible (Brungart, 2001). This form of masking is supposed to take place at the level of the auditory periphery and the recognition process relies mainly on stream segregation and selective attention. Adult listeners experience an advantage in speech reception for temporally fluctuating maskers compared with steady-state maskers presented at the same noise level. This so-called "masking release" originates from a combination of factors (see Füllgrabe et al., 2006 for a complete review), including dip listening, or the listener's ability to exploit short periods with high signal-to-noise ratios (SNR), when the fluctuating noise was lowest, to detect speech cues. The fluctuations in the background noise may also interfere with the temporal fluctuations in the speech, giving raise to the modulation masking, which counterbalances dip listening. Informational masking is believed to have consequences on speech recognition that go beyond its energetic effect, such as attentional capture, semantic interference, and increased cognitive load. Background speech with intelligible and meaningful content may result in informational masking, as its interference directly affects working memory by competing with the target speech. Non-speech sounds may produce informational masking as well. As Klatte et al. (2010b) pointed out, however, the various effects of non-speech sound cannot be explained by a single mechanism. Depending on its characteristics, a sound may have a changing state effect (e.g., when the sound consists of distinct auditory objects that vary consecutively; see Hughes and Jones, 2001), or

an attentional capture effect (e.g., salient, unexpected, or deviant auditory events; see Klatte et al., 2013), or a mixture of both.

With specific reference to the effect of background noise on children in classrooms, Klatte et al. (2007) found higher-level cognitive processing more affected by unintelligible background speech than by traffic noise, when the two noises were presented at the same level; the authors related the difference to the changing-state characteristics of the background speech. Dockrell and Shield (2006) compared quiet, babble, and babble plus environmental noise conditions, testing 7- to 8-year-old children with verbal tasks (reading and spelling). They found the children's performance accuracy negatively affected by classroom babble, and suggested that verbal tasks involving working memory processes are more vulnerable to the interference of concurrent speech.

Like background noise, reverberation in the classroom can also increase the speech processing burden. Normative values have been established for optimal reverberation times, which depend on the classroom's volume and the use made of the space (Deutsche Institut für Normung, 2016). Several studies have demonstrated the importance of assessing the combined effects of noise and reverberation in classrooms, given the greater effect of adverse listening conditions on children than on adults. Prior research indicated that speech recognition in noisy and reverberating conditions improves with age (Neuman et al., 2010) and consonant identification does not reach adult-like performance accuracy until the age of 14 years (Johnson, 2000). Children are also more easily distracted by auditory events due to their less robust and less developed attentional abilities (Klatte et al., 2013; Meinhardt-Injac et al., 2014), and their performance accuracy deteriorates the most in speech-in-speech tasks (with competing speech from two talkers, see Corbin et al., 2016). Masking release is also more limited in children (up to 13–14 years old) than in adults, when a speech-shaped, amplitude-modulated noise is presented in reverberating conditions (Wróblewski et al., 2012). Leibold (2017) suggested that this latter finding might indicate that children are not as good as adults at glimpsing speech in fluctuating noise.

Most of the available data about children's speech processing in the classroom are based on their accuracy in completing tasks, while few studies have also considered their response times (RTs) measured using a single-task paradigm in order to judge their listening effort. In this context, RT is intended as a measure of speed of processing, and provides information on the amount of cognitive capacity allocated to processing the auditory signal (Pichora-Fuller et al., 2016). Several published studies indicate that, like other measures of listening effort, changes in RT may mirror changes in task performance accuracy (e.g., Lewis et al., 2016; McGarrigle et al., 2019), but they may also occur when accuracy is at or near ceiling level (Hällgren et al., 2001), or kept constant (Uslar et al., 2013; Sahlén et al., 2017). On the whole, the literature supports the hypothesis that accuracy and listening effort might represent two different constructs in the general frame of speech processing: the two measures are not always related (Wendt et al., 2018), and factors affecting task performance accuracy do not affect listening effort to the same degree (Picou et al., 2016). Measures of listening effort are

generally considered valuable to complement traditional speech-in-noise tests, and provide additional information beyond task performance accuracy.

With specific reference to the use of RTs in the pediatric population, Lewis et al. (2016) used verbal RTs as a proxy for listening effort in a study on normally-hearing children from 5 to 12 years old, and children with hearing loss. The children with a normal hearing function had longer RTs with decreasing SNR. These results were confirmed by McGarrigle et al. (2019), who also found that verbal RTs were more effective than visual, dual-task RTs for children 6 to 13 years old. Prodi et al. (2013) combined SI with RTs for 8- to 11-year-olds. This method enabled a ranking of the interference of different types of noise, and revealed changes in the balance between signal-driven and knowledge-driven processes. SI improved and RTs decreased with increasing age, but the changes in the two metrics followed different patterns. The increase in task performance accuracy with older age came first, and it was only after accuracy reached the ceiling that a decrease in RTs with increasing age became apparent.

The general mechanisms governing the effects of noise and reverberation on speech reception are sufficiently well-known and documented for primary school children, but there is a need to extend what we know to less well-researched age ranges, such as 11- to 13-year-olds. The ability to hear and understand speech in adverse conditions matures during childhood, but the age at which an adult-like performance is reached depends on the nature of the background noise (Leibold, 2017). In complex acoustic environments, with non-stationary noises and reverberation, 13- to 14-year-olds perform less well than adults (Wróblewski et al., 2012): this gives the impression that children up to this age might continue to be at a particular disadvantage when listening in adverse conditions. In addition, the comparison between performance accuracy results in SI and SC has been pursued for adults (Hustad, 2008; Fontan et al., 2015), and for children aged 7 and 9 years (Klatte et al., 2010a), but no investigations have been conducted on older school-age children. A better understanding of how noise, age and task may interact would be valuable when tailoring classroom acoustics to optimize learning performance and reduce listening effort.

Previous studies on developmental changes in speech processing ability in the classroom have also considered the issue of gender differences. Ross et al. (2015) tested a group of typically-developing children from 5 to 17 years old over a fairly wide range of SNRs using a speech recognition task with isolated, monosyllabic words. They found that females performed better than their male peers in both audio-only and audio-visual presentation modes. When Boman (2004) investigated the interaction between gender and noise in 13-to 14-year-olds using episodic and semantic memory tasks, girls had a better recall performance than boys, and this finding was consistent across different verbal materials. No interaction emerged between gender and noise as the presence of noise affected the boys' and girls' performance to the same degree. Listening effort has only been considered in relation to gender in the case of voice quality deterioration, and for 8-year-olds

(Sahlén et al., 2017). In the study by Sahlén et al. (2017), a SC test was administered in multi-talker babble noise and the RTs for listening conditions in which girls and boys performed equally well were considered (Lyberg-Åhlander et al., 2015). Unlike task performance accuracy, latencies were longer for girls than for boys. Considering these results together, it is unclear whether the girls' better performance accuracy – reported by Boman (2004) and Ross et al. (2015) – coincided with slower processing times, or whether the findings of Sahlén et al. (2017) concerning listening effort related to the particular testing conditions (dysphonic voice) or to differences in the strategies used by girls and boys to solve the task.

The present work reports on SI and SC tasks presented in real reverberating classrooms. The participants consisted of a fairly large group of children 11 to 13 years old, who collectively performed the tasks in three listening conditions: quiet; traffic noise; and classroom noise (speech-like noise plus typical classroom sounds). Both tasks were presented in a closed-set format, using personal portable devices (tablets). Two outcome measures were considered (task performance accuracy and RTs), and used to obtain a comprehensive view of the speech processing phenomenon. RTs were used as a behavioral measure to quantify listening effort, assuming that slower RTs reflect a greater listening effort.

The tasks were presented to 11- to 13-year-old children in their classrooms. The research questions addressed were as follows:

(1) Depending on the task and the type of noise, what is the interplay between task performance accuracy and listening effort when children have to cope with noise? Does age have any effect?
(2) Are there gender-related differences in SC and SI task performance? Do these differences regard task accuracy alone, or listening effort as well?
(3) When both SI and SC are evaluated under the same acoustic conditions, does age and type of noise similarly influence performance accuracy and listening effort in the two task?

MATERIALS AND METHODS

Description of the Classrooms

The experiment took place in the first half of the school year (November–December, 2018) at two schools in Ferrara, Italy. One classroom was chosen at each school for use as a laboratory during the test sessions. Both classrooms were box-shaped, with similar volumes (152 and 155 m^3), and dimensions (7.3 m long × 7.0 m wide × 3.0 m high; and 8.3 m × 6.0 m × 3.1 m). During the experiments, the classrooms were set up as for regular lessons, with wooden desks and chairs arranged in rows and facing the teacher's desk.

Only one of the classrooms had sound-absorbing ceiling tiles, so the other classroom was temporarily fitted with sound-absorbing polyester fiber blankets to balance the acoustic conditions in the two rooms. This temporary solution ensured

the same reverberation times across the octave band frequencies in both classrooms: the T_{mid} (average reverberation time for the octave bands 500–2000 Hz) in occupied conditions was 0.68 and 0.69 s respectively. At the time of testing, the number of pupils sitting in the classrooms ranged between 14 and 23, depending on the number of students belonging to each class.

Participants

A total of 171 pupils between 11 and 13 years old belonging to nine different classes at two different schools took part in the study. The school administrations gave their permission for the study. The study was approved by the Ethics Committee of the University of Padova (Italy). Written informed parental consent was obtained prior to any testing.

After the experiment, the teachers provided details about children with intellectual disabilities and hearing impairments (as certified by the National Healthcare System). There were six such children (three at each school), who were excluded from the subsequent data analysis. The results for another six children were also omitted from the analysis due to: the baseline comprehension score in four cases (two children did not complete the assessment, and two scored lower than the threshold); and an extremely low performance in the SI task (quiet condition) in two, indicating that the children misunderstood the instructions.

The final sample of participants is detailed in **Table 1**.

Reading Comprehension Assessment

Before conducting the experiment, pupils were screened for comprehension problems that could influence the study outcomes. Given the association between listening and reading comprehension (Wolf et al., 2019), a measure of reading comprehension was used for this purpose.

Students were collectively presented with the measures in a quiet condition. The assessment took place nearly 1 week after presenting the SI and SC tasks. A standardized reading comprehension test based on the participants' school grade was administered (derived from Cornoldi et al., 2017). Participants were given text passages to read silently. Then they had to answer 15 multiple-choice questions without any time constraints, and could refer back to the passage while answering. Cronbach's alpha was higher than 0.71 for all tasks, indicating an acceptable internal consistency.

For each age group, differences between classes and genders were examined with reference to the reading comprehension assessment. No significant differences emerged between the genders, whereas there were significant differences between the classes (see **Table 2**).

TABLE 1 | Characteristics of the children participating in the study.

Age group	No. of participants	% Male/female	Age [M (SD); range]
11 years	53	49/51	11.0 (0.3); 10–12
12 years	49	53/47	11.9 (0.2); 11–12
13 years	57	58/42	12.9 (0.3); 12–13
All	159	53/47	12.0 (0.9); 10–13

TABLE 2 | Significance tests for the reading comprehension task, by class (three for each age group) and gender: Mann–Whitney's *U*-test on gender, Kruskal–Wallis test for classes.

Age group	Class	Gender
11 years	$\chi^2(2) = 7.15$, $p = 0.03$	$W = 372$, $p = 0.92$
12 years	$\chi^2(2) = 7.42$, $p = 0.02$	$W = 450$, $p = 0.21$
13 years	$\chi^2(2) = 13.42$, $p < 0.001$	$W = 463$, $p = 0.91$

Speech Intelligibility Task – Stimuli, Procedure, and Dependent Variables

Speech intelligibility was assessed with the Matrix Sentence Test in the Italian language (ITAmatrix, see Puglisi et al., 2015). This test is based on five-word sentences, with a fixed syntactic structure but no semantic predictability (e.g., *Sofia compra poche scatole rosse* [Sophie buys few red boxes]). Each sentence is generated from a 10×5 base-word matrix, with 10 options for each word in the sentence.

Digital recordings of the sentences were acquired by agreement with the producer, Hoertech GmbH. The average sentence duration was 2.3 s. Three lists of 16 sentences were created for the experiment, plus four additional sentences for the training phase.

For each trial comprising the task, participants were presented aurally with the playback of a sentence. After the audio offset, the base-word matrix was displayed on the tablets and participants had to select the words they had heard in serial order (i.e., in the same order in which the words were played back). It was impossible to change a response once the selection had been made. Participants were allowed a maximum of 15 s to select the five words.

The score (right/wrong) for each word comprising the sentence was recorded and used to evaluate the SI score, defined as the percentage of words correctly recognized in the sentence. RTs (i.e., the time elapsing between the end of the waveform of the last word in the sentence heard and the selection of the first word on the tablet) was automatically recorded for each participant and trial.

Sentence Comprehension Task – Stimuli, Procedure, and Dependent Variables

Sentence comprehension was examined using the COMPRENDO Test (Cecchetto et al., 2012), which is designed to assess comprehension of a series of sentences in the Italian language. The sentences differ in their syntactic complexity: transitive active sentences (e.g., *La mamma sta inseguendo il bambino* [The mother is chasing the child]), dative sentences (e.g., *Il papà dà il latte alla bambina* [The father gives milk to the little girl]), active sentences with two objects (e.g., *Il bambino insegue il cane e il gatto* [The child chases the dog and the cat]), coordination between active sentences (e.g., *Il bambino guarda il gatto e la mamma accarezza il cane* [The child looks at the cat and the mother strokes the dog]), sentences with subject relative clauses (e.g., *Il bambino che saluta il nonno guarda la televisione* [The child who greets his grandfather is watching television]), and sentences with object relative clauses

(e.g., *Il nonno spinge il cane che morde il gatto* [The grandfather pushes the dog that is biting the cat]). All the sentences (10 for each type) were generated using 20 nouns and 20 verbs that were easy to understand and in very common use. Material selection occurred in two phases. In the first phase, 200 nouns and 200 verbs with higher frequency were selected from the Laudanna et al. (1995) database. In the second phase, a group comprised by one psychologist, one speech-language pathologist, and one neuropsychologist, selected the nouns and verbs to use for the material of the study among the 400 words obtained in phase one.

The sentences were recorded in a silent room by a native Italian, female, adult speaker. A B&K Type 4189 1/2 inch microphone was placed about 20 cm from the speaker's mouth and routed to a B&K Type 5935 signal conditioner. The digital recordings had a 16-bit resolution and a 44100 Hz sampling rate. The sentences were digitally filtered to match the long-term spectrum of the female speaker in the ITAMatrix. The sentences lasted a mean 3.4 s. Three different lists of 16 sentences each were prepared using a pseudo-randomized procedure to ensure that the same number of sentences was presented for each level of syntactic complexity in each list.

During the experimental session, the sentences were aurally presented to participants. After the audio offset of each sentence, four images appeared (one for each quadrant on the screen), and participants were asked to touch the image that properly described the sentence they had just heard (**Figure 1**). RTs and accuracy were recorded for each sentence. A time-out of 12 s was set for selecting an answer.

Background Noises and Listening Conditions

Three listening conditions were considered in the study: quiet, traffic noise, and classroom noise. For the traffic noise, recordings were obtained alongside a busy road in conditions of dense traffic, including cars and trucks. The recordings were spectrally filtered to account for the sound insulation properties of a typical building façade. For the classroom noise, Italian phrases spoken by a native female speaker were processed according to the established ICRA procedure (Dreschler et al., 2001). The resulting signal had speech-like fluctuations and the same spectrum as Italian speech, but was not intelligible. Sound events typical of a busy classroom were added to this signal by digital mixing (e.g., a pen rolling off a desk onto the floor, chairs scraping, pages being turned over in a book).

The long-term averaged spectral characteristics of the two types of background noise are shown in **Figure 2**. The classroom noise had typical speech-like components plus higher frequencies due to sounds common in classrooms being mixed with the babble. The traffic noise had a more balanced frequency trend up to 2 kHz, then sloped down. **Figure 3** shows the temporal pattern of the two types of background noise, recorded in anechoic conditions. The classroom noise had faster fluctuations, showing shallow depths and sparse peaks, whereas the traffic noise had slower fluctuations. The amount of fluctuation over time of the noise levels was also qualified using the difference in the
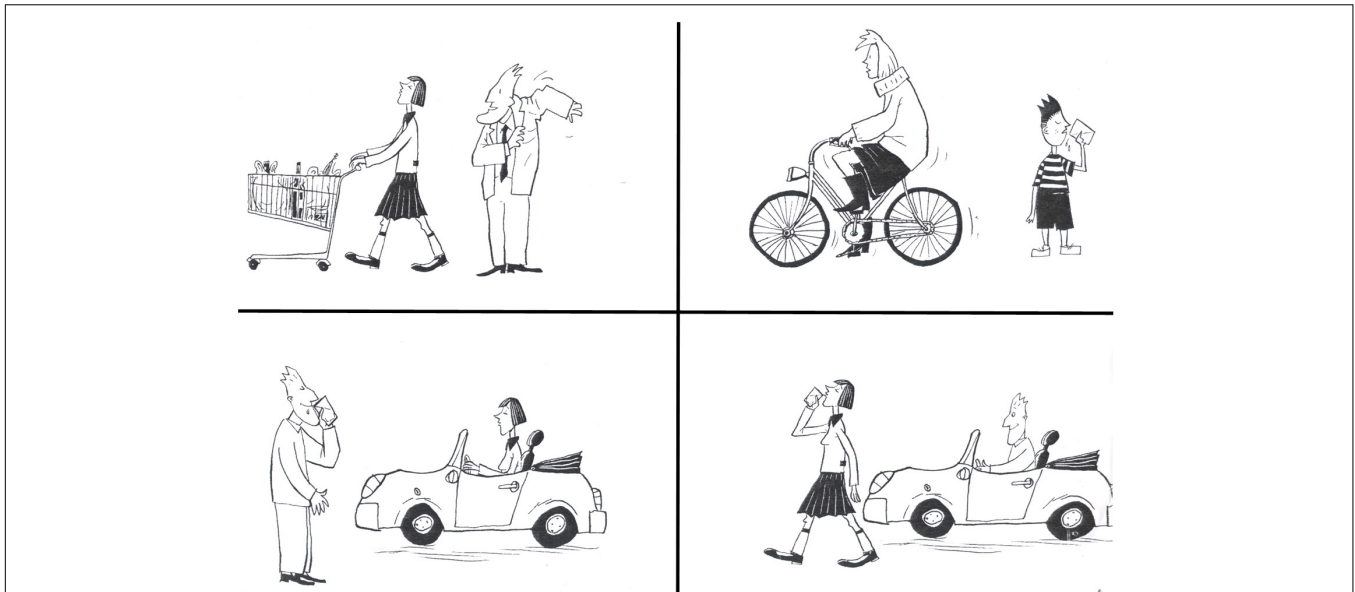
FIGURE 1 | Example of a trial in the sentence comprehension (SC) task, with the four images displayed on the individual tablets. The spoken sentence was "The man drives the car and the woman drinks the milk."

percentile sound levels (i.e., $L_{A,10}$ – $L_{A,90}$). By definition the $L_{A,10}$ value is the level exceeded for 10% of the measurement time, and takes into account the presence of peaks of noise. $L_{A,90}$ is the level exceeded for 90% of the measurement time, and accounts for the residual noise level. The difference between the two percentile sound levels gives an indication of the stationarity of the noise: the difference is low for stationary noise, while it increases for noises with temporal fluctuations. In anechoic conditions the difference was 7.0 and 8.1 dB for the traffic and classroom noise, respectively.
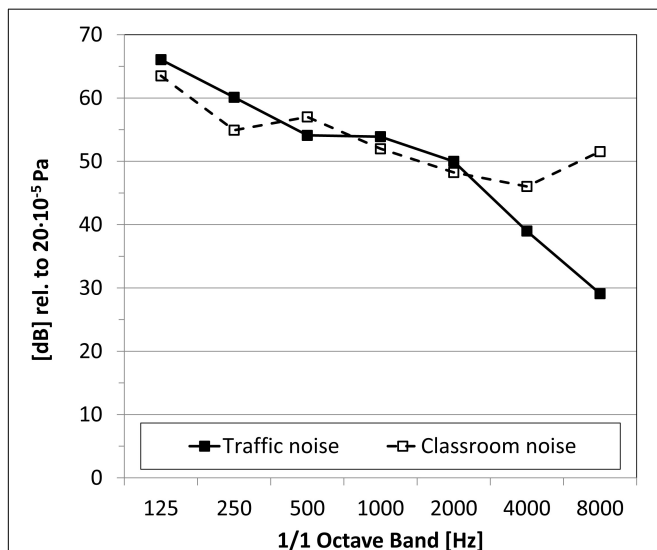
FIGURE 2 | Octave band, long-term average spectra of background noises. The overall A-weighted level is set to 60 dB(A).

For the test sessions, two loudspeakers were placed inside the classroom. A Gras 44AB mouth simulator used to deliver the speech signals was placed close to the teacher's desk, at a height of 1.5 m (assumed as the height of a standing teacher's mouth), and it was oriented toward the audience. The background noises were played back with a Look Line D303 omnidirectional source placed on the floor near the corner of the room closest to the teacher's desk.

In all listening conditions, the speech signal was fixed to a level of 63 dB(A), measured at 1 m in front of the mouth simulator. This corresponds to a speaker talking with a vocal effort qualified as intermediate between "normal" and "raised" (International Organization of Standardization, 2003). This choice of sound pressure level was based on the findings of Bottalico and Astolfi (2012), who measured the average vocal effort of female teachers during the working day, finding a mean sound pressure level of 62.1 dB(A) at 1 m from the speaker's mouth.

In the quiet condition the speech signals were presented against the background ambient noise of the classroom, which consisted of noises coming from adjacent classrooms, where students were engaging in quiet activities. When the tasks were presented in traffic or classroom noise, the playback level was fixed at 60 dB(A), measured as the spatial average over four positions defined in the seating area. This value was chosen to represent a typical level measured in occupied classrooms during lessons, in accordance with the report from Shield et al. (2015), who found that the levels measured during lessons in secondary schools vary between 50 and 70 dB(A).

An objective description of the acoustic conditions experienced by the audience during the test session was obtained with the Speech Transmission Index (STI; International Electrotechnical Commission, 2011). The metric quantifies the loss of modulation of the speech signal during its transmission
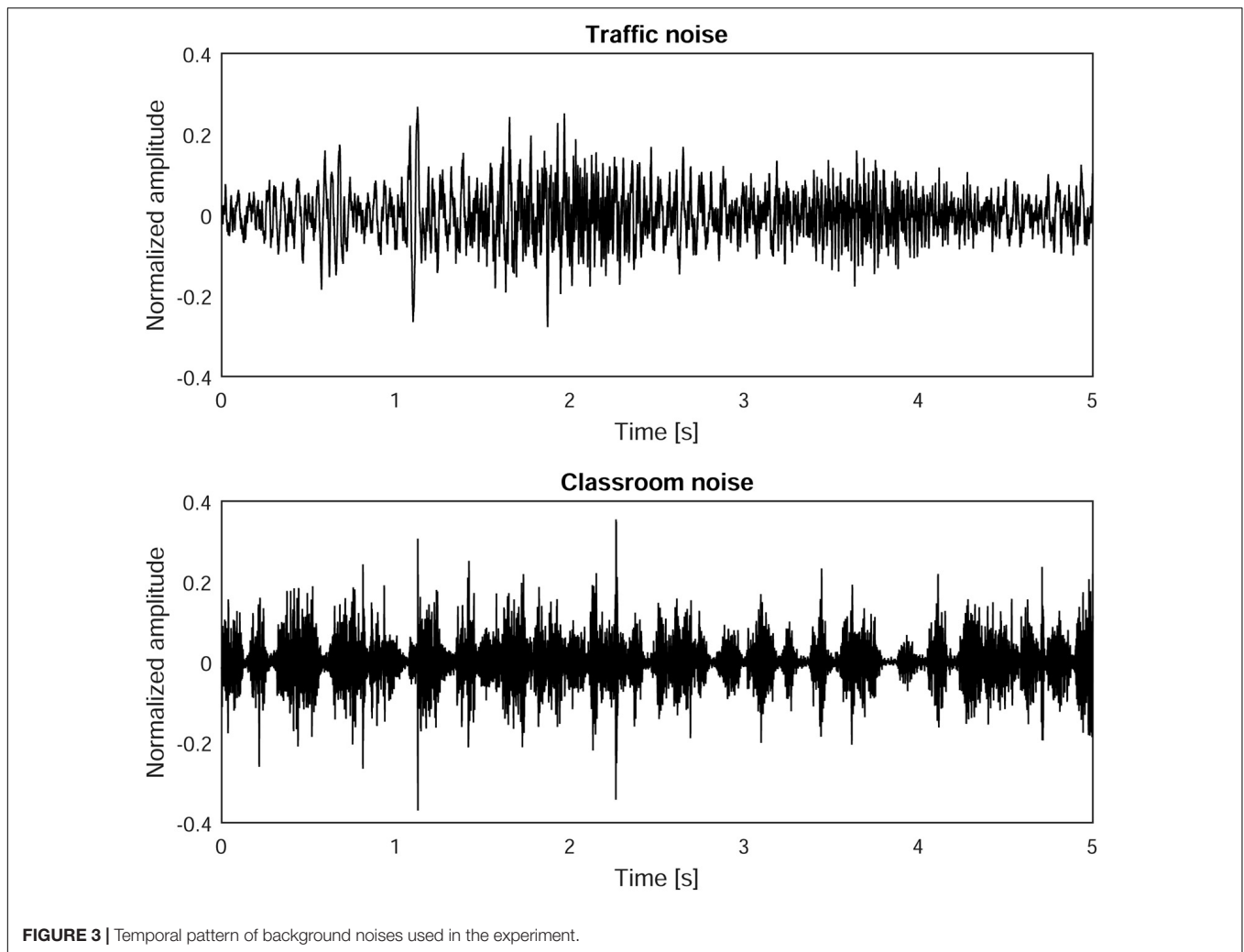
FIGURE 3 | Temporal pattern of background noises used in the experiment.

from the source to the receiver, accounting for the adverse effects of background noise and reverberation. The STI is in the range of [0; 1], the upper limit corresponding to perfect speech transmission.

All measurements were obtained using a B&K type 4189 1/2 inch microphone plus a B&K Type 4231 calibrator, connected to a B&K Type 5935 signal conditioner and a RME Fireface UC full-duplex sound card. The impulse responses and sound pressure levels were measured for each class participating in the study. These measurements were obtained at the end of the experimental session, with the classroom still occupied (see **Figure 4**). Four receiver positions were defined in each classroom, evenly distributed in the area where the students were seated during the experiment, at representative seats. Each microphone was placed at least 1.00 m away from the walls and at a height of 1.20 m (assumed as the height of a student's ears when seated). Care was taken to ensure that the microphone was not shielded by the head or body of the student seated in the row ahead. The students were asked to remain quiet during the measurements.

For each class, the spatial deviation of the acoustic parameters (T_{30}, sound levels, STI) was considered first. The values measured

at the four receivers always differed by a quantity smaller than the corresponding "just noticeable difference" (JND): 5% for the reverberation time, 1 dB for the sound pressure level (International Organization of Standardization, 2009), and 0.03 for the STI (Bradley et al., 1999). This result demonstrates a rather uniform spatial behavior at the seating positions in the classroom, in line with previous studies considering classrooms with sizes comparable to ours (Astolfi et al., 2008, 2012; Prodi et al., 2013). It should be noted that all seating positions were located outside the critical radius (r_c) of the classrooms (i.e., the distance from a sound source at which the level of the direct sound equals the reflected sound level), which was 1.5 m for both classrooms. The seating position closest to the speech source (in the first row of desks, directly facing the source) was 2.10 m from the speech source in one room, and 1.95 m in the other. In the reverberant field, which takes over outside r_c, the sound field is primarily driven by the multiple reflections from the room boundaries. The small dimensions of the classrooms and the presence of a reverberant sound field thus meant that the acoustic parameters had very similar values (no more than 1 JND) in the various seating positions. The spatial uniformity of

FIGURE 4 | The experimental design for the three age groups (11, 12, and 13 years old). Three tests were presented for each task [speech intelligibility (SI) and sentence comprehension], one for each listening condition. The order of the two tasks and the order of the listening conditions were balanced across classes in each age group.
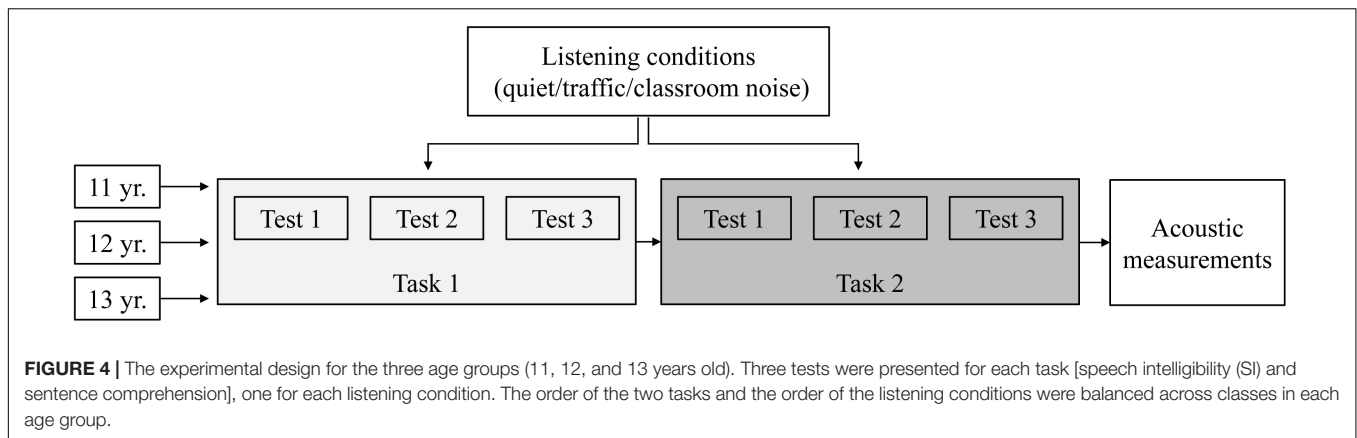
the acoustic parameters in the two rooms is a guarantee that, for these classrooms and seating areas, the listening conditions were equivalent in the different seating positions.

Then the deviation in the acoustic parameters between different groups of students was considered. The differences in the acoustic parameters between repetitions over the classes were always smaller than 1 JND, so the final values for the acoustic parameters in the classrooms were averaged across the repetitions (**Table 3**).

It is worth emphasizing that the differences between the listening conditions in the two classrooms were always smaller than the JND for all the acoustic parameters, except for the sound pressure level in the quiet condition. So, for the purpose of our study, the two rooms can be considered as equivalent from the acoustic perception standpoint (Bradley et al., 1999; Postma and Katz, 2016).

Procedures

Participants completed the experiment in groups consisting of whole classes, which took turns in the laboratory classroom over the course of their morning lessons. The numbers of students in

TABLE 3 | Listening conditions in the two classrooms (A, B) during the experiment: reverberation time T_{mid} (averaged across 500–2000 Hz octave bands), A-weighted sound pressure level $L_{A,eq}$ dB(A), Speech Transmission Index (STI).

Acoustic parameter	Classroom A	Classroom B
Reverberation time T_{mid} [s]	0.68 (0.66; 0.69)	0.69 (0.68; 0.71)
Speech: $L_{A,eq}$ dB(A)	60.4 (59.9; 60.6)	59.5 (59.2; 59.9)
Quiet: $L_{A,eq}$ dB(A)	43.3 (43.2; 43.4)	40.5 (39.7; 40.6)
Traffic noise: $L_{A,eq}$ dB(A)	60.9 (60.7; 61.2)	59.9 (59.4; 60.1)
Classroom noise: $L_{A,eq}$ dB(A)	60.5 (60.3; 60.6)	60.1 (59.6; 60.4)
Quiet: STI	0.64 (0.63; 0.64)	0.65 (0.64; 0.65)
Traffic noise: STI	0.46 (0.44; 0.46)	0.48 (0.47; 0.49)
Classroom noise: STI	0.40 (0.39; 0.40)	0.39 (0.39; 0.40)

All measurements were taken with the rooms occupied. The reported values are spatial averages across four positions in the audience, and across repetitions over the classes. In brackets, maximum and minimum values measured with different groups of children.

each class ranged between 14 and 23. The test session (including the presentation of the task and the acoustic measurements) took 1 h for each class.

At the start of the test session, each child was given a tablet, and was randomly assigned to a seat. Then participants were instructed to enter their age in years and the identification code they found on their desk on their tablets. Using this code ensured that listening positions, test devices and participants were matched correctly, and also ensured anonymity when handling the results. Each child was asked to remember their code and write it on the booklet used for their reading comprehension assessment, which took place on the following days. The same teacher supervised both sessions and ensured the correct matching between participants and codes.

Before starting the experiment, participants were briefly informed about the aim of the study. Then the two tasks were performed, one after the other. To avoid order and fatigue effects, the order of the two tasks was balanced across the classes in each age group. Before each task, participants were given verbal instructions and familiarized with the task and the data collection system by presenting a set of four trials in quiet conditions. Then they completed three tests (one for each listening condition). The listening conditions were balanced across the classes in each age group. The test lists were pseudo-randomized to avoid coupling the same test list with the same listening condition. An outline of the experimental design is shown in **Figure 4**.

During the tests the background noises (traffic or classrooms noise) started approximately 1 s before the target sentence and ended simultaneously with the speech signal. In the quiet condition, an acoustic signal (brief pure tone at 500 Hz) was played back 1 s before the spoken sentence. Each experimental trial was time-limited (to 12 or 15 s, depending on the task). It was only once all participants had responded or reached the time-out that the next target sentence was automatically played back.

Participants were instructed to pay attention to the task, and to respond as accurately as possible. They were not told that RT data would be acquired, nor were they urged to respond as quickly as possible.

The whole experiment was managed by using a wireless test bench (Prodi et al., 2013), based on a server application which

simultaneously controlled the audio playback, the presentation of the base-matrix/images on the tablets, and the data collection.

Data Analysis

Two outcome variables were considered for each task: task performance accuracy and RT.

Before any analysis, data points where technical errors occurred (e.g., loss of the connection between the server and a tablet) were removed from the databases: altogether, 1.2% of the SI trials and 0.7% of the SC trials were discarded for such reasons. Data points corresponding to trials for which the time-out was reached were also removed: this applied to 5.9% of the trials in the SI task and 0.7% of the trials in the SC task.

The statistical analysis was performed using generalized linear mixed-effects models (GLMMs). This statistical method was chosen because it can be used to deal with non-independent individual responses (repeated-measures design) and data for which the normality assumption is not met (Lo and Andrews, 2015; Gordon, 2019). A binomial distribution was adopted in the statistical model for accuracy data, which are bound within the [0; 1] interval, while a Gamma distribution with a log link function was used for the raw RT data.

To analyze each outcome variable in each task, four separate GLMMs were set up (2 tasks $\times$ 2 outcome variables). The fixed effects considered in the models were: listening condition (quiet, traffic, classroom noise); age (11, 12, 13 years); gender (male, female); and all two- and three-way interactions. Because the participants differed significantly in their baseline scores (see **Table 2**), the score in the reading comprehension test was included in the models as a covariate. In all the models, the participant variable was included as a random intercept. The listening condition within-subject factor was also included in the random effects as a random slope. The GLMM thus allowed for the listening condition to have a different effect for each participant.

Then, a second analysis was run to compare the tasks directly in the different listening conditions. This was done by setting up a linear mixed-effects model (LMM), with the relative change in RTs as the outcome variable. The quantity was defined by the ratio of the median RT in noise to the median RT in quiet for each task. The distribution of the raw RTs across the trials was skewed, so the median of the 16 trials was calculated for each combination of participant, listening condition and task, and this was used to calculate the ratio. The resulting quantity reflects the amount of change in processing time due to the addition of background noise. The quiet condition took a value of one for all participant-task combinations, while higher values indicated longer RTs compared with the quiet condition. The fixed effects considered in the LMM were: listening condition (traffic and classroom noise; as quiet was assigned a value of one by definition, it was not included in the model); age (11, 12, 13 years); gender (male, female); task (speech intelligibility, sentence comprehension); the two-way interactions including task and listening condition, and the three-way interaction between age, listening condition and task. The score in the reading comprehension task was added to the models as a

covariate. A random intercept (participant) and two random slopes (the within-participant variables listening condition and task) were also specified.

Values for the GLMMs and LMM were obtained using likelihood ratio tests. The consistency of the models was investigated by checking their assumptions, which meant controlling the normality of the random effect terms and the residuals, as suggested by Everitt and Hothorn (2010).

The analysis was conducted using the *R* software (R Core Team, 2017) and the *lme4* package (Bates et al., 2015). *Post hoc* pairwise comparisons were performed using least-squares means tests with the *lsmeans* package (Lenth, 2016). In the case of multiple comparisons, the Bonferroni method was applied to adjust the *p*-values. The statistical significance threshold was set at 0.05.

RESULTS

Speech Intelligibility: Accuracy

Figure 5 shows the SI scores by age and listening condition, for boys and girls. The analysis revealed a statistically significant main effect of listening condition [$\chi^2(2) = 189.23$, $p < 0.001$]. *Post hoc* tests comparing listening conditions collapsed across age and gender revealed that task performance accuracy was significantly better in quiet than in noisy conditions (quiet > traffic noise, $z = 4.11$, $p < 0.001$; quiet > classroom noise, $z = 11.82$, $p < 0.001$), and in classroom noise than in traffic noise (traffic noise > classroom noise, $z = 10.25$, $p < 0.001$). The SI scores were 1.6% higher in quiet than in traffic noise, and 5.5% higher in traffic noise than in classroom noise.

The analysis also revealed a significant main effect of age [$\chi^2(2) = 56.42$, $p < 0.001$]. *Post hoc* tests with the results collapsed across listening condition and gender showed a worse performance accuracy for the youngest children than for the others (11 < 12 years, $z = -5.66$, $p < 0.001$; 11 < 13 years, $z = -6.88$, $p < 0.001$). The mean results were 85.7% ($SD = 11.7\%$), 91.8% ($SD = 7.3\%$) and 94.1% ($SD = 6.0\%$) for 11-, 12-, and 13-year-olds, respectively.

Finally, the analysis showed a significant main effect of gender [$\chi^2(1) = 56.42$, $p < 0.001$], with girls performing significantly better ($M = 91.8\%$, $SD = 8.3\%$) than boys ($M = 89.6\%$, $SD = 10.1\%$). The main effect of the reading comprehension score [$\chi^2(2) = 20.72$, $p < 0.001$] was significant as well.

There were no interactions between listening condition and age ($p = 0.84$), between listening condition and gender ($p = 0.59$), or between age and gender ($p = 0.84$). There was also no significant three-way interaction between listening condition, age and gender ($p = 0.12$).

Speech Intelligibility: RTs

Figure 6 shows the RTs (median across the trials) for each listening condition and age, for boys and girls. The analysis revealed a significant main effect of listening condition [$\chi^2(2) = 25.41$, $p < 0.001$], a main effect of age [$\chi^2(2) = 6.61$, $p < 0.001$], and a main effect of gender [$\chi^2(1) = 8.66$, $p = 0.003$]. The two-way interactions between listening condition and age
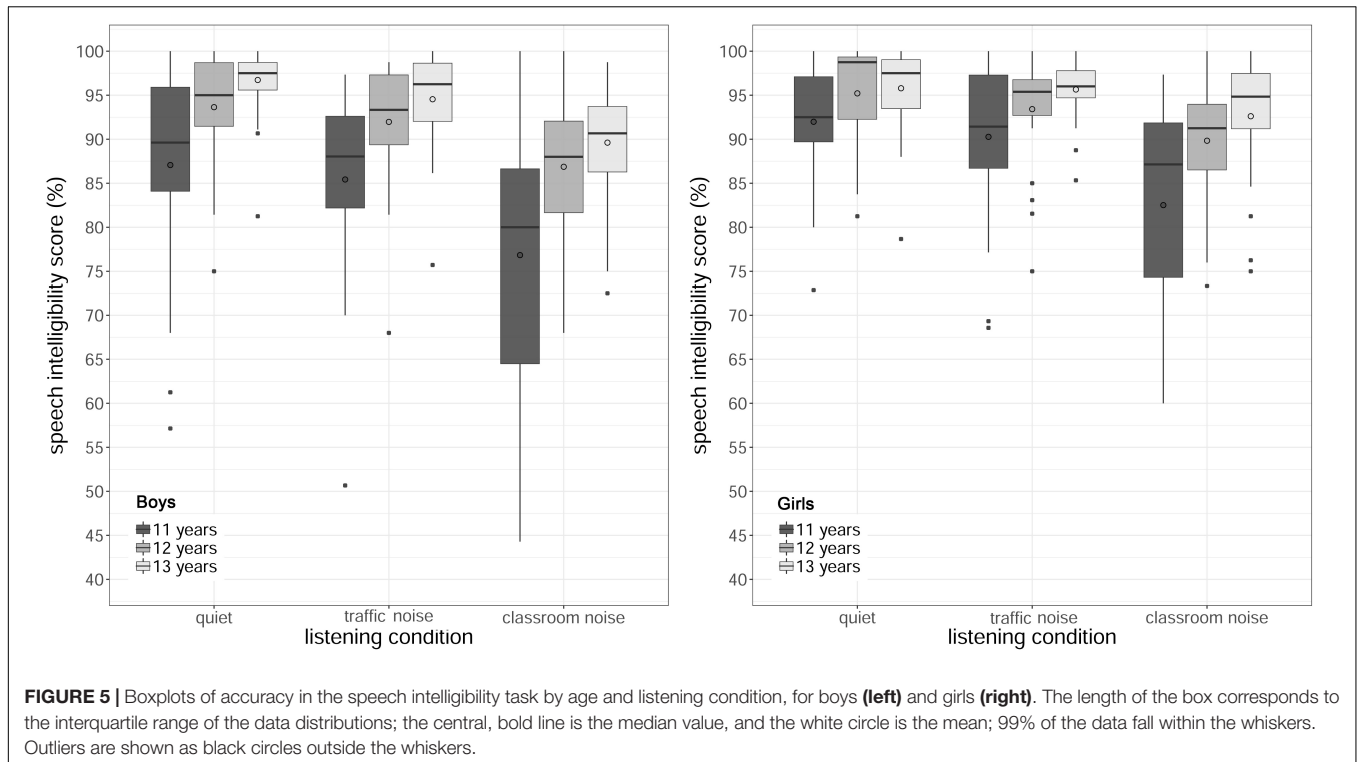
FIGURE 5 | Boxplots of accuracy in the speech intelligibility task by age and listening condition, for boys **(left)** and girls **(right)**. The length of the box corresponds to the interquartile range of the data distributions; the central, bold line is the median value, and the white circle is the mean; 99% of the data fall within the whiskers. Outliers are shown as black circles outside the whiskers.
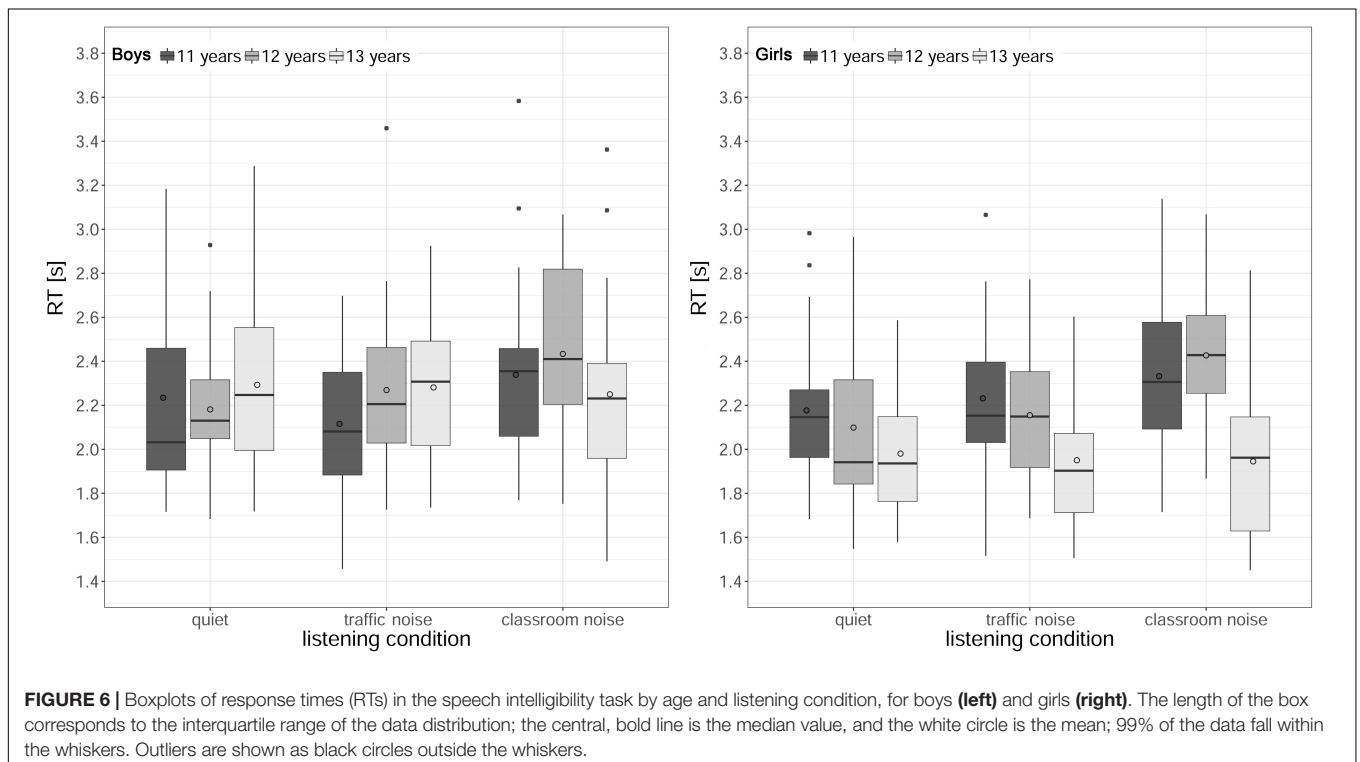
FIGURE 6 | Boxplots of response times (RTs) in the speech intelligibility task by age and listening condition, for boys **(left)** and girls **(right)**. The length of the box corresponds to the interquartile range of the data distribution; the central, bold line is the median value, and the white circle is the mean; 99% of the data fall within the whiskers. Outliers are shown as black circles outside the whiskers.

$[\chi^2(2) = 25.41, p < 0.001]$, and between age and gender $[\chi^2(2) = 25.41, p < 0.001]$ were significant as well. The main effect of the baseline comprehension score and the remaining two- and three-way interactions were not significant (all $ps > 0.15$).

The significant interaction between listening condition and age was considered first, with data collapsed across genders. When the effect of noise was analyzed for each age group, the RTs for the 11- and 12-year-olds were significantly slower

in classroom noise than in quiet or traffic noise conditions, while there was no difference between quiet and traffic noise (11 years: quiet < classroom noise, $z = -3.20$, $p = 0.004$, $\Delta RT = 130$ ms; traffic noise < classroom noise, $z = -2.74$, $p = 0.018$, $\Delta RT = 160$ ms; 12 years: quiet < classroom noise, $z = -4.85$, $p < 0.001$, $\Delta RT = 288$ ms; traffic noise < classroom noise, $z = -3.47$, $p = 0.002$, $\Delta RT = 214$ ms). For the 13-year-olds, on the other hand, there was no difference between listening conditions. When the effect of age was analyzed for each listening condition, pairwise comparisons revealed that RTs only differed across ages in classroom noise, being faster for the oldest students (11 > 13 years, $z = 3.29$, $p = 0.003$, $\Delta RT = 213$ ms; 12 > 13 years, $z = 3.45$, $p = 0.002$, $\Delta RT = 308$ ms). When the interaction between age and gender was analyzed, with data collapsed across listening conditions, *post hoc* tests indicated that it was only among the 13-year-olds that RTs for girls were a mean 316 ms faster than for boys (girls < boys, $z = -3.97$, $p < 0.001$).

Sentence Comprehension

Table 4 shows SC performance accuracy as the percentage of correct answers across ages for the three listening conditions. The results showed a strong ceiling effect, with most pupils achieving or coming close to the highest score in all listening condition. Given this ceiling effect, and the small degree of variance in accuracy in the SC task, only the corresponding RTs were included in the analysis.

Figure 7 shows the RTs in the SC task (median across the trials) for each listening condition and age, for boys and girls. The analysis identified a significant main effect of listening condition [$\chi^2(2) = 30.64$, $p < 0.001$], a main effect of age [$\chi^2(2) = 25.68$, $p < 0.001$], and a main effect of gender [$\chi^2(1) = 7.21$, $p = 0.007$]. The main effect of reading comprehension score was not significant ($p = 0.051$), nor were there any significant two- or three-way interactions (all $ps > 0.38$).

Post hoc tests comparing the listening conditions collapsed across age and gender showed that RTs were significantly slower in classroom noise than in quiet or traffic noise (quiet < classroom noise, $z = -5.30$, $p < 0.001$, $\Delta RT = 314$ ms, traffic noise < classroom noise, $z = -3.19$, $p < 0.001$, $\Delta RT = 239$ ms).

Comparisons between age groups, with data collapsed across listening condition and gender, revealed that RTs were faster for the oldest children (11 > 13 years, $z = 4.95$, $p = < 0.001$, $\Delta RT = 638$ ms; 12 > 13 years, $z = 3.24$, $p = 0.004$, $\Delta RT = 543$ ms).

TABLE 4 | Mean percentage of correct answers and standard deviations (in brackets) in the sentence comprehension task, in the three listening conditions and age groups.

Age group	Listening condition		
	Quiet	Traffic noise	Classroom noise
11 years	95.0 (6.9)	95.3 (6.4)	92.2 (7.5)
12 years	95.7 (4.9)	94.4 (5.5)	93.4 (7.8)
13 years	96.5 (5.7)	95.9 (5.0)	94.7 (6.4)

As for the effect of gender, the boys' RTs were, on average, 319 ms longer than those of the girls.

Comparison of the Effects of Background Noise and Age on the Two Tasks: RTs

Figure 8 shows the RT relative to quiet for each age group, task and noisy listening conditions (traffic noise, classroom noise). Our analysis found a significant main effect of listening condition [$\chi^2(1) = 30.47$, $p < 0.001$], a significant interaction between age and task [$\chi^2(2) = 8.46$, $p = 0.015$], a significant interaction between listening condition and age [$\chi^2(2) = 8.09$, $p = 0.017$], and a significant three-way interaction between listening condition, age and task [$\chi^2(2) = 8.80$, $p = 0.012$]. The main effects of age, gender, task, and baseline comprehension score, and the interaction between listening condition and task were not significant (all $ps > 0.25$).

As shown in **Figure 8**, the three-way interaction was due to a different impact of the two background noises, which depended both on the type of task and on the children's age. For each age group and task, pairwise comparisons were run to analyze the effect of the listening condition. For the 11-year-olds, there was a significant difference between the two noisy listening conditions in both tasks, with traffic noise less invasive than classroom noise (speech intelligibility: $t = -3.31$, $p = 0.006$; sentence comprehension: $t = -3.72$, $p = 0.001$). For the 12-year-olds, the difference between the two listening conditions was only significant for SI (traffic < classroom noise, $t = -4.31$, $p < 0.001$), and no difference was found for the 13-year-olds (all $ps > 0.25$). Whenever a significant difference emerged, it always pointed to classroom noise having a greater impact (prompting a greater increase in RT) than traffic noise.

DISCUSSION

The main aim of this study was to compare SI and SC in lower middle-school students, under three listening conditions (quiet, traffic noise, and classroom noise). Children from 11 to 13 years old were tested to clarify the effects of background noise, whether and how they may be influenced by the listener's age or gender, and whether SI and SC are affected differently. The main findings of our study are discussed below.

Effects of Noise

For both the tasks administered, the children in our sample performed best, and had the fastest RTs in the quiet listening condition. Adding background noise at a sound pressure level typical of a working classroom generally reduced the students' accuracy in the tasks and increased their listening effort (according to their slower RTs). When SI was considered, there was a main effect of listening condition on task accuracy that discriminated between the specific effects of each condition: classroom noise disrupted SI significantly more than traffic noise, which was still more impairing than quiet. In the SC task, on the other hand, a strong ceiling effect emerged for accuracy, probably attributable to the additional cues provided by the pictorial

FIGURE 7 | Boxplots of response times (RTs) in the sentence comprehension task by age and listening condition, for boys **(left)** and girls **(right)**. The length of the box corresponds to the interquartile range of the data distribution; the central, bold line is the median value, and the white circle is the mean; 99% of the data fall within the whiskers. Outliers are shown as black circles outside the whiskers.

representation of the actions. The visual, closed-set format of the test allowed for the inclusion of sentences of different linguistic complexity, but strongly supported listeners trying to complete the task, making the SC task easier than the SI.

As expected, classroom noise impaired performance accuracy in the SI tasks more than traffic noise. The presence of speech-like temporal fluctuations in the masker adversely affects task performance accuracy in verbal tasks by competing with the target speech (Dockrell and Shield, 2006). It should be noted that even notionally steady-state maskers (like the traffic noise used in the present study) can produce modulation masking – which interferes with the target speech processing – for adult listeners (Stone et al., 2011, 2012), but there is no evidence of the same effect in children. The adverse effect of the classroom noise used in the present study may also relate to a capture of attention. In fact, salient sound events (like the events mixed with the ICRA signal) further impair performance accuracy by capturing the listener's attention (Klatte et al., 2010b). This latter mechanism is known to depend on individual attentional abilities (Klatte et al., 2013), which may explain the greater variability in accuracy (i.e., larger standard deviations) seen in the SC task associated with classroom noise (see **Table 4**).

RTs were recorded to see whether the type of noise had the same effect on listening effort as on task performance accuracy. A main effect of listening condition on RTs was found in the SC task, indicating that the children took longer to process what they heard (240 ms) in classroom noise as opposed to quiet or traffic noise. A more complex pattern emerged for the SI task, for which a significant interaction emerged between listening condition and age. The RTs were slower in classroom noise than in quiet or

traffic noise, but only for the 11- and 12-year-olds, not for the 13-year-olds. This would suggest a developmental effect on the strategies for coping with noise, which is discussed in more detail in the next section.

In the SC task, the children in our study were able to cope with traffic noise, which impaired neither their performance accuracy nor their RTs by comparison with the quiet condition. In the SI task traffic noise did not impair the children's RT and only slightly decreased their performance accuracy (by 1.6 percentage points) by comparison with the quiet condition. In classroom noise, however, the increase in the 11- and 12-year-olds' RTs reflected the worsening of the task performance accuracy. This finding is consistent with previous studies on children using RT as a behavioral proxy for listening effort (Prodi et al., 2013, 2019; Lewis et al., 2016; McGarrigle et al., 2019). The latency before a response includes the time listeners take to decode and process the auditory information they have received, so it can be considered informative on the effort invested in the task, or the cognitive resources needed to process the stimulus (Gatehouse and Gordon, 1990; Houben et al., 2013; McGarrigle et al., 2014; Pichora-Fuller et al., 2016). A slower RT is interpreted as a sign of a greater listening effort, and several studies have already found the measure sensitive to adverse conditions, such as a worsening of the SNR. More cognitive resources are needed to process auditory information in degraded listening conditions, leaving fewer resources available for the actual task, and leading to a weaker performance.

Overall, the findings of the present study support the existing literature on the harmful effects of background noise with a fluctuating temporal envelope and salient sound events on

FIGURE 8 | Boxplots of the normalized response times (RTs), by task (speech intelligibility, sentence comprehension), age (11, 12, 13 years) and listening condition (traffic noise, classroom noise). The length of the box corresponds to the interquartile range of the data distribution; the central, bold line is the median value, and the white circle is the mean; 99% of the data fall within the whiskers. Outliers are shown as black circles outside the whiskers.

children performing SI and SC tasks (Klatte et al., 2010a; Prodi et al., 2013), confirming that this also applies to 11- to 13-year-olds.

Effects of Age

Another question addressed in this study was whether children from 11 to 13 years old show any developmental effect on how they cope with background noise in SI and SC tasks. Our interest lay in investigating whether age interacted with type of noise and, if so, whether task performance accuracy and listening effort showed the same pattern of results.

Concerning SC, age had a significant main effect on RTs, the 13-year-old students always answering faster than the 11- or 12-year-olds: the former took 500 ms less time to process the sentences than the latter. This developmental effect in the SC task was unaffected by listening condition, as no interaction emerged between the two factors. This would suggest that the effect of age is due to more basic developmental processes, involving memory functioning or language competences, for instance. Sullivan et al. (2015) found that working memory and vocabulary size (both of which increase with age) contributed to children's comprehension, in both quiet and noise.

It is also worth emphasizing that this difference in RTs in the SC task was seen despite a ceiling effect in the results for task accuracy. This result is in line with studies indicating

that RTs may vary for the same level of task accuracy, and even when listeners have already reached their highest possible level of accuracy. Listening effort may therefore be a totally different construct from task performance accuracy. Several studies witnessed this effect for adults (Houben et al., 2013; Picou et al., 2013), but few have explored it in children (Sahlén et al., 2017; Prodi et al., 2019).

As for the SI task, performance accuracy was significantly lower for 11-year-olds than for the older children already in the quiet condition, and the same difference applied to the noisy conditions – as indicated by the absence of any interaction between age and listening condition. This finding might suggest that 11-year-olds found the ITAMatrix (administered in real classrooms using a fixed-stimuli procedure) more difficult than the older students. In the quiet condition, in which the extremely favorable SNR and the modest contribution of reverberation led us to expect the highest SI results, the 11-year-olds fared significantly worse than the older children, while the 12- and 13-year-olds reached a near-ceiling accuracy – possibly meaning that in a quiet condition an adult-like performance accuracy is acquired by 12 years of age. The age effect observed in the SI task would be in line with many published reports of the ability to perceptually segregate speech from a noise masker being immature in childhood, but adult-like by adolescence. For instance, Leibold and Buss (2013) found that adult-level

performance accuracy was reached already at around 8 years old in a consonant identification task conducted in speech-shaped noise. A mature performance was observed a little later on, by about 9–10 years of age, in other studies (Corbin et al., 2016). This ability appears to develop at different rates, however, also depending on the characteristics of the masker (Wróblewski et al., 2012; Leibold, 2017), and on the stimulus type (Lewis et al., 2016).

When RTs in the SI task are considered, a picture complementary to task performance accuracy can be drawn. No effect of age was seen in quiet or in traffic noise, but in classroom noise the 13-year-olds' RTs were significantly faster. Based on these results, the effects of age on SI in noise would depend on the nature of the masker for listening effort as well. The absence of an age effect in traffic noise could relate to the temporal characteristics of this masker, which is essentially steady-state, with no salient sound events that may capture a child's attention (Klatte et al., 2013). Using a similar traffic noise and SI task, Prodi et al. (2013) found no difference in the RTs of children between 8 and 10 years old, but longer RTs for children aged 6 or 7. The similarity of the experimental setups enable the findings of the two studies to be compared. It may be that, by 8 years old, the presence of traffic noise during a SI task mainly impairs "bottom–up" processing, with less call for additional, explicit cognitive processing.

In classroom noise, there was a significant effect of age on RTs, with older students responding faster. Younger students are more susceptible to sound-induced distractors (e.g., salient sound events) due to their more limited attentional control (Klatte et al., 2010b, Klatte et al., 2013). This means that our 11- and 12-year-old children needed to dedicate more active resources to the task, and this increased their processing time. This finding confirms – and extends up to 12 years of age – a trend already seen in children 6 to 10 years old by Prodi et al. (2013): RTs were significantly slower, under the same masker, the younger the age of the respondent. No difference in RTs emerged between the two background noise conditions for our 13-year-old sample, suggesting that they had already developed the key cognitive abilities needed to cope with speech in noise. No adult group was included in our study, which could have served as a benchmark against which to compare the 13 year olds' results, and judge the age at which processing time may plateau. The age of 12 years seemed crucial to both accuracy and RTs in the SI task: this age group's task performance accuracy was better than that of the younger children, and comparable with that of the older ones, but the 12-year-olds still needed more processing time than the 13-year-olds.

Effects of Gender

Significant differences emerged in the present study between boys' and girls' task performance accuracy and RTs. In the SC task, girls always had shorter processing times than boys. The averaged RT gap was quite large (319 ms), representing 9.4% of the average duration of the COMPRENDO sentences. In the SI task, the girls were 2.2 percentage points more accurate than the boys, but their RTs were only significantly shorter (by 316 ms; 13.7% of the average duration of the ITAMatrix stimuli) at 13 years of age.

Our findings of a better performance in girls confirm the uneven developmental course of speech reception for males and females, and are in line with previous reports on accuracy (Ross et al., 2015). As gender no longer makes a significant difference when adult groups are considered (Ross et al., 2015), this effect may be driven by the development of underlying abilities in the age range considered here, and particularly by gender-related differences in the processing of verbal tasks (Burman et al., 2008; Etchell et al., 2018).

It is worth noting that despite the statistically significant main effect of gender on SI performance accuracy, the difference in the SI scores of male and female was very small (2.2 percentage points referred to a mean SI of 90.7%) and might have a limited relevance in the classroom setting. Differently, the present study shows that RTs can provide some interesting additional information, which have practical implications for the children's performance in classrooms. An interaction between age and gender was found for the SI task, but was not significant for SC. When listening effort was considered, and the analysis was limited to the reception of multiple words (as in the SI tasks), the advantage of females was confined to the 13-year-old group. When a more comprehensive display of processing capacity was needed, however, as in the SC task, the gap between females and males applied at all the ages considered. Given the fast pace of communication in classrooms, and the amount of new information that pupils face during lessons, a slowing down in the processing time of the verbal message would likely have a negative impact on the students' learning. In addition, the RT to a task give information on the effort invested, and an increase in RTs can be taken to reflect an increase in listening effort. A prolonged effort (as requested over the time of a lesson or over the school hours) may lead to an outcome of mental stress and fatigue, which is often associated with slower information processing, decreased level of goal-directed attention, difficulties in focusing on the task, and increased involuntary shifts of attention (Key et al., 2017).

It should be noted that the present RT results (referring to 11- to 13-year-old children) contrast with the report from Sahlén et al. (2017) of slower RTs for girls than for boys when 8-year-olds are considered. Given the similarity of the SC tasks employed in the two studies, the reasons for this discrepancy probably lie in the different age ranges considered, and the dysphonic voice used by Sahlén et al. (2017).

Finally, it is also worth noting that, both in the present study and in the one by Boman (2004), the effect of gender on task performance accuracy did not interact with the listening condition. This would suggest that the effect was not driven by a different sensitivity to noise, but by a more basic difference between the two genders in the 11–13 age range.

Speech Intelligibility Versus Sentence Comprehension

This work compared SI and SC using a standardized audiological test for SI and a standardized test battery for SC. The two tests rely on different levels of speech processing. In the SC task, listeners first have to construct a coherent integrated mental representation of a sentence's meaning by combining lexical,

semantic and syntactic information; then they must choose the appropriate image on the screen after comparing with confusing competitors. In the SI task, listeners have to recognize and sequentially select all the words of a sentence, without contextual or semantic cues to support the recall phase. It would therefore be inappropriate to compare the absolute results of the two tasks directly, so changes in RT in noisy conditions relative to quiet were considered. Using normalized quantities, the additional negative effects of noise on response latencies in the two tasks were compared after the effects of age and gender had been partialled out of the analysis.

The results indicated that the type of noise affected RTs differently depending on the participants' age. In particular, a significant three-way interaction was found between task, age and noise, reflecting a developmental effect on how the children coped with the more challenging classroom noise. This suggests that, when the burden on cognitive processes is considered, the comparison between the two tasks might be even more challenging than the one revealed by accuracy alone, as reported in previous studies. When SI and SC were compared in both adults (Hustad, 2008; Fontan et al., 2015) and primary school children (Klatte et al., 2007), SI scores proved to be poor predictors of comprehension performance accuracy in quiet conditions (Hustad, 2008). In addition, the two tasks were differently affected by background noise level (Fontan et al., 2015) and the spectro-temporal characteristics of the masker (Klatte et al., 2007). Generally speaking, transposing SI results (in quiet or in noise) directly to SC might not be meaningful, and acoustic conditions that guarantee optimal SI might not be equally adequate for SC. This issue needs clarification because most currently-used technical means for assessing room acoustics rely on SI, and have no clear and unambiguous connection with SC.

Judging from what we know for now, it does not seem that a simple relationship can capture the link between SI and SC tasks (as hypothesized, for instance, by Hygge, 2014), as it is strongly affected by the characteristics of the tasks themselves. The choice of using tasks based on different speech materials and the presence of a strong ceiling effect on the accuracy on SC task, prevented the possibility of directly exploring the relationship between SI and SC in the present study. However, the SC method applied here presents two main advantages: its easy pictorial implementation and the chance to obtain accuracy and RT data simultaneously – features that make the SC test appropriate for different categories of listeners, and students in particular.

Study Limitations and Future Directions

The present study has some limitations. The hearing sensitivity was not measured for the children participating in the study, and the presence of possible hearing impairments was based only on the parent and teacher's reports. In addition, the SC performance accuracy results showed a strong ceiling effect in all listening condition and for all ages. This happened despite the test being based on sentences of different lexical difficulty. Given the limited number of sentences in each list, a reliable statistical analysis including complexity as an explanatory variable could not be pursued. That said, exploratory analysis suggested a significant

trend of declining performance accuracy (and slowing RTs) with increasing sentence difficulty. Aiming to investigate the effect of syntactic complexity and its possible interaction with the noise type, future studies might consider more sentences for each complexity level and include the sentence difficulty as a factor in the analysis of the task performance accuracy.

The near-ceiling results also prevented any direct comparison between SC and SI, as concerns performance accuracy. The interactions identified by our analysis on the normalized RTs give us the impression that a more extensive comparison would be worthwhile. In particular, it would be important to explore a wider range of reverberations and SNRs, using maskers comprising more competing talkers or intelligible speech. These manipulations would improve our understanding of the objective characteristic of maskers that mediate the relationship between the two tasks.

The results of our study indicate that the ITAMatrix may not be suitable for 11-year-old children in classrooms, because they were unable to perform as well as the 12- and 13-year-olds even in quiet condition. The reasons behind this finding warrant further investigation, the first step being to see whether the same pattern of results is seen at this age in anechoic conditions too. It may be that this age group would manage better with the simplified version of the Matrix Sentence Test (with three- instead of five-word sentences). The applicability of the simplified ITAMatrix has been demonstrated in clinical settings for children 5 to 10 years old (Puglisi et al., 2018), and in both noisy and anechoic conditions the performance of 10-year-olds already approached that of adults. Using this simplified test for older pupils (12–13 years old) as well would level the task difficulty between the age groups. Finally, Puglisi et al. (2015) established the presence of a practice effect when the ITAMatrix is presented in a clinical setting, using an adaptive procedure converging at a SI = 50%; two test lists of 20 sentences are recommended to account for the effect. In the present study higher SI values were targeted (due to the realistic listening conditions selected for the experiment), a constant stimuli paradigm was used, and the test was presented collectively and not at the individual level. Given the much simpler procedure than in a clinical setting, the children were expected to accustom to it more easily reducing the practice effect, and only four sentences were presented during the training phase of the task. Even though the potential presence of training effects was addressed by counterbalancing the listening conditions among the classes, there might be remaining training effects depending on the age of the children.

CONCLUSION

The present study provides evidence that supports previous reports, and also better frames the relationships between type of noise, age, gender, and task. The main results can be summarized as follows.

Effects of age and listening condition were found mainly for the SI task, on both accuracy and RTs. The most demanding condition was in classroom noise, when the SI scores were lowest and the RTs slowest. In this condition, 11- and 12-year-olds

needed the same processing time, but the former group scored lower for accuracy. The 12-year-olds already performed as well as the 13-year olds in terms of accuracy, but with slower RTs. The oldest students had the fastest RTs. A pattern for SI thus emerged, with improvements in task performance accuracy preceding improvements in processing time. This is consistent with findings in younger children and presumably due to a mechanism whereby the cognitive processes underpinning speech reception are first acquired and later consolidated. In the SC task, accuracy scores neared the ceiling, meaning that merging accuracy and RT data was not as informative as in the SI task.

This study also confirmed the effects of gender on the SI and SC tasks. In particular, a main effect of gender was found on the latter task, indicating that the gap between girls and boys was wider for the task of greater linguistic complexity that engaged the pupils in a listening situation more closely resembling actual communication in classrooms. Standardized tests should be developed to include the assessment of this competence when designing for classroom acoustics. Mitigating the gender bias in SC could prove difficult, however, as it may involve class management and how classes are organized.

Finally, our study showed that classroom noise slowed response latencies by comparison with the quiet condition in both SC and SI. Since several factors – such as the nature of background noise, and children's age – appear to affect differently the two tasks, it will be necessary to develop specific test settings to investigate a possible model linking SC and SI.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

REFERENCES

Astolfi, A., Bottalico, P., and Barbato, G. (2012). Subjective and objective speech intelligibility investigations in primary school classrooms. *J. Acoust. Soc. Am.* 131, 247–257. doi: 10.1121/1.3662060

Astolfi, A., Corrado, V., and Griginis, A. (2008). Comparison between measured and calculated parameters for the acoustical characterization of small classrooms. *Appl. Acoust.* 69, 966–976. doi: 10.1016/j.apacoust.2007.08.001

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Boman, E. (2004). The effects of noise and gender on children's episodic and semantic memory. *Scand. J. Psychol.* 45, 407–416. doi: 10.1111/j.1467-9450.2004.00422.x

Bottalico, P., and Astolfi, A. (2012). Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms. *J. Acoust. Soc. Am.* 131, 2817–2827. doi: 10.1121/1.3689549

Bradley, J. S., Reich, R. D., and Norcross, S. G. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. *J. Acoust. Soc. Am.* 106, 1820–1828. doi: 10.1121/1.427932

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Burman, D. D., Bitan, T., and Booth, J. R. (2008). Sex differences in neural processing of language among children. *Neuropsychologia* 46, 1349–1362. doi: 10.1016/j.neuropsychologia.2007.12.021

Cecchetto, C., Di Domenico, A., Garraffa, M., and Papagno, C. (2012). *Comprendo. Batteria per la Comprensione di frasi negli adulti.* Milan: Raffaele Cortina Editore.

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (2016). Development of open-set word recognition in children: speech-shaped noise and two-talker speech maskers. *Ear Hear.* 37, 55–63. doi: 10.1097/AUD.0000000000000201

Cornoldi, C., Carretti, B., and Colpo, C. (2017). *Prove MT-Kit Scuola. Dalla valutazione degli apprendimenti di lettura e comprensione al potenziamento.* Firenze: Organizzazioni Speciali Editore.

Deutsche Institut für Normung. (2016). *Hörsamkeit in Räumen – Anforderungen, Empfehlungen und Hinweise für die Planung (Acoustic quality in rooms – Specifications and instructions for room acoustics design): DIN18041:2016.* Berlin: Deutsche Institut für Normung.

Dockrell, J., and Shield, B. (2006). Acoustical barriers in classrooms: the impact of noise on performance in the classroom. *Br. Educ. Res. J* 32, 509–525. doi: 10.1080/01411920600635494

Downs, D. W., and Crum, M. A. (1978). Processing demands during auditory learning under degraded listening conditions. *J. Speech Lang. Hear. Res.* 21, 702–714. doi: 10.1044/jshr.2104.702

Dreschler, W. A., Verschure, H., Ludvigsen, C., and Westermann, S. (2001). ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. *Audiology* 40, 148–157. doi: 10.3109/00206090109073110

Etchell, A., Adhikari, A., Weinberg, L. S., Choo, A. L., Garnett, E. O., Chow, H. M., et al. (2018). A systematic literature review of sex differences in childhood

language and brain development. *Neuropsychologia* 114, 19–31. doi: 10.1016/j.neuropsychologia.2018.04.011

Everitt, B. S., and Hothorn, T. (2010). *A Handbook of Statistical Analyses Using R*. New York, NY: Chapman and Hall.

Fontan, L., Ferrané, I., Farinas, J., Pinquier, J., Tardieu, J., Magnen, C., et al. (2017). Automatic speech recognition predicts speech intelligibility and comprehension for listeners with simulated age-related hearing loss. *J. Speech Lang. Hear. Res.* 60, 2394–2405. doi: 10.1044/2017_JSLHR-S-16-0269

Fontan, L., Tardieu, J., Gaillard, P., Woisard, V., and Ruiz, R. (2015). Relationship between speech intelligibility and speech comprehension in babble Noise. *J. Speech Lang. Hear. Res.* 58, 977–986. doi: 10.1044/2015_JSLHR-H-13-0335

Füllgrabe, C., Berthommier, F., and Lorenzi, C. (2006). Masking release for consonant features in temporally fluctuating background noise. *Hear. Res.* 211, 74–84. doi: 10.1016/j.heares.2005.09.001

Gatehouse, S., and Gordon, J. (1990). Response times to speech stimuli as measures of benefit from amplification. *Br. J. Audiol.* 24, 63–68. doi: 10.3109/03005369009077843

Gordon, K. R. (2019). How mixed-effects modeling can advance our understanding of learning and memory and improve clinical and educational practice. *J. Speech Lang. Hear. Res.* 62, 507–524. doi: 10.1044/2018_JSLHR-L-ASTM-18-0240

Hällgren, M., Larsby, B., Lyxell, B., and Arlinger, S. (2001). Evaluation of a cognitive test battery in young and elderly normal-hearing and hearing-impaired persons. *J. Am. Acad. Audiol.* 12, 357–370.

Houben, R., van Doorn-Bierman, M., and Dreschler, W. A. (2013). Using response time to speech as a measure for listening effort. *Int. J. Audiol.* 52, 753–761. doi: 10.3109/14992027.2013.832415

Hughes, R., and Jones, D. M. (2001). The intrusiveness of sound: laboratory findings and their implications for noise abatement. *Noise Health* 4, 51–70.

Hustad, K. (2008). The relationship between listener comprehension and intelligibility scores for speakers with dysarthria. *J. Speech Lang. Hear. Res.* 51, 562–573. doi: 10.1044/1092-4388

Hygge, S. (2014). "Classroom noise and its effect on learning," in *Proceedings of the 11th International Congress on Noise as a Public Health Problem (ICBEN)*, Vol. 4, Nara.

International Electrotechnical Commission, (2011). *Sound system equipment. Part 16: Objective rating of speech intelligibility by speech transmission index: IEC 60286-16*. Geneva: International Electrotechnical Commission.

International Organization of Standardization, (2003). *Ergonomics – Assessment of Speech Communication: ISO9921:2003*. Geneva: International Organization of Standardization.

International Organization of Standardization, (2009). *Acoustics – Measurement of room acoustic parameters. In Part 1: Performance spaces: ISO3382-1:2009*. Geneva: International Organization of Standardization.

Johnson, C. E. (2000). Children's phoneme identification in reverberation and noise. *J. Speech Lang. Hear. Res.* 43, 144–157. doi: 10.1044/jslhr.4301.144

Key, A. P., Gustafson, S. J., Rentmeester, L., Hornsby, B. W., and Bess, F. H. (2017). Speech-processing fatigue in children: auditory event-related potential and behavioral measures. *J. Speech Lang. Hear. Res.* 60, 2090–2104. doi: 10.1044/2016_JSLHR-H-16-0052

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4:578. doi: 10.3389/fpsyg.2013.00578

Klatte, M., Lachmann, T., and Meis, M. (2010a). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise Health* 12, 270–282. doi: 10.4103/1463-1741.70506

Klatte, M., Lachmann, T., Schlittmeier, S., and Hellbrück, J. (2010b). The irrelevant sound effect in short-term memory: is there developmental change? *Eur. J. Cogn. Psychol.* 22, 1168–1191. doi: 10.1080/09541440903378250

Klatte, M., Meis, M., Sukowski, H., and Schick, A. (2007). Effects of irrelevant speech and traffic noise on speech perception and cognitive performance in elementary school children. *Noise Health* 9, 64–74.

Laudanna, A., Thornton, A. M., Brown, G., Burani, C., and Marconi, L. (1995). "Un corpus dell'italiano scritto contemporaneo dalla parte del ricevente," in *III Giornate internazionali di Analisi Statistica dei Dati Testuali*, Vol. I, eds S. Bolasco, L. Lebart, and A. Salem, (Roma: Cisu), 103–109.

Leibold, L. J. (2017). Speech perception in complex acoustic environments: developmental effects. *J. Speech Lang. Hear. Res.* 60, 3001–3008. doi: 10.1044/2017_JSLHR-17-0070

Leibold, L. J., and Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *J. Speech Lang. Hear. Res.* 56, 1144–1155. doi: 10.1044/1092-4388(2012/12-0011)

Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33. doi: 10.18637/jss.v069.i01

Lewis, D., Schmid, K., O'Leary, S., Spalding, J., Heinrichs-Graham, E., and High, R. (2016). Effects of noise on speech recognition and listening effort in children with normal hearing and children with mild bilateral or unilateral hearing loss. *J. Speech Lang. Hear. Res.* 59, 1218–1232. doi: 10.1044/2016_JSLHR-H-15-0207

Lo, S., and Andrews, S. (2015). To transform or not to transform: using generalized linear mixed models to analyse reaction time data. *Front. Psychol.* 6:1171. doi: 10.3389/fpsyg.2015.01171

Lyberg-Åhlander, V., Holm, L., Kastberg, T., Haake, M., Brännström, K. J., and Sahlén, B. (2015). Are children with stronger cognitive capacity more or less disturbed by classroom noise and dysphonic teachers? *Int. J. Speech Lang. Pathol.* 17, 577–588. doi: 10.3109/17549507.2015.1024172

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

McGarrigle, R., Gustafson, S. J., Hornsby, B. W., and Bess, F. H. (2019). Behavioral measures of listening effort in school-age children: examining the effects of signal-to-noise ratio, hearing loss, and amplification. *Ear Hear.* 40, 381–392. doi: 10.1097/AUD.0000000000000623

McGarrigle, R., Munro, K. J., Dawes, P., Stewart, A. J., Moore, D. R., Barry, J. G., et al. (2014). Listening effort and fatigue: what exactly are we measuring? A British society of audiology cognition in hearing special interest group 'white paper'. *Int. J. Audiol.* 53, 433–440. doi: 10.3109/14992027.2014.890296

Meinhardt-Injac, B., Schlittmeier, S. J., Klatte, M., Otto, A., Persike, M., and Imhof, M. (2014). Auditory distraction by meaningless irrelevant speech. A developmental study. *Appl. Cogn. Psychol.* 29, 217–225. doi: 10.1002/acp.3098

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1097/AUD.0b013e3181d3d514

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W., Humes, L. E., et al. (2016). Hearing impairment and cognitive energy: the framework for understanding effortful listening (FUEL). *Ear Hear.* 37, 5S–27S. doi: 10.1097/AUD.0000000000000312

Picou, E. M., Gordon, J., and Ricketts, T. A. (2016). The effects of noise and reverberation on listening effort for adults with normal hearing. *Ear Hear.* 37, 1–13. doi: 10.1097/AUD.0000000000000222

Picou, E. M., Ricketts, T. A., and Hornsby, B. W. (2013). How hearing aids, background noise, and visual cues influence objective listening effort. *Ear Hear.* 34, e52–e64. doi: 10.1097/AUD.0b013e31827f0431

Postma, B. N., and Katz, B. F. (2016). Perceptive and objective evaluation of calibrated room acoustic simulation auralizations. *J. Acoust. Soc. Am.* 140, 4326–4337. doi: 10.1121/1.4971422

Prodi, N., Visentin, C., and Feletti, A. (2013). On the perception of speech in primary school classrooms: ranking of noise interference and of age influence. *J. Acoust. Soc. Am.* 133, 255–268. doi: 10.1121/1.4770259

Prodi, N., Visentin, C., Peretti, A., Griguolo, J., and Bartolucci, G. B. (2019). Investigating listening effort in classrooms for 5-to 7-year-old children. *Lang. Speech Hear. Serv. Sch.* 50, 196–210. doi: 10.1044/2018_LSHSS-18-0039

Puglisi, G. E., Calleri, C., Astolfi, A., Warzybok, A., Sellami, F., et al. (2018). "An effective test for the assessment of speech intelligibility in noise with children," in *Proceedings of the EURONOISE2018 Conference*, Heraklion.

Puglisi, G. E., Warzybok, A., Hochmuth, S., Visentin, C., Astolfi, A., Prodi, N., et al. (2015). An Italian matrix sentence test for the evaluation of speech intelligibility in noise. *Int. J. Audiol.* 54(Suppl. 2), 44–50. doi: 10.3109/14992027.2015.1061709

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available at: https://www.R-project.org/

Ross, L. A., Del Bene, V. A., Molholm, S., Frey, H. P., and Foxe, J. J. (2015). Sex differences in multisensory speech processing in both typically developing children and those on the autism spectrum. *Front. Neurosci.* 9:185. doi: 10.3389/fnins.2015.00185

Rudner, M., and Lunner, T. (2014). Cognitive spare capacity and speech communication: a narrative overview. *BioMed Res. Int.* 2014:869726. doi: 10.1155/2014/869726

Rudner, M., Lyberg-Åhlander, V., Brännström, J., Nirme, J., Pichora-Fuller, M. K., and Sahlén, B. (2018). Listening comprehension and listening effort in the primary school classroom. *Front. Psychol.* 9:1193. doi: 10.3389/fpsyg.2018.01193

Sahlén, B., Haake, M., von Lochow, H., Holm, L., Kastberg, T., Brännström, K. J., et al. (2017). Is children's listening effort in background noise influenced by the speaker's voice quality? *Logoped. Phoniatr. Vocol.* 43, 47–55. doi: 10.1080/14015439.2017.1324914

Shield, B., Conetta, R., Dockrell, J., Connolly, D., Cox, T., and Mydlarz, C. (2015). A survey of acoustic conditions and noise levels in secondary school classrooms in England. *J. Acoust. Soc. Am.* 137, 177–188. doi: 10.1121/1.4904528

Stone, M. A., Füllgrabe, C., Mackinnon, R. C., and Moore, B. C. (2011). The importance for speech intelligibility of random fluctuations in "steady" background noise. *J. Acoust. Soc. Am.* 130, 2874–2881. doi: 10.1121/1.3641371

Stone, M. A., Füllgrabe, C., and Moore, B. C. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317–326. doi: 10.1121/1.4725766

Sullivan, J. R., Osman, H., and Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *J. Speech Lang. Hear. Res.* 58, 1043–1051. doi: 10.1044/2015_JSLHR-H-14-0204

Uslar, V. N., Carroll, R., Hanke, M., Hamann, C., Ruigendijk, E., Brand, T., et al. (2013). Development and evaluation of a linguistically and audiologically controlled sentence intelligibility test. *J. Acoust. Soc. Am.* 134, 3039–3056. doi: 10.1121/1.4818760

Valente, L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., and Lewis, D. E. (2012). Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and

learning in children. *J. Acoust. Soc. Am.* 131, 232–246. doi: 10.1121/1.3662059

von Lochow, H., Lyberg-Åhlander, V., Sahlén, B., Kastberg, T., and Brännström, K. J. (2018). The effect of voice quality and competing speakers in a passage comprehension task: performance in relation to cognitive functioning in children with normal hearing. *Logop. Phoniatr. Voco.* 43, 11–19. doi: 10.1080/14015439.2017.1298835

Wagener, K., Kuhnel, V., and Kollmeier, B. (1999). Entwicklung und Evaluation eines Satztests für die deutsche Sprache I: Design des Oldenburger Satztests (Development and evaluation of a speech intelligibility test for German I: Design of the Oldenburg sentence test). *Z. Audiologie* 38, 4–15.

Wendt, D., Koelewijn, T., Książek, P., Kramer, S. E., and Lunner, T. (2018). Toward a more comprehensive understanding of the impact of masker type and signal-to-noise ratio on the pupillary response while performing a speech-in-noise test. *Hear. Res.* 369, 67–78. doi: 10.1016/j.heares.2018.05.006

Wolf, M. C., Muijselaar, M. M. L., Boonstra, A. M., and de Bree, E. H. (2019). The relationship between reading and listening comprehension: shared and modality-specific components. *Read. Writ.* 32, 1747–1767. doi: 10.1007/s11145-018-9924-8

Wróblewski, M., Lewis, D. E., Valente, D. L., and Stelmachowicz, P. G. (2012). Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. *Ear Hear.* 33, 731–744. doi: 10.1097/AUD.0b013e31825aecad

Early Cognitive Predictors of 9-Year-Old Spoken Language in Children With Mild to Severe Hearing Loss Using Hearing Aids

Teresa Y. C. Ching[1,2]*, Linda Cupples[3,4] and Vivienne Marnane[1,2]

[1] *National Acoustic Laboratories, Sydney, NSW, Australia,* [2] *The Hearing CRC, Melbourne, VIC, Australia,* [3] *Department of Linguistics, Macquarie University, Sydney, NSW, Australia,* [4] *Centre for Language Sciences, Macquarie University, Sydney, NSW, Australia*

This study examined the extent to which cognitive ability at 5 years of age predicted language development from 5 to 9 years of age in a population-based sample of children with hearing loss who participated in the Longitudinal Outcomes of Children with Hearing Impairment (LOCHI) study. The developmental outcomes of 81 children with hearing loss were evaluated at 5 and 9 years of age. Hearing loss ranged from mild to severe degrees, and all participants used hearing aids. They all used spoken language as the primary mode of communication and education. Nine-year-old language was assessed using the Clinical Evaluation of Language Fundamentals – 4th edition (CELF-4), the Peabody Picture Vocabulary Test – 4th edition (PPVT-4), and the Expressive Vocabulary Test – 2nd edition (EVT-2). Multiple regression analyses were conducted to examine the extent to which children's scores on these standardized assessments were predicted by their cognitive ability (non-verbal IQ and verbal working memory) measured at 5 years of age. The influence of early language scores at 5 years and a range of demographic characteristics on language scores at 9 years of age was evaluated. We found that 5-year-old digit span score was a significant predictor of receptive and expressive language, but not receptive or expressive vocabulary, at 9 years of age. Also, 5-year-old non-word repetition test score was a significant predictor of only expressive language and vocabulary, but not receptive language or vocabulary at 9 years of age. After allowing for the effects of non-verbal IQ and 5-year-old receptive vocabulary, early digit span score (but not non-word repetition score) was a significant predictor of expressive and receptive language scores at 9 years of age. The findings shed light on the unique role of early verbal working memory in predicting the development of receptive and expressive language skills and vocabulary skills in children who use hearing aids.

Keywords: short-term memory, language, cognitive predictors, hearing aids, children with hearing loss

INTRODUCTION

Children with hearing loss achieve lower language outcomes, on average, than children with normal hearing. Findings from a recent, population-based study, the Longitudinal Outcomes of Children with Hearing Impairment or "LOCHI" study, show that 5-year-old children with hearing loss are about 0.5–1 SD on average behind their normally hearing peers on standardized tests of receptive and expressive language and receptive vocabulary (Cupples et al., 2018). Average scores conceal

marked variability, however, in the outcomes achieved by individual children with hearing loss. The demographic and audiological variables that contribute to this variation have been widely studied in recent research, but questions remain regarding the possible influence of early cognitive variables. The current research aimed to shed light on this unresolved issue. In doing so, the intention was not to evaluate the reciprocal view, that early language impacts later cognitive ability, because that view has not been cast into doubt on the basis of recent research (e.g., Botting et al., 2017; Jones et al., 2019). On the other hand, neither Botting et al. (2017) nor Jones et al. (2019) found evidence for an influence of cognitive ability on language outcomes, an important finding that warrants further attention.

Cognition and Language in Cochlear Implant Users

Much of the past research on this topic has focused on the association between executive function and language outcomes achieved by cochlear implant users. This focus has often been underpinned by a theoretical perspective in which early exposure to sound is seen as vital for the typical development of cognitive abilities that relate to the representation and processing of sequential information, which in turn is associated with language development (Conway et al., 2009, 2011a,b; Pisoni et al., 2016). Although there have been several failures to replicate critical empirical findings cited in support of the theory (e.g., Giustolisi and Emmorey, 2018; Hall et al., 2018; von Koss Torkildsen et al., 2018), evidence for a link between cognitive ability and language outcomes remains.

In a study of 64 cochlear implant users and 74 normally hearing peers, ages 7–27 years, Kronenberger et al. (2014) examined the association between four composite measures of executive function (verbal working memory, spatial working memory, fluency speed skills, and inhibition-concentration) and a composite language measure, which encompassed receptive vocabulary (the Peabody Picture Vocabulary Test – 4th edition; Dunn and Dunn, 2007) and the core language score from the Clinical Evaluation of Language Fundamentals – 4th edition (CELF-4; Semel et al., 2003). A regression analysis, conducted across the two groups of participants combined, showed that non-verbal ability, hearing status, verbal working memory, and spatial working memory each accounted for significant unique variance in language outcomes measured concurrently. However, the strength of the relationship between language and working memory varied according to hearing status, such that verbal working memory was a stronger predictor of language outcomes in cochlear implant users than their normally hearing peers, and spatial working memory was a relatively weaker predictor in cochlear implant users.

In accordance with the need for longitudinal studies to address questions of development, Pisoni et al. (2011) examined the extent to which later language outcomes (at 15–18 years of age) could be predicted by earlier digit span and verbal rehearsal speed (at 8–9 years of age) in a sample of 112 cochlear implant users. The results showed that digit span standard scores were significantly below that of a normative age-matched sample at both assessment time points, and that digit span forward, digit span backward, and verbal rehearsal speed were all significantly positively correlated with later language scores on the PPVT and CELF. The researchers noted further that similar correlations were obtained even when controlling for a range of relevant audiological variables including age at cochlear implantation. An obvious question, however, is whether early cognitive variables would still be associated with later language variables if early language variables were controlled.

Early Language and Later Language in Cochlear Implant Users

Several recent studies of cochlear implant users have provided evidence that early language abilities are a good predictor of later language and cognitive outcomes. Castellanos et al. (2016a) examined the association between early receptive vocabulary (assessed using the PPVT-3 at 3–6 years of age) and later receptive vocabulary (using the PPVT-4 at 7.8–23.4 years of age) in a small sample of 19 cochlear implant users. They reported a significant positive association between the two vocabulary measures, and noted that, in a regression analysis, demographic and audiological variables explained little additional unique variance in later PPVT-4 once early PPVT-3 was included. Similarly, a study of 51 cochlear implant users reported by Nittrouer et al. (2016) revealed that morpho-syntactic ability at 8.6 years of age, which comprised a set of narrative measures including mean length of utterance (MLU) in morphemes, number of conjunctions, and number of pronouns, was best predicted by the same narrative measures collected at 3 and 6 years of age, and by expressive vocabulary measured at 4 years of age. Castellanos et al. (2016b) found that a parent-reported measure of expressive vocabulary collected within 2.5 years of cochlear implantation strongly predicted long-term language (assessed using PPVT-4 and CELF-4) and memory outcomes (assessed using digit span forward and backward, and visual digit span) from 5 to 16 years later in a sample of 32 cochlear implant users (9–22 years of age at follow-up). Collectively, these results underscore the importance of controlling for early language ability when examining the association between early cognitive predictors and later language outcomes.

Further supporting evidence comes from Hunter et al. (2017) who investigated whether early language, assessed within 1 year of cochlear implantation, could predict later outcomes in language and executive functioning, in particular, verbal working memory. A sample of 36 adolescent and young adult cochlear implant users (ages 11.6–27.4 years) completed assessments of executive functioning (including verbal and visuo-spatial working memory), receptive and expressive language (CELF-4 core language subtests), and receptive vocabulary (PPVT-4). Regression analyses showed that measures of early speech and language ability accounted for significant unique variance in later language outcomes, while controlling for variation in age at implantation (which was also significant), degree of hearing loss, household income, and non-verbal IQ (which were not significant). A total of 46% of variance in later language outcomes was explained. Using the same set of predictor variables, however,

a total of 71% of variance in later verbal working memory was explained, with all predictor variables accounting for significant unique variance. Hunter et al. (2017) findings are therefore indicative of an association between early language and later cognitive outcomes, a view that gains further support from the literature via empirical studies examining outcomes in groups of deaf and hard-of-hearing (DHH) people more broadly, rather than cochlear implant users in particular.

Cognition and Language in Diverse Groups of DHH Individuals

Figueras et al. (2008) reported evidence to suggest that hearing-related group differences in executive function were underpinned by differences in language ability. They found a significant positive association between language ability (receptive vocabulary and sentence comprehension) and executive function (planning, set shifting, working memory, impulse regulation, and visual attention) in both normally hearing children, after controlling for age, and children with hearing loss, after controlling for age, degree of hearing loss, and duration of use of the current device (22 children used cochlear implants and 25 used hearing aids). They also reported, however, that hearing-related group differences in measures of executive function were non-significant once language ability was controlled.

In a similar vein, as noted earlier, both Botting et al. (2017) and Jones et al. (2019) reported evidence to suggest that language ability predicts executive function in children with hearing loss, but executive function does not, in turn, predict language outcomes. Botting et al. (2017) conducted a mediation analysis using concurrent measures of language and executive function, and showed that language ability mediated differences in executive function between groups of deaf and hearing children. More specifically, after controlling for language, group no longer predicted executive function; whereas after controlling for executive function, group differences in language remained significant. On the other hand, Jones et al. (2019) used longitudinal data to demonstrate that early expressive vocabulary (assessed at around 8 years of age) predicted later executive function (at around 10 years old), but not the reverse.

In comparing the results of the studies by Botting et al. (2017) and Jones et al. (2019) with the findings from studies of cochlear implant users, two important methodological differences are noteworthy. As already mentioned, participants in the studies by Botting et al. (2017) and Jones et al. (2019) constituted a more diverse group of children using hearing aids or cochlear implants and various modes of communication, including British Sign Language, Sign supported English, or spoken English. In addition, language was assessed using a single measure of *expressive* vocabulary and executive function using a battery of explicitly *non-verbal* measures. By contrast, assessments of cochlear implant users often targeted *receptive* vocabulary and *verbal* measures of cognitive ability. It is unknown the extent to which these differences in participant samples and methodology might have influenced the findings, but what is apparent is the need for further systematic investigation into

whether early measures of cognitive ability predict later language outcomes in children with mild to severe hearing loss who use hearing aids.

The Current Study

The current study addressed this gap in the literature. The participant sample was drawn from the cohort taking part in the population-based LOCHI study in which children with hearing aids outnumber those with cochlear implants approximately 2:1. Although the advent of universal newborn hearing screening has made it possible for early detection and fitting of hearing devices to children with permanent childhood hearing loss, it remains uncertain as to whether those at risk for suboptimal long-term outcomes for speech and language may be identified through early measures of speech, language, and working memory. We took advantage of the prospective nature of the LOCHI study to examine the influence of early cognitive and language abilities on later language abilities for children with permanent hearing loss using hearing aids. The study measured outcomes of a population-based cohort of about 450 children in Australia who were born with hearing loss and received hearing intervention before 3 years of age. Details of the study have been reported in Ching et al. (2013). As part of the study, the demographic characteristics and developmental outcomes of children were examined at chronological ages of 3, 5, and 9 years of age. The current study draws on data collected at 5 years of age for predicting outcomes collected at 9 years of age.

From a theoretical perspective, we drew on the multi-component model of working memory described by Baddeley et al. (1998) and Baddeley and Hitch (2019), with a particular focus on the role of the phonological short-term store or phonological loop in language learning. According to Baddeley et al. (1998), the phonological loop mediates language learning, especially vocabulary development, by enabling the temporary storage of new phonological forms while long-term representations are established. They also acknowledge, however, that as knowledge of language increases, learners can use that language knowledge to support new word learning and thereby reduce reliance on the phonological loop. In accordance with this view, an evaluation of the role of early phonological memory in later language development should include a control for the impact of early language ability.

In the current study, the capacity of the phonological loop was measured using both memory span for digits and non-word repetition (the ability to repeat an unfamiliar spoken form). These measures were selected in light of positive results from previous research with normally hearing children (e.g., Gathercole et al., 1992; Avons et al., 1998). Language was assessed using standardized measures of receptive and expressive skills, including two measures of vocabulary development. Finally, a measure of non-verbal cognitive ability was included along with other relevant demographic variables (age at hearing aid fitting, degree of hearing loss, and maternal education) to evaluate the unique contribution of our primary predictors to language outcomes.

The research questions addressed in the study were as follows:

1. Does the capacity of the phonological loop, assessed at 5 years of age, predict 9-year-old outcomes in receptive and expressive language and vocabulary in a population-based sample of hearing aid users after controlling for variation in non-verbal cognitive ability and relevant demographic variables? If so, which measure of 5-year-old phonological short-term memory, digit span, or non-word repetition, is more strongly associated with 9-year-old outcomes?

2. Does the capacity of the phonological loop, assessed at 5 years of age, predict 9-year-old outcomes in receptive and expressive language and vocabulary in a population-based sample of hearing aid users after controlling for 5-year-old receptive vocabulary, non-verbal cognitive ability, and relevant demographic variables? If so, which measure of 5-year-old short-term memory, digit span, or non-word repetition, is more strongly associated with 9-year-old outcomes?

In accordance with the working memory theoretical framework and empirical findings from previous literature with cochlear implant users (e.g., Pisoni et al., 2011), we hypothesized that (1) higher digit span and non-word repetition scores at 5 years of age would be associated with better language and vocabulary outcomes at 9 years of age after controlling for 5-year-old non-verbal ability and relevant demographic variables; and that non-word repetition would be a stronger predictor than digit span (Baddeley et al., 1998). We also hypothesized that (2) these associations would remain significant after allowing for the influence of early receptive vocabulary in addition to non-verbal ability and relevant demographic variables.

MATERIALS AND METHODS

Participants

The Australian Hearing Human Research Ethics Committee approved the protocols used in the current study. Participants in the LOCHI study were included if they continued to use hearing aids by 9 years of age, and completed direct assessments of cognitive and spoken language abilities at 5 and 9 years. Data on measures of 81 participants in the LOCHI study were included in this report. **Table 1** provides descriptive statistics of the demographic characteristics of the current sample.

Procedure

Parents of participants provided written informed consent to the protocol approved by the local institutional human research review board. As part of the LOCHI study, each child was assessed directly by research speech pathologists on norm-referenced tests using standard protocols when they turned 5 and 9 years of age. All data were audited and checked for reliability by double scoring 10% of the evaluations.

Measures

The 5-year-old assessment battery included the PPVT-4 (Dunn and Dunn, 2007), the Memory for Digits (MD) subtest of the Comprehensive Test of Phonological Processing (CTOPP;

TABLE 1 | Demographic characteristics of study participants.

Variable	Number of participants (%)	
Gender (male)	37	(45.7)
Degree of hearing loss (4FA HL)		
Mild ($\leq$40 dB)	29	(35.8)
Moderate (41–60 dB)	40	(49.4)
Severe (61–80 dB)	12	(14.8)
Maternal education (n = 144)		
1. University qualification	34	(42.0)
2. Diploma or certificate	22	(27.2)
3. 12 years or less of schooling	25	(30.9)
Age at first fitting of hearing aids (months)		
Mean (SD)	7.8 (8.2)	
Median	3.8	
75th percentile	10.1	

4FA HL, the average of hearing threshold levels at 0.5, 1, 2, and 4 KHz in the better ear.

Wagner et al., 1999), the non-word repetition test (NRT) (Dollaghan and Campbell, 1998), and the Wechsler Non-verbal Scale of Ability (WNV; Wechsler and Naglieri, 2006). The 9-year-old assessment battery included the CELF-4 (Semel et al., 2003), the PPVT-4, and the Expressive Vocabulary Test (EVT; Williams, 2007).

The PPVT-4 is a standardized test of receptive vocabulary, using a four-alternative-forced choice, picture-pointing format in administration. It gives an overall score on receptive vocabulary.

The EVT is a standardized test of expressive vocabulary. It gives an overall score on expressive vocabulary.

The MD subtest of the CTOPP is a standardized test of capacity of the phonological loop. Recorded digits are presented at a rate of two per second, and forward-only recall is measured. It gives an overall score on phonological short-term memory.

The NRT is another measure of the phonological loop (Gathercole and Baddeley, 1990) in short-term memory. Recorded non-words are presented at a comfortable listening level and the participant is required to repeat back the non-words heard. Responses are recorded and transcribed phonetically for scoring of the number of vowels and consonants correctly repeated. It gives an overall score on phonological short-term memory in terms of phoneme correct score.

The WNV is a standardized test of non-verbal cognitive ability. It gives a full-scale IQ score.

The CELF is a standardized test of spoken English. The test includes verbal tasks which enable children to demonstrate understanding of and ability to produce English language structures. It gives an overall core language score, and two subtest scores – receptive language and expressive language. It also gives a language memory score.

Parents were requested to complete a custom-designed questionnaire to provide demographic information. Audiological information was retrieved from individual clinical files, with permission from parents. All hearing level information and hearing device information were current within 6 months of the evaluation, and at a time closest to the actual evaluation date for each child.

Statistical Analyses

Descriptive statistics were used to report quantitative outcomes for each measure. To examine the relations between early measures of language and working memory at 5 years of age and later language outcomes, correlational analyses were carried out. To determine whether any relations found between early working memory remained after accounting for the effects of early language abilities and other demographic variables and non-verbal intelligence, multiple linear regression analyses were conducted. Two models were fitted with the 9-year CELF Receptive Language and Expressive Language standard scores as dependent variables with repeated measures. In the first model, age at hearing aid fitting, maternal educational level (three categories: university vs. certificate or diploma vs. schooling of 12 years or less), degree of hearing loss [averaged hearing levels at 0.5, 1, 2, and 4 kHz, 4FA in decibel hearing level (4FA dB HL)], and 5-year-old standard scores for WNV, MD, and NRT were used as independent variables to predict 9-year-old language outcomes. The second model included all predictor variables together with 5-year-old PPVT scores to examine the effects of early cognitive measures after allowing for the effect of early language ability. To investigate the relationship between early measures and later vocabulary outcomes, two separate models were fitted in the same manner, but using 9-year-old PPVT scores and EVT scores as dependent variables. Statistical significance was set at the 0.05 level.

RESULTS

Demographic characteristics of participants in this study are reported in **Table 1**.

Descriptive statistics for the scores of each of the outcome measures are shown in **Table 2**. The mean scores on the PPVT-4, EVT, and the CELF-4 measures of receptive and expressive language were within 1 SD (15) of the norm-referenced mean score of 100. The mean scores on the CTOPP Memories for Digit test is within 1 SD (3) of the norm-referenced mean score of 10.

Correlations between language scores and non-verbal IQ at 9 years of age and characteristics at 5 years of age are shown in **Table 3**. Early receptive vocabulary scores, non-verbal IQ, digit span, and non-word repetition scores were significantly correlated with receptive and expressive language and vocabulary scores at 9 years of age. There were no significant associations between language performance at 9 years of age and age at first fitting of hearing aids, or degree of hearing loss of the children. There was no significant relation between maternal education at 5 years of age and non-verbal cognitive ability at 9 years of age.

Regression models predicting language and vocabulary at 9 years of age are shown in **Table 4**. Non-verbal cognitive ability accounted for significant variance in language and vocabulary abilities at 9 years of age. Phonological memory as measured by a digit span test accounted for significant variance in 9-year-old language scores, but not vocabulary scores. The NRT score was a significant predictor of expressive language and vocabulary, but not receptive language and vocabulary at 9 years of age. For both 9-year-old language and vocabulary measures,

TABLE 2 | Descriptive statistics for language and cognitive measures.

| | Measures at 5 years of age | | Measures at 9 years of age | |
	Mean (SD)	Range	Mean (SD)	Range
Y5PPVT	93.7 (14.7)	58–134		
Y5WNV	106.6 (14.3)	58–132		
Y5MD	9.1 (3.1)	1–16		
Y5NRT	57.4 (13.3)	22–86		
Y9RecLg			86.6 (14.9)	50–112
Y9ExpLg			88.8 (18.5)	49–118
Y9PPVT			94.0 (14.7)	61–134
Y9EVT			95.1 (12.2)	69–125
Y9WNV			101.1 (16.9)	43–132

Y5PPVT, PPVT-4 Receptive Vocabulary standard score at 5 years of age; Y5WNV, Wechsler Non-Verbal Full Scale IQ at 5 years of age; Y5MD, CTOPP Memory for digits subtest standard score at 5 years of age; Y5NRT, non-word repetition test phoneme correct score at 5 years of age; Y9RecLg, CELF Receptive Language standard score at 9 years of age (n = 78); Y9ExpLg, CELF Expressive Language standard score at 9 years of age (n = 78); Y9PPVT, PPVT-4 Receptive Vocabulary standard score at 9 years of age; Y9EVT, EVT Expressive Vocabulary standard score at 9 years of age (n = 80), and Y9WNV, Wechsler Non-Verbal Full Scale IQ at 9 years of age (n = 78).

TABLE 3 | Correlations (Pearson's r) between demographic and early predictors and long-term language outcomes.

| | 9-Year-old outcomes | | | | |
	Y9RecLg (n = 78)	Y9ExpLg (n = 78)	Y9PPVT (n = 80)	Y9EVT (n = 80)	Y9WNV (n = 78)
AgeHA	0.04	−0.07	0.07	−0.01	0.05
BE4FA	−0.14	−0.08	−0.09	−0.13	0.01
Y5MEdn	−0.38*	−0.42**	−0.36*	−0.43**	0.15
Y5PPVT	0.62**	0.72**	0.72**	0.72**	0.31*
Y5WNV	0.59**	0.55**	0.43**	0.56**	0.65**
Y5MD	0.54**	0.58**	0.42**	0.39*	0.31*
Y5NRT	0.37*	0.51**	0.37*	0.43**	0.21

*AgeHA, Age at Hearing Aid Fitting; BE4FA, Four Frequency Average Hearing Loss in the better ear; Y5MatEd, Maternal Education (1 = university; 2 = diploma/certificate; 3 = ≤12 years formal schooling) at 5 years of age; Y5PPVT, PPVT-4 Receptive Vocabulary standard score at 5 years of age; Y5WNV, Wechsler Non-Verbal Full Scale IQ at 5 years of age; Y5MD, CTOPP Memory for digits subtest standard score at 5 years of age; Y5NRT, non-word repetition test phoneme correct score at 5 years of age; Y9RecLg, CELF Receptive Language standard score at 9 years of age; Y9ExpLg, CELF Expressive Language standard score at 9 years of age; Y9PPVT, PPVT-4 Receptive Vocabulary standard score at 9 years of age; Y9EVT, EVT Expressive Vocabulary standard score at 9 years of age; and Y9WNV, Wechsler Non-Verbal Full Scale IQ at 9 years of age. *p < 0.01, **p < 0.001.*

adding early receptive vocabulary score at 5 years of age resulted in a significant increase in the variance accounted for by the models. After adding the early receptive vocabulary score as a predictor variable, non-word repetition score was no longer a significant predictor of expressive language or vocabulary at 9 years of age. In summary, the full models incorporating early cognitive ability and language measured at 5 years of age together with demographic characteristics accounted for 61% variance in receptive language and 68% in expressive language scores at 9 years of age. Significant predictors included non-verbal IQ at 5 years, phonological short-term memory measured by a digit

TABLE 4 | Multiple regression summary table showing unstandardized coefficient estimates (*b*-values) and significance levels (*p*-values) of predictor variables for outcomes of children with hearing aids (*p*-values in parentheses).

Predictor	9-Year-old language				9-Year-old vocabulary			
	Receptive (Y9RecLg)		Expressive (Y9ExpLg)		Receptive (Y9PPVT)		Expressive (Y9EVT)	
(ln)AgeHA[a]	0.01 (0.89)	0.05 (0.52)	−0.06 (0.42)	−0.02 (0.77)	0.07 (0.48)	0.13 (0.09)	−0.01 (0.87)	0.04 (0.59)
BE4FA	−0.16 (0.05)	−0.13 (0.08)	−0.12 (0.13)	−0.08 (0.22)	−0.06 (0.54)	−0.03 (0.70)	−0.12 (0.17)	−0.10 (0.18)
Y5WNV	**0.43 (<0.001)**	**0.35 (<0.001)**	**0.38 (<0.001)**	**0.28 (<0.001)**	**0.31 (0.004)**	0.17 (0.05)	**0.46 (<0.001)**	**0.33 (<0.001)**
Y5MD	**0.37 (<0.001)**	**0.31 (0.001)**	**0.34 (0.001)**	**0.27 (0.002)**	0.20 (0.12)	0.09 (0.36)	0.05 (0.62)	−0.04 (0.68)
Y5NRT	0.12 (0.20)	0.01 (0.92)	**0.25 (0.008)**	0.11 (0.16)	0.16 (0.19)	−0.02 (0.85)	**0.24 (0.02)**	0.10 (0.27)
Y5MatEdn[b]	−0.003, 0.14 (0.97)	−0.04, 0.14 (0.70)	0.05, 0.07 (0.64)	0.01, 0.08 (0.91)	0.12, 0.03 (0.37)	0.08, 0.03 (0.45)	0.17, 0.03 (0.14)	0.14, 0.03 (0.14)
Y5PPVT		**0.36 (<0.001)**		**0.43 (<0.001)**		**0.62 (<0.001)**		**0.53 (<0.001)**
Sample size	78	78	78	78	80	80	80	80
Adjusted R^2	0.52	0.61	0.55	0.68	0.26	0.55	0.42	0.63

Bolded entries indicate significance at <0.05 level. AgeHA, Age at Hearing Aid Fitting; BE4FA, Four Frequency Average Hearing Loss in the better ear; Y5WNV, Wechsler Non-Verbal Full Scale IQ at 5 years of age; Y5MD, CTOPP Memory for digits subtest standard score at 5 years of age; Y5NRT, non-word repetition test phoneme correct score at 5 years of age; Y5MatEd, Maternal Education (1 = university; 2 = diploma/certificate; 3 = ≤12 years formal schooling) at 5 years of age; Y5PPVT, PPVT-4 Receptive Vocabulary standard score at 5 years of age; Y9RecLg, CELF Receptive Language standard score at 9 years of age; Y9ExpLg, CELF Expressive Language standard score at 9 years of age; Y9PPVT, PPVT-4 Receptive Vocabulary standard score at 9 years of age; Y9EVT, EVT Expressive Vocabulary standard score at 9 years of age. [a](ln) indicates that AgeHA was transformed using the natural logarithm. [b]For maternal education, the first coefficient estimate is for diploma relative to school, and the second coefficient estimate is for university relative to school. The p-value for MatEd is for the overall test of MatEd.

span test at 5 years, and receptive vocabulary at 5 years of age. The full models also accounted for 55% in receptive vocabulary and 63% in expressive vocabulary at 9 years of age. The only significant predictor for 9-year-old receptive vocabulary was 5-year-old receptive vocabulary. Significant predictors for 9-year-old expressive vocabulary included non-verbal IQ and receptive vocabulary measured at 5 years of age.

DISCUSSION

This study reports findings that extend current knowledge, focusing on the early cognitive predictors of later language abilities in a prospective cohort of children who received early intervention for mild to severe hearing loss using HAs. On average, children's receptive and expressive language scores were around 1 SD below the normative mean, whereas vocabulary scores were within −0.5 SD of the mean.

To address the first research question of whether the capacity of the phonological loop assessed at 5 years of age predicted 9-year-old language outcomes, we found that higher digit span and non-word repetition scores at 5 years of age were significantly associated with better language and vocabulary skills measured at 9 years of age. Of the early cognitive predictors, 5-year-old digit span score was a significant predictor of receptive and expressive language but not receptive or expressive vocabulary. Further, 5-year-old NRT score was a significant predictor of only expressive language and vocabulary, but not receptive language or vocabulary. We also found that higher maternal education and higher non-verbal ability were significantly associated with higher language and vocabulary scores at 9 years of age. However, the regression analyses revealed that only non-verbal ability accounted for unique variance. The failure to find a unique contribution of maternal education is probably not due to its association with non-verbal IQ because the correlation between them is not significant (*p* > 0.05; **Table 3**). Regression analyses

revealed that non-verbal IQ measured at 5 years of age was a significant predictor of 9-year-old language and vocabulary.

To address the second research question of whether 5-year-old phonological short-term memory predicted 9-year-old language outcomes after controlling for 5-year-old receptive vocabulary, we found that 5-year-old digit span score was a significant predictor of 9-year-old expressive and receptive language score, after allowing for the effects of non-verbal IQ and 5-year-old receptive vocabulary. Non-word repetition was no longer a significant predictor of expressive language or expressive vocabulary after allowing for the effects of non-verbal IQ and 5-year-old receptive vocabulary.

The significant association between phonological short-term memory, measured using forward digit span, and 9-year-old language is consistent with findings reported for cochlear implant users (Kronenberger et al., 2014), and for a combined group of hearing aid and cochlear implant users (Figueras et al., 2008). Whereas these previous studies involved concurrent assessments and used composite measures of verbal working memory and language in analyses, the current study used a longitudinal design and showed that early digit span forward, rather than non-word repetition, significantly predicted later expressive and receptive language skills but not vocabulary skills. Importantly, digit span forward was a significant predictor after allowing for the effect of early receptive vocabulary score. The predictive relationship between early cognitive measures and later language abilities is also consistent with findings in Pisoni et al. (2011) longitudinal study of children and adolescents using cochlear implants.

These findings are also broadly consistent with the theoretical framework described by Baddeley et al. (1998) and Baddeley and Hitch (2019). However, whereas Baddeley et al. (1998) suggested that non-word repetition may be a more sensitive measure of the capacity of the phonological loop than digit span, the current data suggest a stronger role for digit span in predicting later language outcomes. Furthermore, by contrast

with Baddeley et al.'s (1998) focus on the particular role of the phonological loop in vocabulary acquisition, the current findings provide stronger evidence that early digit span is associated with later language ability considered more broadly (as assessed using the CELF-4). Despite these relatively minor departures from specific theoretical expectations, we interpret our findings as generally consistent with the working memory model and with the assertion that, for children with mild to severe hearing loss who use hearing aids and communicate using spoken language, the capacity of the phonological loop at 5 years of age appears to play an important role in language development.

The question remains: why was digit span forward a stronger predictor of later language abilities than the NRT score? Non-word repetition is a complex task requiring a child to identify a novel string of heard phonemes, retain this string in phonological short-term memory, and produce the same sequence as speech output. Although both digit span forward and non-word repetition tasks require a child to plan and execute the sequence of articulatory gestures to yield a phonological output that corresponds to a retrieved memory representation, articulatory accuracy has a potentially greater effect on the non-word repetition score than a digit span score because a single phoneme deviation is scored as an error in the former but not the latter (Gathercole and Baddeley, 1996). Many children with hearing loss have impoverished phonological representations as a consequence of their auditory deficits and the distorted signals received through hearing devices (e.g., Kronenberger et al., 2014), and may have phonological production systems that will never be fully accurate. As such, performance in the non-word repetition task may be limited by children's speech output skills (e.g., Snowling and Hulme, 1989). That said, it is important to note that Avons et al. (1998) also reported a stronger association between early digit span and later vocabulary scores than between non-word repetition and later vocabulary scores in a sample of normally hearing children assessed at 5 and 6 years of age. Furthermore, in that study, articulation rate was not a significant predictor.

A striking aspect of the findings reported here is the strong and significant association between early receptive vocabulary and later expressive and receptive language and vocabulary. When added to the regression models for the four 9-year-old language and vocabulary measures, 5-year-old receptive vocabulary accounted for significant unique variance ranging from 9 to 29% after controlling for all other predictor variables. This result is consistent with findings reported for users of cochlear implants (e.g., Castellanos et al., 2016a,b; Nittrouer et al., 2016; Hunter et al., 2017). It also provides support for the proposed role of language knowledge in supporting further language development, and thereby potentially reducing reliance on the phonological loop as the primary language learning device (Baddeley et al., 1998). Some evidence for a reduction in the role of phonological short-term memory in predicting later language ability comes from our findings for non-word repetition, which was a significant predictor of expressive language and vocabulary after controlling for digit span, non-verbal ability, and relevant demographic variables, but became non-significant with the addition of receptive vocabulary as a predictor.

The observed relation between early cognitive abilities and later expressive and receptive language and vocabulary development implies that early forward digit span may be used as a screening tool for audiologists to identify children who may be at risk of later language difficulties, so that the children can be referred for professional assessment and treatment. The significant association between 5-year-old receptive vocabulary and 9-year-old language abilities supports the use of early language assessments to inform what the targets of language intervention should be. Despite early intervention for the cohort of children with hearing loss reported in this study, **Table 2** shows that the mean language scores were within 1 SD below the mean of the normal population, suggesting that some children with hearing loss exhibit language deficits at school age that is potentially avoidable if a digit span test could be used as an early screener to expedite referral for language assessment and intervention. This offers the opportunity to capitalize on the benefits due to early detection and treatment of hearing loss (Ching et al., 2017) by optimizing habitation strategies, including considerations for increase in intensity of early language intervention and considerations for alternative hearing devices, such as cochlear implantation, for those who may be in need.

The findings reported in this study are drawn from a cohort of children born with hearing loss who used spoken language as the primary mode of communication and early education. Therefore, these findings should not be generalized to children who communicate using sign language.

Age at hearing aid fitting did not reach significance level in the regression analyses. This finding may be explained in terms of the restricted range of age at hearing aid fitting for the cohort in this report, and the significant association between age at fitting and 5-year-old language outcomes. The current results were drawn from a sub-sample of the LOCHI cohort who use hearing aids, who completed all the spoken language and cognitive measures at 5 years and 9 years of age, and who used speech to communicate. The sample received very early fitting of hearing aids (median: 3.8 months, upper quartile: 10.1 months). In an earlier report (Ching et al., 2017), we showed that earlier fitting was significantly associated with better language at 5 years of age. In the current investigation, 5-year-old receptive vocabulary and cognitive abilities were included as predictors in the regression analyses. Therefore, it is not surprising that after allowing for the effects of early language and cognitive abilities, age at fitting does not account for unique variance in 9-year-old language and vocabulary outcomes.

Future investigations on predictors of language outcomes of children using hearing aids at 9 years will need to include a wider range of factors reported in the literature (e.g., Tomblin et al., 2015) than was used in this study. For example, we did not include hearing device use or aided audibility in this study partly because the current focus is on the role of early cognitive factors on later language development; and partly because our earlier reports on the LOCHI cohort at 3 and 5 years of age showed that these factors did not account for unique variance after allowing for the effects of a range of child- and family-related factors (Ching et al., 2013, 2018a,b). Even though the question of the link between cognitive abilities and language development is not new,

the current study is the first to investigate this relationship in a prospective study of a population-based cohort of children with mild to severe hearing loss using hearing aids.

We conclude that early phonological short-term memory assessed using a digit span task significantly predicted later expressive and receptive language abilities, even after allowing for the effect of early receptive vocabulary. Future studies will examine the importance of phonological working memory in speech perception in children with hearing loss.

DATA AVAILABILITY STATEMENT

The datasets for this study will not be made publicly available because the ethics approval for this study does not include the dissemination of the dataset.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the ethics guidelines of the Australian Hearing Human Research Ethics Committee with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Australian Hearing Human Research Ethics Committee.

AUTHOR CONTRIBUTIONS

TC conceptualized and designed the study, took overall responsibility for all aspects of the study, and drafted the manuscript. LC consulted in the design of the study, and drafted and reviewed the manuscript. VM coordinated the acquisition of data and collation of data, and reviewed the manuscript. TC, LC, and VM approved the final manuscript as submitted.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Avons, S. E., Wragg, C. A., Cupples, L., and Lovegrove, W. J. (1998). Measures of phonological short-term memory and their relationship to vocabulary development. *Appl. Psychol.* 19, 583–601. doi: 10.1017/S0142716400010377

Baddeley, A., Gathercole, S., and Papagno, C. (1998). The phonological loop as a language learning device. *Psychol. Rev.* 105, 158–173. doi: 10.1037/0033-295X.105.1.158

Baddeley, A., and Hitch, G. J. (2019). The phonological loop as a buffer store: an update. *Cortex* 112, 91–106. doi: 10.1016/j.cortex.2018.05.015

Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., and Morgan, G. (2017). Nonverbal executive function is mediated by language: a study of deaf and hearing children. *Child Dev.* 88, 1689–1700. doi: 10.1111/cdev.12659

Castellanos, I., Kronenberger, W. G., Beer, J., Henning, S. C., Colson, B. G., and Pisoni, D. B. (2016a). Preschool speech intelligibility and vocabulary skills predict long-term speech and language outcomes following cochlear implantation in early childhood. *Cochlear Implants Int.* 15, 200–210. doi: 10.1179/1754762813Y.0000000043

Castellanos, I., Pisoni, D. B., Kronenberger, W. G., and Beer, J. (2016b). Early expressive language skills predict long-term neurocognitive outcomes in cochlear implant users: evidence from the macarthur–bates communicative development inventories. *Am. J. Speech Lang. Pathol.* 25, 381–392. doi: 10.1044/2016_AJSLP-15-0023

Ching, T. Y. C., Dillon, H., Marnane, V., Hou, S., Day, J., Seeto, M., et al. (2013). Outcomes of early- and late-identified children at 3 years: findings from a population study. *Ear. Hear.* 34, 535–552. doi: 10.1097/AUD.0b013e3182857718

Ching, T. Y. C., Dillon, H., Button, L., Seeto, M., Van Buynder, P., Marnane, V., et al. (2017). Age at intervention for permanent hearing loss and 5-year language outcomes. *Pediatrics* 140:e20164274. doi: 10.1542/peds.2016-4274

Ching, T. Y. C., Zhang, V. W., Flynn, C., Burns, L., Button, L., Hou, S., et al. (2018a). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57(Suppl. 2), S70–S80.

Ching, T. Y. C., Zhang, V. W., Johnson, E. E., Van Buynder, P., Hou, S., Burns, L., et al. (2018b). Hearing aid fitting and developmental outcomes of children fit according to either the NAL or DSL prescription: fit-to-target, audibility, speech and language abilities. *Int. J. Audiol.* 57(Suppl. 2), S41–S54. doi: 10.1080/14992027.2017.1380851

Conway, C. M., Karpicke, J., Anaya, E. M., Henning, S. C., Kronenberger, W. G., and Pisoni, D. B. (2011a). Nonverbal cognition in deaf children following cochlear implantation: motor sequencing disturbances mediate language delays. *Dev. Neuropsychol.* 36, 237–254. doi: 10.1080/87565641.2010.549869

Conway, C. M., Pisoni, D. B., Anaya, E. M., Karpicke, J., and Henning, S. C. (2011b). Implicit sequence learning in deaf children with cochlear implants. *Dev. Sci.* 14, 69–82. doi: 10.1111/j.1467-7687.2010.00960.x

Conway, C. M., Pisoni, D. B., and Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: the auditory scaffolding hypothesis. *Curr. Direct Psychol. Sci.* 18, 275–279. doi: 10.1111/j.1467-8721.2009.01651.x

Cupples, L., Ching, T. Y. C., Button, L., Seeto, M., Zhang, V., Whitfield, J., et al. (2018). Spoken language and everyday functioning in 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57, S55–S69. doi: 10.1080/14992027.2017.1370140

Dunn, L. M., and Dunn, D. M. (2007). *Peabody Picture Vocabulary Test,* 4th Edn, Bloomington, MN: Pearson.

Dollaghan, C., and Campbell, T. F. (1998). Nonword repetition and child language impairment. *J. Speech Lang. Hear. Res.* 41, 1136–1146. doi: 10.1044/jslhr.4105. 1136

Figueras, B., Edwards, L., and Langdon, D. (2008). Executive function and language in deaf children. *J. Deaf Stud. Deaf Educ.* 13, 362–377. doi: 10.1093/deafed/ enm067

Gathercole, S. E., and Baddeley, A. D. (1990). The role of phonological memory in vocabulary acquisition: a study of young children learning new names. *Br. J. Psychol.* 81, 439–454. doi: 10.1111/j.2044-8295.1990.tb02371.x

Gathercole, S. E., and Baddeley, A. D. (1996). *The children's test of nonword repetition*. London: Psychological Corporation.

Gathercole, S. E., Willis, C. S., Emslie, H., and Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: a longitudinal study. *Dev. Psychol.* 28, 887–898. doi: 10.1037/0012-1649.28.5.887

Giustolisi, B., and Emmorey, K. (2018). Visual statistical learning with stimuli presented sequentially across space and time in deaf and hearing adults. *Cogn. Sci.* 42, 3177–3190. doi: 10.1111/cogs.12691

Hall, M. L., Eigsti, I.-M., Bortfeld, H., and Lillo-Martin, D. (2018). Auditory access, language access, and implicit sequence learning in deaf children. *Dev. Sci.* 21:e12575. doi: 10.1111/desc.12575

Hunter, C. R., Kronenberger, W. G., Castellanos, I., and Pisoni, D. B. (2017). Early postimplant speech perception and language skills predict long-term language and neurocognitive outcomes following pediatric cochlear implantation. *J. Speech Lang. Hear. Res.* 60, 2321–2336. doi: 10.1044/2017_JSLHR-H-16-0152

Jones, A., Atkinson, J., Marshall, C., Botting, N., St Clair, M. C., and Morgan, G. (2019). Expressive vocabulary predicts nonverbal executive function: a 2-year longitudinal study of deaf and hearing children. *Child Dev.* doi: 10.1111/cdev. 13226 [Epub ahead of print].

Kronenberger, W. G., Colson, B. G., Henning, S. C., and Pisoni, D. B. (2014). Executive functioning and speech-language skills following long-term use of cochlear implants. *J. Deaf Stud. Deaf Edu.* 19, 456–470. doi: 10.1093/deafed/ enu011

Nittrouer, S., Lowenstein, J. H., and Holloman, C. (2016). Early predictors of phonological and morphosyntactic skills in second graders with cochlear implants. *Res. Dev. Disabil.* 55, 143–160. doi: 10.1016/j.ridd.2016.03.020

Pisoni, D. B., Kronenberger, W. G., Chandramouli, S. H., and Conway, C. M. (2016). Learning and memory processes following cochlear implantation: the missing piece of the puzzle. *Front. Psychol.* 7:493. doi: 10.3389/fpsyg.2016. 00493

Pisoni, D. B., Kronenberger, W. G., Roman, A. S., and Geers, A. E. (2011). Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation. *Ear. Hear.* 32, 60S–74S. doi: 10.1097/AUD. 0b013e3181ffd58e

Semel, E., Wiig, E. H., and Secord, E. A. (2003). *Clinical Evaluation of Language Fundamentals*, 4th Edn. San Antonio, TX: PsychCorp.

Snowling, M., and Hulme, C. (1989). A longitudinal case study of developmental phonological dyslexia. *Cogn. Neuropsychol.* 6, 379–401. doi: 10.1080/02643298908253289

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., and Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss. *Ear. Hear.* 36, 76S–91S. doi: 10.1097/aud.00000000000 00219

von Koss Torkildsen, J., Arciuli, J., Haukedal, C. L., and Wie, O. B. (2018). Does a lack of auditory experience affect sequential learning? *Cognition* 170, 123–129. doi: 10.1016/j.cognition.2017.09.017

Wagner, K. R., Torgsesen, K. J., and Rashotte, A. C. (1999). *CTOPP: Comprehensive Test of Phonological Processing*. Austin, TX: Pro-Ed.

Wechsler, D., and Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. San Antonio, TX: Harcourt Assessment.

Williams, K. T. (2007). *Expressive Vocabulary Test*, 2nd Edn. Minneapolis, MN: Pearson Assessment.

Acoustics of Emotional Prosody Produced by Prelingually Deaf Children With Cochlear Implants

Monita Chatterjee*, Aditya M. Kulkarni, Rizwan M. Siddiqui, Julie A. Christensen[†], Mohsen Hozan[†], Jenni L. Sis[†] and Sara A. Damm[†]

Auditory Prostheses and Perception Laboratory, Center for Hearing Research, Boys Town National Research Hospital, Omaha, NE, United States

Edited by:
Mary Rudner,
Linköping University, Sweden

Reviewed by:
Frank A. Russo,
Ryerson University, Canada
Timothy Beechey,
University of Minnesota Twin Cities,
United States

*Correspondence:
Monita Chatterjee
monita.chatterjee@boystown.org

†Present address:
Julie A. Christensen,
Otolaryngology Branch,
National Institute on Deafness and
Other Communicative Disorders,
National Institutes of Health,
Bethesda, MD, United States
Mohsen Hozan,
Department of Special Education and
Communication Disorders, Barkley
Memorial Center, University of
Nebraska-Lincoln, Lincoln,
NE, United States;
Biological Systems Engineering
Department, University of Nebraska-
Lincoln, Lincoln, NE, United States
Jenni L. Sis,
Department of Special Education and
Communication Disorders, Barkley
Memorial Center, University of
Nebraska-Lincoln, Lincoln, NE,
United States;
Biological Systems Engineering
Department, University of Nebraska-
Lincoln, Lincoln, NE, United States
Sara A. Damm,
Omaha Public Schools,
Omaha, NE, United States

Purpose: Cochlear implants (CIs) provide reasonable levels of speech recognition quietly, but voice pitch perception is severely impaired in CI users. The central question addressed here relates to how access to acoustic input pre-implantation influences vocal emotion production by individuals with CIs. The objective of this study was to compare acoustic characteristics of vocal emotions produced by prelingually deaf school-aged children with cochlear implants (CCIs) who were implanted at the age of 2 and had no usable hearing before implantation with those produced by children with normal hearing (CNH), adults with normal hearing (ANH), and postlingually deaf adults with cochlear implants (ACI) who developed with good access to acoustic information prior to losing their hearing and receiving a CI.

Method: A set of 20 sentences without lexically based emotional information was recorded by 13 CCI, 9 CNH, 9 ANH, and 10 ACI, each with a happy emotion and a sad emotion, without training or guidance. The sentences were analyzed for primary acoustic characteristics of the productions.

Results: Significant effects of Emotion were observed in all acoustic features analyzed (mean voice pitch, standard deviation of voice pitch, intensity, duration, and spectral centroid). ACI and ANH did not differ in any of the analyses. Of the four groups, CCI produced the smallest acoustic contrasts between the emotions in voice pitch and emotions in its standard deviation. Effects of developmental age (highly correlated with the duration of device experience) and age at implantation (moderately correlated with duration of device experience) were observed, and interactions with the children's sex were also observed.

Conclusion: Although prelingually deaf CCI and postlingually deaf ACI are listening to similar degraded speech and show similar deficits in vocal emotion perception, these groups are distinct in their productions of contrastive vocal emotions. The results underscore the importance of access to acoustic hearing in early childhood for the production of speech prosody and also suggest the need for a greater role of speech therapy in this area.

Keywords: acoustics, emotion, vocal, production, speech, cochlear implants, children

INTRODUCTION

Emotional communication is a key element of social development, social cognition, and emotional well-being. Studies have shown that in children and adults with cochlear implants (CIs), performance in vocal emotion recognition tasks predicts their self-perceived quality of life, but their general speech recognition does not (Schorr et al., 2009; Luo et al., 2018). This indicates that speech emotion communication is a critical area of deficit in CIs that needs to be addressed. Acoustic cues signaling vocal emotions in speech include voice pitch, timbre, intensity, and speaking rate (e.g., Banse and Scherer, 1996). Among these, voice pitch is a dominant cue. CIs do not represent voice pitch to the listener with adequate fidelity, but other cues to vocal emotions, such as intensity and duration cues, are retained in the electric input. These deficits in vocal pitch perception have been implicated in CI users' poorer performance in pitch-dominant areas of speech perception such as prosody or lexical tones (Peng et al., 2004, 2008, 2017; Green et al., 2005; Chatterjee and Peng, 2008; See et al., 2013; Deroche et al., 2016; Jiam et al., 2017). The importance of voice pitch for spoken emotions is thought to account for the deficits observed in cochlear implant users' ability to identify emotional prosody (Luo et al., 2007; Hopyan-Misakyan et al., 2009; Chatterjee et al., 2015; Paquette et al., 2018). The perceptual deficit observed in CI users in emotion identification suggests that on their own, these secondary cues are not sufficient to provide normal levels of accuracy in vocal emotion identification. Similar deficits have been observed in normally hearing listeners attending to CI-simulated speech (Luo et al., 2007; Chatterjee et al., 2015; Gilbers et al., 2015; Tinnemore et al., 2018).

Prelingually deaf children who received a CI (CCI) within the sensitive period (e.g., by 2 years of age) and are developing oral communication skills through the prosthesis provide a unique opportunity to investigate the impact of the perceptual deficits associated with electric hearing on the development of emotional prosody. This population also provides an important contrast to postlingually deaf adult CI users (ACI) who learned to hear and speak with good hearing in childhood before losing their hearing as teenagers or adults, in many cases in middle age or later years. ACI generally retain excellent speech production skills, despite listening through the distorted input of the CI. In a previous study comparing ACI and CCI in their vocal emotion perception, Chatterjee et al. (2015) noted that they were similar in both the mean and the range of performance. Notably, the stimuli used by Chatterjee et al. (2015) were highly recognizable by normally hearing listeners as they were produced in a child-directed manner, with exaggerated prosody. While few studies have reported deficits in prelingually deaf pediatric CI users' productions of vocal emotions (Nakata et al., 2012; Van De Velde et al., 2019), they have focused on younger children (<10 years of age) and used perceptual ratings of the productions as the outcome measure. Little is known about the factors predicting the acoustic features of these productions as children develop into teenagers, and no studies have reported on a comparison between pre and postlingually deaf CI users. Here, we present acoustic analyses of emotional prosody [a set of 20 emotion-neutral sentences (i.e., without lexically based emotional information) read with "happy" and "sad" emotional prosody] produced by prelingually deaf school-aged children and postlingually deaf adults with CIs, alongside productions by typically developing normally hearing children and young normally hearing adults. We selected happy and sad emotions because these are well-contrasted acoustically (happy is spoken with a higher mean pitch, more fluctuating pitch, higher intensity, and faster than sad). These two emotions are also uncontroversial and relatively easy for school-aged children as young as 6 years old to know and be able to produce without an exemplar. Previous studies have used different methodologies, e.g., Nakata et al. (2012) asked children to imitate the vocal productions of an exemplar, while Van De Velde et al. (2019) asked children to produce a word depicted in a picture with an emotion simultaneously depicted in a picture. Imitative production provides information about vocal capabilities but not about how the participants would normally produce emotions. Van De Velde et al.'s (2019) method avoided imitation but may have imposed additional task complexity in the requirement to generate the word associated with the picture and the emotion associated with the picture, combine them conceptually, and produce the word with the correct emotion. In our task, we avoided imitation and kept the cognitive load to a minimum by asking children to read the list of sentences in a happy way and in a sad way. There was still the remaining task burden of having to combine the emotion with the sentence before producing it, but the participants did not have to generate the words themselves or figure out the emotion required for the production.

Among acoustic cues, we focused on mean voice pitch, variance of voice pitch, mean intensity, mean spectral centroid, and mean duration of each utterance. These cues were found to be important acoustic features of vocal emotions in previous studies (Banse and Scherer, 1996; Scherer, 2003). These cues have also been found to be useful in artificial manipulations of speech designed to represent different human emotions (e.g., Přibilová and Přibil, 2009). Based on pitch and spectral degradations in CIs, we expected the CI users (particularly CCI) to show deficits in the pitch and spectral centroid domains of their productions. We expected to observe smaller acoustic contrasts between "happy" and "sad" emotions in the productions of the CCI than in those by children with normal hearing (CNH) and adults with normal hearing (ANH), but we were interested in the specific acoustic cues that might show such reduced contrasts. We expected CNH and ANH to produce the emotions similarly. A key question of interest was how CCI and ACI would compare in their productions. Specifically, we asked if CCI and/or ACI would emphasize intensity or duration differences between the emotions to compensate for any deficits in the pitch domain. Previous studies have shown that adult and child CI users can trade primary acoustic cues for secondary cues such as duration and intensity in speech recognition, intonation recognition, and lexical tone recognition tasks (Peng et al., 2009, 2017; Winn et al., 2012). Luo et al. (2007) showed that removing intensity cues from the stimuli resulted in much poorer emotion recognition scores in their adult CI listeners, indicating that intensity cues are emphasized in vocal emotion recognition by postlingually

def CI users. The extent to which this would influence their vocal emotion productions is not known, nor is it known whether prelingually deaf CCI would emphasize intensity cues in their productions. Among the CCI, we asked if earlier age at implantation or longer duration of experience with the device would change the acoustic characteristics of their productions. These questions center around the role of neuroplasticity within the more sensitive, early years of brain development and during the developmental period of auditory and language systems, which extends into the teenage years.

MATERIALS AND METHODS

Participants

Participants were comprised of four groups of talkers. All talkers provided informed consent to be recorded, and procedures were approved by Boys Town National Research Hospital's IRB protocol #11-24-XP. The four groups of talkers are as described below. Detailed information about the CI users who participated is shown in **Table 1**. The information in **Table 1** was derived from a questionnaire filled out by participants or (in the case of child participants) by their parents/guardians. Written informed assent was obtained from all child participants, together with written informed parental consent to participate; written informed consent was obtained from all adult participants. Participants were compensated for travel time and for their listening time. In addition, children were offered a toy or a book of their choice after they completed their sessions.

Children With Normal Hearing

Nine children with normal hearing participated. Their ages ranged between 6 and 18 years [mean age 12.5 years, standard deviation (SD) 4.4 years]. Five of the children were females, and four were males. All had normal hearing based on audiometric screening at criterion level of 20 dB HL or better between 250 and 8,000 Hz.

Children With Cochlear Implants

Thirteen children with cochlear implants participated. Their ages ranged between 7 and 18 years (mean age 12.93 years, SD 4.27 years). Four of the children were males, and nine were females. All of the CCI were prelingually deaf, implanted at the age of 2, and none had any usable hearing at birth. Their mean age at implantation was 1.36 years (SD 0.35 years), and their mean duration of device use was 11.57 years (SD 4.04 years).

Adults With Normal Hearing

Nine adults with normal hearing participated. Their ages ranged between 21 and 45 years. Six of the ANH were females; three were males. As with the CNH, normal hearing was confirmed based on audiometric screening at criterion level of 20 dB HL or better between 250 and 8,000 Hz.

Adults With Cochlear Implants

Ten postlingually deaf adults with cochlear implants participated. Their ages ranged between 27 and 75 years. Six of the ACI were females; four were males.

TABLE 1 | Information about CI participants.

	Age at testing (years)	Age of implantation (years)	Duration of CI use (years)	Bilateral implant (yes/no)	Sex	Manufacturer/ device
Prelingually deaf CCI group						
Child CI participant						
CICH02	18.14	2	16.14	No	Male	Cochlear
CICH03	11.89	1.4	10.49	No	Female	Advanced Bionics
CICH13	7.72	0.83	6.89	Yes	Female	Advanced Bionics
CICH18	17.2	1.7	15.5	No	Female	Advanced Bionics
CICH19	7	0.9	6.1	No	Female	Advanced Bionics
CICH20	7.6	1.1	6.5	Yes	Male	Advanced Bionics
CICH22	12.62	1.86	10.76	Yes	Female	Advanced Bionics
CICH35	12.73	1	11.73	Yes	Male	Advanced Bionics
CICH36	16.27	1.5	14.77	Yes	Female	Med-El
CICH37	18.49	1.5	16.99	Yes	Female	Advanced Bionics
CICH38	7.9	1.25	6.65	Yes	Female	Cochlear
CICH39	16.61	1.17	15.44	Yes	Female	Advanced Bionics
CICH40	14.025	1.5	12.53	Yes	Male	Advanced Bionics
Postlingually deaf ACI group						
Adult CI participant						
C01	37	31	6	Yes	Female	Advanced Bionics
C03	67	55	12	No	Male	Advanced Bionics
C05	68	63	5	No	Female	Advanced Bionics
C06	75	55	20	No	Female	Advanced Bionics
C07	68	67	1	No	Female	Advanced Bionics
N5	53	50	3	No	Female	Cochlear
N6	51	44	7	Yes	Male	Cochlear
N7	57	51	6	No	Female	Cochlear
N15	61	59	2	No	Male	Cochlear
N16	27	25	2	Yes	Male	Cochlear

Procedure

The list of materials used for this study is comprised of 20 simple sentences that had no overt semantic cues about emotion. These sentences are provided in **Table 2** (identical to Table 2 in Damm et al., 2019, JSLHR). The sentences were simple enough that the youngest participants (as young as 6 years of age) could read them aloud easily. The protocol for the recordings was as follows: the participant was invited to sit in a soundproof booth at a distance of 12 inches from a recording microphone (AKG C 2000 B) and asked to read the 20 sentences in sequence, first in a happy way (three times) and then in a sad way (three times). They were provided with some initial practice runs and recordings that were initiated when they felt ready. No targeted training or feedback was provided; all feedbacks were encouraging and laudatory in nature. The signal from the microphone was routed through an external A/D converter (Edirol UA-25X) using Adobe Audition v. 3.0 or v. 6.0. Recordings were made at a sampling rate of 44,100 Hz and with 16-bit resolution. The recordings were high-pass filtered using a 75-Hz cut-off frequency. Of the three sets of recordings in each emotion provided by individual talkers, the second set was typically used for acoustic analyses. For instance in which the second recording of a particular sentence was noisy and included non-speech sounds (such as coughing or throat-clearing), the best sample of the other two recordings was selected. An order effect may be present in the data, as *happy* emotions were recorded prior to *sad*. The recordings took very little time overall, so it is unlikely that fatigue played a role. Based on experience, we noted that it was easier for the participants (particularly, the younger children) to begin the session with

TABLE 2 | List of sentences.

1. Time to go.

2. Here we are.

3. This is it.

4. This is mine.

5. The bus is here.

6. It's my turn.

7. They are here.

8. Today is the day.

9. Time for a bath.

10. She is back.

11. It's snowing again.

12. It's Halloween.

13. Time for bed.

14. Time for lunch.

15. I see a dog.

16. I see a car.

17. I see a cat.

18. That is the book.

19. I saw a bug.

20. That is a big tree.

the *happy* productions and to continue recording in a particular emotion, rather than to switch from *happy* to *sad* during the recordings. Any order effect in the data would be expected to be present for all participants. The CI users who were bilaterally implanted were recorded with their earlier-implanted devices activated only.

Acoustic Analyses

Acoustic analyses were performed on the recordings using the Praat software package (Boersma, 2001; Boersma and Weenink, 2019). For the 40 recordings (20 sentences, 2 emotions each) provided by each participant, a Praat script was run to compute the mean pitch (F0, Hz), the F0 variation (standard deviation of F0), the mean intensity (dB), and the duration (sec) of each utterance. The default autocorrelation method in the Praat software program was used to estimate F0. Primary challenges in such analyses are encountered by researchers attempting to determine the onset and offset of the utterances in a consistent way and in setting parameters for pitch estimation appropriately for each utterance. The onset and offset times of each waveform were estimated using similar criteria by at least two of the co-authors so as to obtain consistent measures of duration. The pitch settings were established using the following steps: for each talker and emotion, a set of 4–5 recordings (from the total of 20) was pseudo randomly selected, and the pitch range, silence threshold, voicing threshold, octave cost, octave-jump cost, and voiced/unvoiced cost were set to appropriate levels, ensuring that the pitch contour was properly represented (e.g., avoiding octave jumps, discontinuities in the estimated pitch, or silences in regions of voiced speech). This was done more than once to ensure that the settings were indeed appropriate. Next, an automated Praat script was run on all the 20 recordings for that talker and emotion. The output was then analyzed for consistency (e.g., mean F0 values were compared across the recordings, and the ratio of maximum to minimum F0 for individual recordings was investigated). If these values appeared suspect for any of the recordings (e.g., if the ratio of maximum to minimum F0 values exceeded a value of 3.0 or if the estimated values were obviously different from other recordings by the same talker in the same emotion), they were individually checked again, modifications were made as needed to the settings, and the values were manually computed in Praat for those individual recordings. Two of the authors (RS and MC) were always involved in the final analyses. Some of the analyses of productions by the children had been previously conducted (by authors MC and JS) using a similar but not identical approach. Care was taken to compare these older analyses with the newer ones. When correlations between the two sets of data fell below 0.85, the analyses were again checked to ensure accuracy and modified again as needed.

Spectral centroid analyses were conducted in R using the seewave package (Sueur et al., 2008; Sueur, 2018). A window was first applied to discard the first 10% and last 10% of each waveform, with a bandpass filter with cut-off frequencies at [50, 4,000] Hz to narrow the range of the calculated centroid to speech content. Next, using the meanspec() function, the

short-term Fourier transform (STDFT) of 50-ms long successive time segments (Hann-windowed with 50% overlap) of the waveform was computed and averaged across all segments to obtain the mean spectrum. Finally, using the specprop() function, the spectral centroid of each waveform was computed for its mean spectrum, based on the formula $C = \sum_{i=1}^{N} f_i \times a_i$, where N is the number of frequency bins (STDFT columns), f_i is the center frequency of the ith bin, and a_i is the relative amplitude. Both frequency and amplitude are linearly scaled in the centroid calculations.

Statistical Analyses

Statistical analyses and graphical renderings were conducted using R v. 3.3.2 (R Core Team, 2016). Plots were created using the *ggplot2* package within R (Wickham, 2016). Linear mixed effects models were constructed using the package *lme4* (Bates et al., 2015). A hierarchical approach was used to determine the best-fitting model, and the function *anova()* was used in the *car* package in R to compare models (Fox and Weisberg, 2011). Model residuals were visually inspected (using plots and histograms of residuals) to ensure normality. The *lmerTest* package (Kuznetsova et al., 2017) was used to obtain estimated model results and t-statistic-based significance levels for each parameter of interest. The *optimx* package (Nash and Varadhan, 2011) was used to promote model convergence in one instance.

RESULTS

Group Differences

Figure 1 shows boxplots of the acoustic characteristics of happy and sad emotions produced by each of the four groups of participants. From top to bottom, the different rows show the mean F0, F0 variation [standard deviation of F0 (F0 s.d.)], mean intensity, duration, and spectral centroid of the sentences produced with the two emotions (red and blue). The boxplots in the left-hand panels show the distribution of values computed for each sentence (abscissa) across the participants. The boxplots in the right panel of each row show the mean values computed across the sentences recorded in each emotion.

LME analyses were conducted on these data to investigate effects of Group, Sentence, and Emotion and their interactions. In all cases, the LME model was constructed including Group, Sentence, and Emotion as fixed effects, subject-based random intercepts, and random slopes for the effect of sentence. The dependent variable in each case was the particular acoustic measure under consideration (mean F0, F0 s.d., Intensity, Duration, or Spectral Centroid). The effect of Sentence was included as a fixed effect because systematic differences for individual sentences were expected, based on differences between them in their phonetic and linguistic characteristics.

Mean F0

Results showed a significant interaction between Group and Emotion [$\beta = -17.656$ (SE = 3.879), $t(1599.10) = -4.552$,

$p < 0.0001$] and a significant main effect of Emotion – i.e., higher mean F0 for happy than for sad productions [$\beta = -45.014$ (SE = 6.907), $t(1599.1) = -6.517$, $p < 0.0001$]. No other effects and no other interactions were observed. To follow-up on the interaction, we investigated the effect of Group for the happy and the sad productions separately. LME analyses on the happy productions with fixed effect of Group, by-subject random intercepts, and by-subject random slopes for the effect of individual sentences showed no effect of Group. A similar analysis on the sad productions did show a significant effect of Group [$\beta = -25.96$ (SE = 8.76), $t(41) = -2.96$, $p = 0.005$], explaining the interaction between Group and Emotion. A pairwise t test (Bonferroni correction) to investigate the effect of Group in the sad productions showed no significant differences between the ANH and ACIs' mean F0 values ($p = 0.32$), but all other comparisons showed significant differences ($p < 0.001$ in all cases). Of note, the CCIs' sad productions had the highest mean F0 of the four groups.

F0 Variation

The mean F0 and F0 s.d. values were significantly correlated in all groups. A linear multiple regression analysis confirmed that F0 s.d. was significantly predicted by mean F0 and also showed that there was an interaction with Group (i.e., different correlation coefficients for the different groups). Individual linear regression analyses within the four groups confirmed this observation: estimated coefficients for the ANH, ACI, and CCI groups were 0.266 (SE 0.01), 0.263 (SE 0.012), and 0.259 (SE 0.011), respectively, whereas the coefficient for the CNH group was only 0.162 (SE 0.009).

The LME analysis showed significant effects of Group [$\beta = 9.307$ (SE = 2.661), $t(49.4) = 3.498$, $p = 0.001$], as well as a significant interaction between Group and Emotion [$\beta = -10.44$ (SE = 1.569), $t(1599) = -6.651$, $p < 0.0001$]. No other effects or interactions were observed. Follow-up analyses showed that the effect of Group was significant for the happy emotion [$\beta = 8.759$ (SE = 2.774), $p = 0.003$], but no significant effect of Group was observed for the sad emotion. *Post hoc* pairwise t tests (Bonferroni corrections applied) comparing the F0 s.d. values obtained by the different groups for the happy emotion productions showed significant differences between the CCI group and ACI, ANH, and CNH groups ($p < 0.0001$ in all cases), but no significant differences between the ACI, ANH, and CNH groups. Thus, the CCI group's productions for happy were more monotonous (smaller F0 s.d.) than all other groups.

Mean Intensity

Results showed a significant interaction between Group and Emotion [$\beta = -0.757$ (SE = 0.276), $t(1558) = -2.738$, $p = 0.00625$], a main effect of Emotion [$\beta = -6.041$ (SE = 0.492), $t(1558) = -12.27$, $p < 0.0001$], and a main effect of Sentence [$\beta = -0.0773$ (SE = 0.031), $t(133.9) = 2.526$, $p = 0.0127$].

The interaction between Group and Emotion was not clearly supported by follow-up analyses. When the data were separated out into happy and sad emotions, separate LME analyses with Group as a fixed effect, random subject-based intercepts, and

FIGURE 1 | Continued

random subject-based slopes for the effect of Sentence showed no significant effects of Group for either emotion. However, the estimated effect for Group [$\beta = -1.1224$ (SE = 0.686), $t(41) = -1.635$, $p = 0.11$] was larger for sad productions than for happy productions [$\beta = -0.3044$ (SE = 0.538), $t(41) = 0.566$, $p = 0.574$]. This is likely explained by the somewhat lower intensity levels observed in CNH relative to other groups.

Duration

Results showed significant effects of Emotion [$\beta = 0.2233$ (SE = 0.0336), $t(1599) = 6.8$, $p < 0.0001$] and Sentence [$\beta = 0.01244$ (SE = 0.00194), $t(1443) = 6.4$, $p < 0.0001$] but no effects of Group and no two-way or three-way interactions.

Spectral Centroid

Results showed a significant effect of Emotion [$\beta = -272.591$ (SE = 30.093), $t(1599) = -9.058$, $p < 0.0001$] but no effect of Group or Sentence and no interactions.

Acoustic Contrasts Between Happy and Sad Productions

The acoustic contrast between happy and sad productions was specifically investigated for each acoustic cue. For the mean F0, the contrast was defined as the ratio between the mean F0s for happy and sad productions. For the F0 s.d., the contrast was defined as the ratio between the standard deviations of F0 for happy and sad productions. For Intensity, the contrast was defined as the difference in dBs between mean intensities of happy and sad productions. For Duration, the contrast was defined as the ratio between the durations of happy and sad productions. For Spectral Centroid, the contrast was defined as the ratio between the spectral centroids of happy and sad productions. Ratios between the values for happy and sad productions were chosen over other measures (e.g., simple difference) for consistency with findings in the literature on auditory perception, which indicates that perceptual sensitivity to differences between sounds in specific acoustic dimensions are well modeled by a system that encodes the sensory input using a power law and/or logarithmic representation. LME analyses were conducted with Group and Sentence as fixed effects and by-subject random intercepts and Sentence as by-subject random slopes.

Mean F0 Contrasts

Results of the LME analysis showed a significant effect of Group [$\beta = 0.115$ (SE = 0.05), $t(39.00) = 2.237$, $p = 0.031$]. A pairwise t test with Bonferroni corrections showed significant differences between all Groups ($p < 0.0001$ in all cases). **Figure 2** (upper) shows boxplots of the mean F0 contrast for the four groups and for each of the 20 sentences. The CCI group (blue) shows the smallest contrast of all four groups.
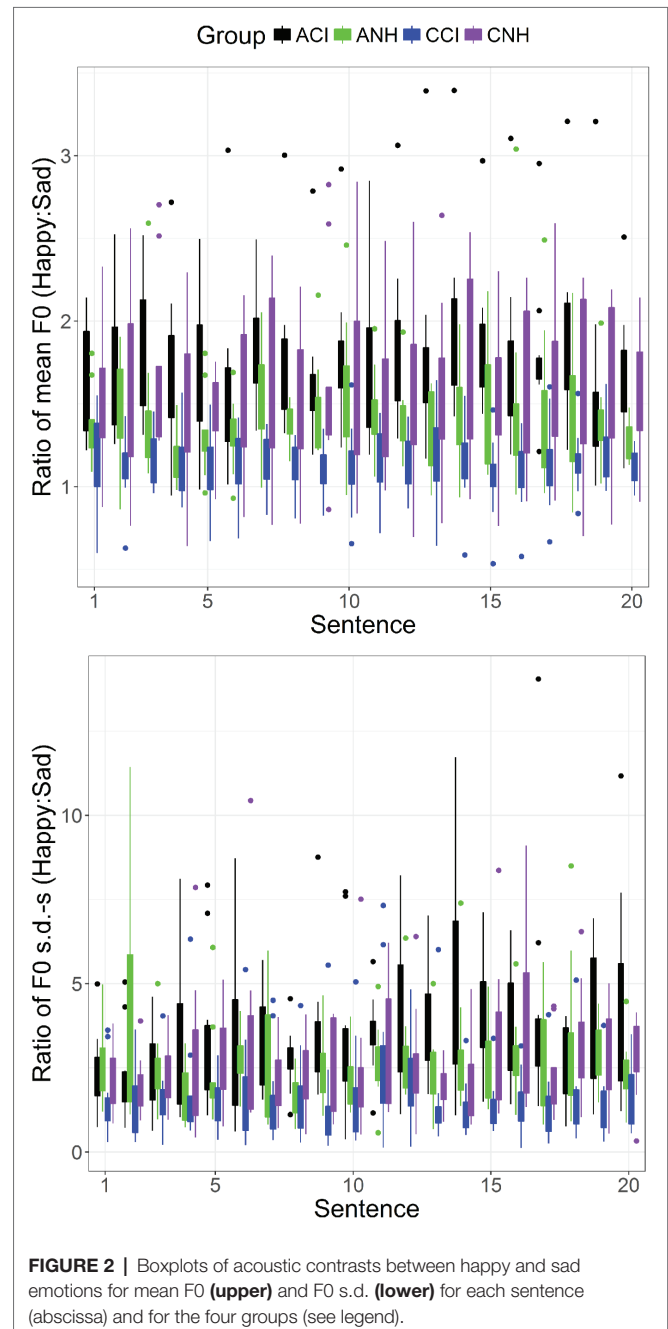
FIGURE 2 | Boxplots of acoustic contrasts between happy and sad emotions for mean F0 **(upper)** and F0 s.d. **(lower)** for each sentence (abscissa) and for the four groups (see legend).

F0 Standard Deviation Contrasts

Results of the LME analysis showed a significant effect of Group [$\beta = 0.35$ (SE = 0.160), $t(39.00) = 2.19$, $p = 0.0345$]. A pairwise t test with Bonferroni correction showed significant differences between all groups ($p < 0.0001$ in all cases). **Figure 2** (lower) shows boxplots of the F0 s.d. contrast for the four groups and

for each of the 20 sentences. The CCI group shows the smallest contrast of all four groups.

Intensity Contrast
The results of the LME analysis showed no significant effects of Sentence or Group and no interactions.

Duration Contrast
Consistent with previous analyses, results of the LME analysis showed no effects of Group or Sentence and no interactions.

Spectral Centroid Contrast
Consistent with previous analyses, the LME analysis showed no effects of Group or Sentence and no interactions.

Analyses of Results Obtained in Child Participants With Normal Hearing and Cochlear Implants
Initial analyses indicated different patterns for CNH and CCI and for female versus male children. Results obtained in NH and CI child participants were therefore analyzed separately for effects of Age and Sex on mean F0, F0 variation, Intensity, Duration, and Spectral Centroid. The data are plotted in **Figure 3**, which shows each acoustic cue as a function of Age, separated out by Sex and Group.

Acoustic Analyses of Productions by Children With Normal Hearing
Mean F0
An LME with fixed effects of Age and Sex and Sentence|Subject random intercepts/slopes showed significant effects of Sex [$\beta = 407.03$ (SE = 172.935), $t(11.4) = 2.354$, $p = 0.0375$], a significant interaction between Age and Sex [$\beta = -37.29$ (SE = 12.72), $t(11.4) = -2.932$, $p = 0.0132$], a significant interaction between Sex and Emotion [$\beta = -242.569$ (SE = 112.977), $t(351) = -2.147$, $p = 0.0325$], and a significant three-way interaction between Age, Sex, and Emotion [$\beta = 19.348$ (SE = 8.31), $t(351) = 2.328$, $p = 0.0205$]. The data are plotted in **Figure 3A** (right-hand panels). Consistent with expected differences in vocal development, the male children showed a larger decrease in F0 with age than female children did. The male children in this sample also showed a decreasing effect of emotion with age compared to the female children (hence, the three-way interaction).

F0 Variation
A parallel analysis to that described above with F0 s.d. as the dependent variable showed significant interactions between Age and Sex [$\beta = -10.77$ (SE = 3.33), $t(13.6) = -3.022$, $p = 0.0094$] and Sex and Emotion [$\beta = -80.765$ (SE = 39.08), $t(342) = -2.067$, $p = 0.0395$]. The pattern of results (**Figure 3B**, right-hand panels) is generally similar to that in **Figure 3A** for CNH, consistent with the correlation between the two variables. The separation between the emotions is somewhat smaller for male than for female participants in this sample, and there are age-related declines in the F0 s.d. in the male participants' productions that are not observed in the female participants' voices.

Intensity
An LME model with random slopes showed a significant effect of Emotion [$\beta = 6.393$ (SE = 2.364), $t(351) = -2.704$, $p = 0.0072$], a marginally significant two-way interaction between Age and Emotion [$\beta = 0.3916$ (SE = 0.209), $t(351) = -1.873$, $p = 0.062$], a significant two-way interaction between Sex and Sentence [$\beta = 0.8233$ (SE = 0.404), $t(351) = 2.037$, $p = 0.0425$], a three-way interaction among Sex, Emotion, and Sentence [$\beta = -1.447$ (SE = 0.572), $t(351) = -2.531$, $p = 0.0118$], a marginally significant three-way interaction among Age, Sex, and Sentence [$\beta = -0.0559$ (SE = 0.0297), $t(347.827) = -1.879$, $p = 0.0611$], and a four-way interaction among Age, Sex, Emotion, and Sentence [$\beta = 0.099$ (SE = 0.042), $t(351) = 2.358$, $p = 0.0189$]. The results are plotted in **Figure 3C** (right-hand panels). The separation between the emotions is clear, but for the male participants, the separation decreases somewhat of their age, more so than for female participants. The interaction with Sentence indicates that the pattern depends on the individual sentence.

Duration
An LME model with random slopes showed significant effects of Age [$\beta = -0.051$ (SE = 0.024), $t(12.30) = -2.459$, $p = 0.0296$] and Emotion [$\beta = 0.420$ (SE = 0.171), $t(351)=2.456$, $p = 0.0145$] with no other effects and no interactions. This is clearly apparent in **Figure 3D** (right-hand panels). The separation between the emotions remains consistent with age, across sentences, and for both sexes.

Spectral Centroid
An LME model with random slopes showed a significant effect of sex [$\beta = 980.364$ (SE = 359.645), $t(14.30) = 2.726$, $p = 0.0162$] and significant two-way interactions between Age and Sex [$\beta = -84.471$ (SE = 26.456), $t(14.30) = -3.193$, $p = 0.0064$], between Age and Emotion [$\beta = -19.712$ (SE = 9.686), $t(352) = -2.035$, $p = 0.0426$], and between Emotion and Sentence [$\beta = 21.47$ (SE = 9.145), $t(352) = 2.348$, $p = 0.0195$]. A three-way significant interaction among Sex, Emotion, and Sentence [$\beta = -84.10$ (SE = 26.493), $t(352) = -3.174$, $p = 0.0016$] and a four-way significant interaction among Age, Sex, Emotion, and Sentence [$\beta = 5.471$ (SE = 1.949), $t(352) = 2.808$, $p = 0.0053$] were also observed. The results are shown in **Figure 3E** (right-hand panels). It is apparent that the separation between the emotions decreases with age for the male participants, more so than for the female participants, and that the pattern varies across sentences.

Acoustic Analyses of Productions by Children With Cochlear Implants
Mean F0
Results obtained in child participants with CIs are plotted in the left-hand panels of **Figure 3A**. It is apparent that the separation between the emotions is smaller in the CI population than in the NH children (right-hand panels). A parallel analysis
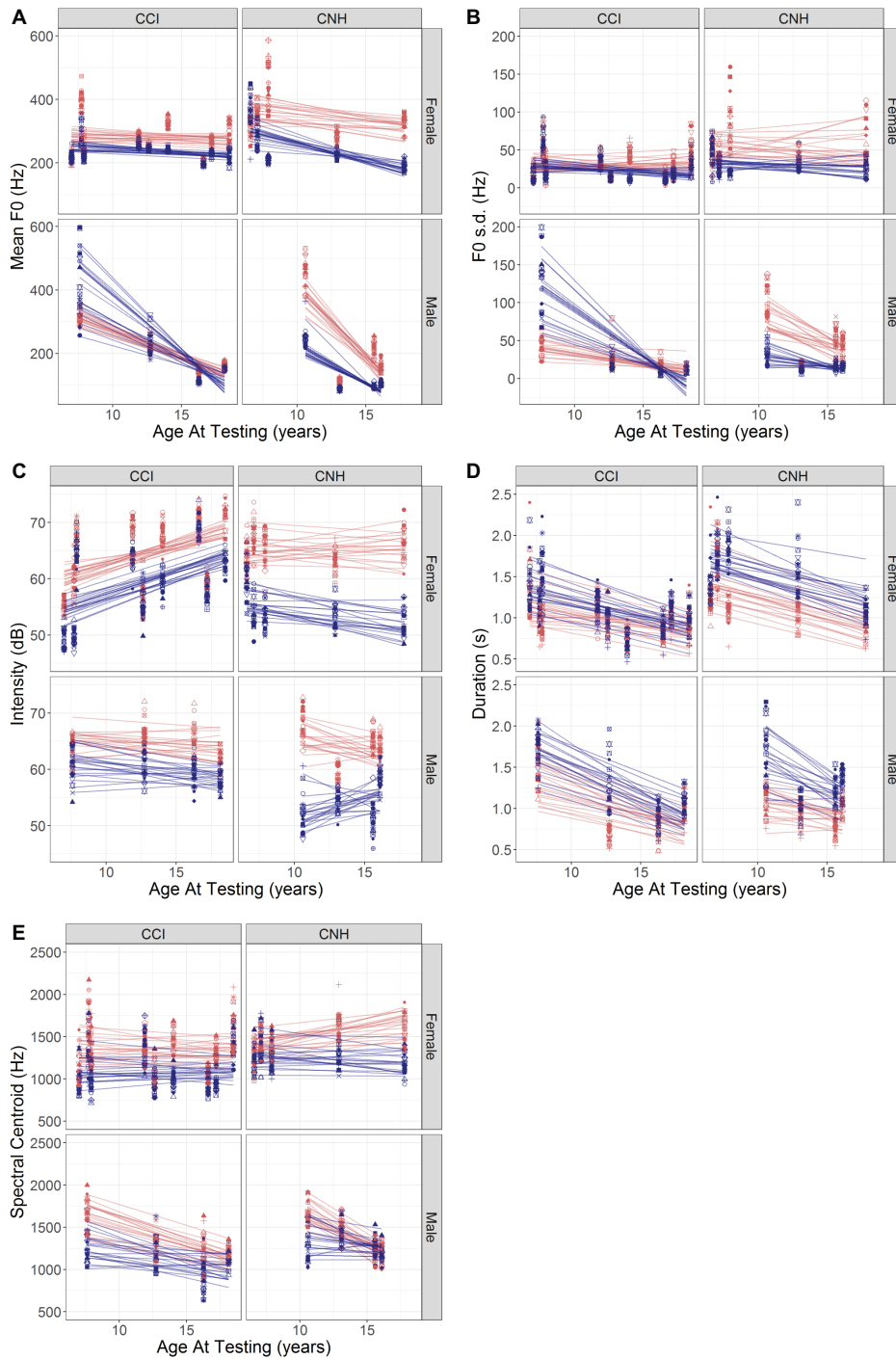
FIGURE 3 | (A–E) Values of acoustic features (**A**: mean F0; **B**: F0 s.d.; **C**: Intensity; **D**: Duration; **E**: Spectral Centroid) of the happy (red) and sad (blue) emotions recorded by CNH and CCI, plotted against their age (abscissa). For each acoustic feature, left- and right-hand panels show results in CCI and CNH, respectively, and upper and lower plots show results in female and male participants, respectively. The differently shaped symbols and lines in each color represent individual sentences recorded in each emotion.

that conducted with CNH showed a significant interaction between Age and Sex [$\beta = -12.971$ (SE = 5.619), $t(15.2) = -2.308$, $p = 0.0355$]. This is clear in the steeper slope obtained with male children with CIs than in the female children with CIs

in **Figure 3A** and parallels the findings with the CNH. A significant two-way interaction was observed between Sex and Emotion [$\beta = 285.99$ (SE = 42.159), $t(507) = 6.784$, $p < 0.0001$], and a three-way interaction among Age, Sex, and Emotion

showed a further effect of Age on the Sex by Emotion interaction [$\beta = -15.814$ (SE = 3.018), $t(507) = -5.24$, $p < 0.0001$]. These interactions are likely explained by the female participants' productions showing a consistent separation between happy and sad emotions with Age, while the male participants' productions show little to no separation, which changes in direction with Age. A marginally significant three-way interaction among Sex, Emotion, and Sentence was also observed [$\beta = -7.03$ (SE = 3.519), $t(507) = -1.998$, $p = 0.046$], likely due to the greater dependence of the mean F0 on Sentence for sad emotions produced by male children relative to their happy emotions and also relative to their female counterparts (**Figure 3A**, left-hand panels).

F0 Variation
Results obtained in the child participants with CIs are plotted in the left-hand panels of **Figure 3B**. It is evident that the separations between the two emotions are smaller in the CCI than in their CNH counterparts (**Figure 3B**, right-hand panels). A similar analysis as described above with F0 s.d. as the dependent variable showed a significant two-way interaction between Sex and Emotion [$\beta = 165.813$ (SE = 18.646), $t(507) = 8.893$, $p < 0.0001$], three-way interactions among Age, Sex, and Emotion [$\beta = -9.91$ (SE = 1.334), $t(507) = -7.425$, $p < 0.0001$] and among Sex, Emotion, and Sentence [$\beta = -3.991$ (SE = 1.557), $t(507) = -2.564$, $p = 0.0106$], and a four-way interaction among Age, Sex, Emotion, and Sentence [$\beta = 0.265$ (SE = 0.111), $t(507) = 2.378$, $p = 0.0178$]. The pattern of results is generally similar to those obtained with mean F0 (**Figure 3B**, left-hand panels).

Intensity
Results (plotted in the left-hand panels of **Figure 3C**) showed a marginally significant effect of Age [$\beta = 0.781$ (SE = 0.386), $t(13.70) = 2.022$, $p = 0.063$] and a significant effect of Emotion [$\beta = -4.198$ (SE = 1.692), $t(494) = -2.481$, $p = 0.0134$] and no other effects or interactions. It is apparent that the separation between the emotions is smaller in the children with CIs than in their counterparts with NH (**Figure 3C**, right-hand panels).

Duration
Results showed a significant negative effect of Age [$\beta = -0.0368$ (SE = 0.0137), $t(18.1) = -2.662$, $p = 0.0158$], indicating an overall faster speaking rate in older children, but no other effects or interactions. The effect of Age is similar to that observed in the children with NH.

Spectral Centroid
Results showed a significant effect of Emotion [$\beta = -253.695$ (SE = 121.996), $t(508.2) = -2.08$, $p = 0.0381$] but no other effects and no interactions. No obvious differences are apparent between the children with CIs and their NH counterparts.

Children With Cochlear Implants: Effects of Age at Implantation and Duration of Device Experience
The results obtained with CCI were separately analyzed for effects of Age at Implantation and Duration of Device Experience

on individual acoustic cues to emotion. Age at implantation was significantly correlated with Duration of Device Experience ($r = 0.63$, $p < 0.0001$), so these variables were considered separately in the statistical analyses. Consistent with the Duration of Device Experience being highly correlated with Age ($r = 0.99$), the statistical analyses with Duration of Device Experience as a fixed effect produced almost identical results to those previously described with Age as the fixed effect and are not reported here in the interest of space. Results with Age at Implantation as the fixed effect of interest are described below.

LME analyses were conducted with Age at Implantation, Emotion, and Sentence as fixed effects, random intercepts by subject, and random slopes for the effect of Sentence.

Mean F0
An LME analysis as described above with mean F0 as the dependent variable showed a significant effect of Emotion [$\beta = -73.119$ (SE = 30.5273), $t(507) = -2.395$, $p = 0.017$], a significant interaction between Emotion and Sex [$\beta = 251.6315$ (SE = 50.006), $t(507) = 5.032$, $p < 0.0001$], and a significant three-way interaction between Age at Implantation, Emotion, and Sex [$\beta = -131.525$ (SE = 35.046), $t(507) = -3.753$, $p = 0.0002$]. These interactions can be observed in **Figure 4** (top panel), which plots the ratio of mean F0 values for happy and sad emotions against Age at Implantation. Left- and right-hand panels show data obtained in female and male children. The acoustic contrast for mean pitch is relatively unchanging for female children but increases for male children with increasing Age at Implantation. This likely simply reflects the developmental effects in the male children observed in **Figure 3** (recall that Age at Implantation is correlated with age at testing).

F0 Variation
An LME analysis as described above with F0 s.d. as the dependent variable showed a significant interaction between Emotion and Sex [$\beta = 94.7914$ (SE = 24.711), $t(507) = 3.836$, $p = 0.0001$] and a three-way interaction between Age at Implantation, Emotion, and Sex [$\beta = -46.195$ (SE = 17.318), $t(507) = -2.667$, $p = 0.0079$]. **Figure 4** (middle panel) shows the F0 s.d. ratio between happy and sad emotions plotted against Age at Implantation. The patterns are similar to those observed with mean F0 and also consistent with the effects of Age in **Figure 3**.

Intensity
An LME analysis as described above with Intensity as the dependent variable showed a significant effect of Emotion [$\beta = -10.712$ (SE = 2.169), $t(494) = -4.938$, $p < 0.0001$] and a significant interaction between Age at Implantation and Emotion [$\beta = 4.715$ (SE = 1.567), $t(494) = 3.009$, $p = 0.0028$]. A marginally significant three-way interaction between Age at Implantation, Emotion, and Sex was also observed [$\beta = -4.71$ (SE = 2.49), $t(494) = -1.892$, $p = 0.0591$]. **Figure 4** (bottom panel) shows the intensity difference between happy and sad emotions plotted against Age at Implantation. The interaction between Age and Emotion appears to be determined by the

FIGURE 4 | (Top to bottom) Mean F0, F0 s.d., and Intensity of productions by CCI, plotted against their age at implantation. Left- and right-hand panels show results in female and male participants, respectively. Red and blue symbols represent happy and sad emotions, respectively, and the differently shaped symbols and the lines represent individual sentences recorded in each emotion.

female children who produce larger intensity differences (happy > sad) at earlier ages of implantation. This pattern of results is distinct from that observed in **Figure 3** with Age. Note, however, that a nonlinear fit may have better captured the trends in this dataset, specifically, the elevated intensities observed at some ages at implantation: however, given the small sample size, we refrained from attempting such a fit to avoid problems with overfitting.

Duration
An LME analysis as described above with Duration as the dependent variable showed a marginally significant effect of Emotion [$\beta = 0.3311$ (SE = 0.18), $t(507) = 1.842$, $p = 0.066$], but no other effects or interactions reached significance.

Spectral Centroid
An LME analysis as described above with Spectral Centroid as the dependent variable showed a significant effect of Emotion [$\beta = -380.227$ (SE = 162.677), $t(507.8) = -2.337$, $p = 0.0198$] but no other effects and no interactions.

DISCUSSION

Summary of the Results
Analysis of the mean F0, F0 variation (F0 s.d.), Intensity, Duration, and Spectral Centroid for the happy and sad emotions showed significant effects of Emotion on each of the cues measured. The mean F0, F0 s.d., mean Intensity, and Spectral Centroid were each higher for happy than for sad emotion productions, whereas Duration was shorter for happy than for sad. These basic findings are consistent with acoustic analyses reported in the literature in typical adult populations (e.g., Banse and Scherer, 1996). We were particularly interested in differences between the groups in the effect of the individual emotions. An interaction between Emotion and Group was observed for mean F0, F0 s.d., and Intensity but not for Duration or Spectral Centroid. The Group by Emotion interaction for Intensity was not well supported in *post hoc* analyses and was not reflected in the analysis of Intensity contrasts between the two emotions, which showed no significant effect of Group. Thus, the only reliably strong Emotion by Group interactions were those observed in F0 and F0 s.d. measures. The Emotion by Group interaction for mean F0 was explained by *post hoc* analyses indicating that mean F0 values were not significantly different for happy productions across the groups, but there was a significant difference between the pitch of sad productions between groups: while the adult NH and CI groups did not differ significantly, the CCI group produced a higher mean F0 for sad emotions than all others. *Post hoc* analyses on the F0 s.d. measures showed that the primary factor driving the Group by Emotion interaction was that the CCI group's happy productions were the most monotonous of the four groups. Analyses of the acoustic contrasts between the two emotions further confirmed these findings.

The spectral centroid of an utterance provides information about the overall shape of the spectrum and is expected to be reflective of the phonetic content of the utterance, but it also provides information about emotion. Specifically, the relative energy in the lower and higher portions of the spectrum changes with emotion. As an example, Banse and Scherer (1996) showed that the decrease in energy at frequencies higher than 1,000 Hz is one of the important acoustic cues for emotion. These differences are reasonably well captured in the spectral centroid measure. For instance, positive emotions tend to be associated with more energy in the higher frequencies (higher spectral centroid), while negative or unpleasant emotions are associated with more energy at lower frequencies (lower spectral centroid). The present results suggest that all four groups showed similar changes in spectral centroid between happy and sad productions.

Consistent with the fact that the duration cue is well represented in CI processing, all four groups showed similar changes in duration with emotion, reflecting the expected faster speaking rate for happy emotion and a slower speaking rate for sad emotion. The Intensity cue is also represented in CIs, although the limited dynamic range and the effects of the automatic gain control do distort intensity-domain information, and this is consistent with the results showing that, similar to the other groups, the CCI also produced louder speech for happy than for sad emotions.

Taken together, the analyses of the group data indicate that the CCI produce happy and sad emotions with normal-range distinctions in duration, intensity, and spectral shape. The deficit appears to be focused on F0 (voice pitch)-related parameters in this dataset. Specifically, CCI produce smaller contrasts in mean F0 and in F0 variation than other groups. The reduced production of F0 contrasts is consistent with a degraded perception of voice pitch through CIs. The reduced F0 s.d. for happy emotions in CCI suggests a more monotonous speaking style overall, which may impose difficulties in social communication by this population. These data also suggest that CCI do not exaggerate contrasts between the cues as they are more perceptually sensitive (e.g., duration, intensity) to distinguish emotions in their speech. However, it is possible that differences do exist between CCI and other groups in these parameters and that a study with a larger sample size might reveal such differences. Based on the present dataset, it appears that F0-related cues are more strongly and more consistently impacted in CCIs' productions than other cues.

The analyses of the CNH and CCIs' productions were conducted separately to investigate developmental effects and effects of sex. Results in the CNH group showed interactions among Age, Sex, and Emotion, with the male children's mean F0 decreasing more than the female children's as they reached their upper teenage years. With the deepening voices, visual inspection of the data further suggested that the older male children also produced smaller contrasts between happy and sad emotions than did their female peers.

The CCI's productions showed similar effects of Age and Sex, although the acoustic contrasts were clearly smaller for the CCIs' productions than for the CNHs' productions. Male

CCI showed a deepening voice pitch with increasing age, while female CCI showed relatively small changes in voice pitch with age. The two younger male children with CIs showed a strong dispersion of mean F0 across sentences, particularly for sad productions, and higher mean F0 for sad than for happy productions for some of the sentences. The trend reversed in the older male children who showed the expected lower mean F0 for sad than for happy productions, but the separation remained small (**Figure 3A**). Note, however, that the limited sample size precludes the drawing of firm conclusions. Measures of F0 s.d. showed similar patterns. Intensity, Duration, and Spectral Centroid did not show any interactions between Age and Emotion in the CCI.

Analyses of effects of Age at Implantation and Duration of Device Experience were conducted separately because these two variables were correlated with one another. Duration of Device experience was highly correlated with Age, and the patterns of findings were virtually identical. Effects of mean F0 and F0 s.d. showed similar patterns with increasing Age at Implantation as those observed with Age and with Duration of Device Experience. The correlations between these variables preclude clear inferences regarding the underlying mechanisms. It is likely that the deepening mean F0 with Age at Implantation in male CCI is simply a reflection of developmental changes with Age.

The analysis of Intensity showed a different effect of Age at Implantation than did Age, and therefore, this effect is more likely to be unique to Age at Implantation. There was a significant two-way interaction between Age at Implantation and Emotion modified by Sex in a further three-way interaction. Visual inspection of **Figure 4** (lower panels) suggests that the interaction was due to a greater separation of the emotions in earlier-implanted children than in later-implanted children, an effect that is stronger in female than in the male children in the present sample.

Comparison Between Children With Cochlear Implants' and Adults With Cochlear Implants' Production of Emotions

Although both CCI and ACI hear speech through the degradation of CI processing combined with electric stimulation, the present results indicate that the two groups produce vocal emotions very differently. While the ACIs' productions showed clear separations between the emotions in all measures considered, the CCI showed significantly smaller acoustic contrasts in F0 and in F0 variation than all other groups. On the other hand, ACIs' perceptions of vocal emotions have been shown to be comparable to CCIs' productions, even with the exaggerated prosody of child-directed speech (Chatterjee et al., 2015). This suggests that perception and production of vocal emotions may be linked in CCI who learned to speak through electric hearing, but not in ACI who learned to speak through acoustic hearing. We conclude that access to acoustic information in the early developmental years is crucial for the development of vocal motor patterns. In the ACI, these patterns seem to

have been retained despite years of listening to a highly degraded, abnormal speech input. The CCI, on the other hand, had no access to usable hearing prior to implantation, and this is reflected in their atypical patterns of emotional prosody. We note here that the CCI produced the words in the sentences with high accuracy (this was separately verified by asking normally hearing listeners to listen to the recordings and repeat back the words in the recorded productions without regard to emotion). The ACI also produced the words with high accuracy. It is possible that speech therapy in CCI focuses more on speech phonetics of words than on speech prosody and that a greater focus on prosody in general may be beneficial to CCI. We note that ACI in the United States do not receive more than minimal speech therapy after implantation.

Links to Related Studies in the Literature

Similar to the present study, other studies of vocal emotion production by children with CIs have also focused on primary emotions such as happy and sad primarily because they are highly contrastive in multiple acoustical dimensions as well as in their conceptual meaning. The present study focused on acoustic analyses, while other studies have investigated the intelligibility of the emotional productions. Additionally, in the majority of other studies, the child participants were tasked to imitate the emotional productions of an exemplar, while in the present study, participants were not provided with any examples, training, or targeted feedback. A recent study (Van De Velde et al., 2019) did not use imitative productions, but their methodology was quite different, and as discussed in the section Introduction, the task was more complex. These differences notwithstanding the present findings of reduced acoustic contrasts between the emotions in CCI are consistent with the findings of previous studies showing impaired or less recognizable emotions produced by CCI. These findings are also consistent with previous findings of impaired production of question/statement contrasts and lexical tones by children with CIs. Studies of singing by children with CIs also show impairments, although music requires a far greater sense of pitch, and therefore, singing may be considered a far more difficult task than producing speech intonations. Our finding that Age at Implantation had modest effects on the productions whereas Age at Testing (highly correlated with Duration of Device Experience) had a stronger impact is consistent with the findings of Van De Velde et al. (2019), who also found improvements with increased hearing age in their cohort of children with CIs.

In a recent investigation (Damm et al., 2019), the identifiability of these identical recordings made by the same participants was measured by asking normally hearing child and adult listeners to indicate whether each recording sounded happy or sad. In contrast to the normally hearing talkers and the postlingually deaf adult talkers, the CCI group's recordings showed deficits in how well their recorded emotions were identified. In that study, Age at Implantation was found to be a significant predictor, with the earlier-implanted CCIs' emotions being significantly better identified than the later-implanted CCIs' productions. The group results are consistent

between the two studies (i.e., the CCI in the present study produced smaller acoustic contrasts than other groups, and their emotions were also more poorly identified than other groups in the Damm et al. study). However, a larger dataset would be needed to establish direct relationships between acoustic features in individual talkers' emotion productions and how well they can be identified by listeners.

Limitations, Strengths, and Clinical Implications of the Present Study

The present study suffers from several limitations. First, the limited sample size leads us to treat these findings with caution. Thus, it is possible that a larger sample size might reveal differences in acoustic features such as intensity, duration, or spectral centroid, which are significant but cannot be captured with a small dataset. Second, the information about Age at Implantation was obtained from parents or guardians of the child participants and could not be verified independently. Third, perceptual data on this cohort of CCI's emotion recognition abilities were not obtained, nor were data on their general or social cognition or other linguistic abilities. Fourth, the correlations between specific variables of interest (such as age at implantation and duration of device experience) precluded investigations of their combined effects. This, however, is a problem that is inherent to CI studies and not easily remedied in experimental design. Further, information about access/use of speech therapy in the CCI was not obtained. Finally, the method used to elicit the emotions had some limitations in that spontaneous expression of emotions was not achieved. There may well be differences between the emotions recorded using brief sentences in the laboratory and natural emotions communicated by the participants in their everyday life. Differences in the prosody of read or scripted speech as opposed to spontaneous speech have been reported in the literature (Laan, 1997). Although Damm et al. (2019) found that these methods evoked highly identifiable emotions in the CNH, ANH, and ACI groups, the differences between laboratory-recorded and naturally spoken emotional speech may further modify the group differences observed here. These limitations should be addressed in future studies.

Despite these limitations, the present results represent the first attempt to compare emotional productions by prelingually deaf children with CIs with postlingually deaf adult CI users, alongside normally hearing peers. One strength of the design was the careful selection of CCI who – with the caveat that the information was based on self- and parent-report and could not be independently verified – had no prior usable hearing at birth to more clearly separate them from postlingually deaf ACI who had good hearing in their early years. The findings suggest a key role of access to acoustic information during development for the production of prosodic cues. They also shed new light on specific sources of the impairment in emotional productions that could help develop improved speech therapy tools for children with CIs. For instance, the data suggest that the CCIs' small acoustic contrast between mean F0 for happy and sad emotions in the present study was driven by an insufficiently low mean F0 for sad emotions compared

to other populations. Additionally, CCIs' small acoustic contrast between F0 variations for happy and sad emotions was driven by an overly monotonous production of happy emotions compared to other populations. These impairments may be addressed in targeted speech therapy. Finally, although the sample size was small, the findings suggest the possibility of differences between male and female children in their productions of vocal emotions and speech intonations. Specifically, the results suggest that male children with CIs may encounter difficulties adjusting to their changing vocal pitch with increasing age. This is an aspect that needs further investigation with larger sample sizes.

The emotions selected for the present study were chosen for their high acoustic and conceptual contrast. We speak with a higher pitch, with more pitch modulations, louder and faster when we communicate in a happy way. By contrast, we speak with a lower pitch, more monotonously, softer and slower when we communicate in a sad way. The vocal tract changes in a contrastive way between these emotions as well. The deficits observed in the CCI in the present study with these highly contrastive emotions may underestimate the true nature of the deficit when more subtle emotions are to be communicated through prosodic cues in speech. A study investigating how well these participants' recordings were heard as happy or sad by normally hearing listeners (Damm et al., 2019) showed strong variability among the CCI talkers. Although some were very well understood, others' emotions were mislabeled more frequently. Overall, the CCIs' productions were less correctly identified than the ACIs', the CNHs', and the ANHs' productions. On the other hand, the CCIs' productions of the words in the sentences were highly recognizable. It is worth noting that present-day clinical protocols are designed with a focus on word and sentence recognition, with little to no emphasis on speech prosody. These findings, and others in the current literature, underscore a crucial need to address vocal pitch and emotion communication in the pediatric CI population in both the realms of scientific research and clinical intervention. The positive findings with ACI indicate that the presence of acoustic hearing (particularly at low frequencies) at birth and during development provides a supportive role in vocal emotion production, which is retained long after that hearing is lost. This result suggests a benefit to retaining any residual acoustic hearing in CCI alongside cochlear implantation, at least in the area of the production of emotional (and likely other forms) of speech prosody.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Boys Town Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

MC conceptualized the study, led the design and set-up, led the acoustic analyses, analyzed the data, and wrote the manuscript. AK contributed to the study design and set-up, led the data collection, and processing. RS conducted acoustic analyses and helped with manuscript writing. JC contributed to study design and data collection and conducted preliminary acoustic analyses. MH contributed to study design and acoustic analyses. JS contributed to study design, data collection, and acoustic analyses. SD contributed to study design, data collection, and acoustic analyses.

FUNDING

REFERENCES

Banse, R., and Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *J. Pers. Soc. Psychol.* 70, 614–636. doi: 10.1037/0022-3514.70.3.614

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot Int.* 5, 341–345.

Boersma, P., and Weenink, D. (2019). Praat: doing phonetics by computer [Computer program]. Version 6.0.52. Available at: http://www.praat.org/ (Accessed May 2, 2019).

Chatterjee, M., and Peng, S. C. (2008). Processing F0 with cochlear implants: modulation frequency discrimination and speech intonation recognition. *Hear. Res.* 235, 143–156. doi: 10.1016/j.heares.2007.11.004

Chatterjee, M., Zion, D. J., Deroche, M. L., Burianek, B. A., Limb, C. J., Goren, A. P., et al. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. *Hear. Res.* 322, 151–162. doi: 10.1016/j.heares.2014.10.003

Damm, S. A., Sis, J. L., Kulkarni, A. M., and Chatterjee, M. (2019). How vocal emotions produced by children with cochlear implants are perceived by their hearing peers. *J. Speech Lang. Hear. Res.* (in press).

Deroche, M. L., Kulkarni, A. M., Christensen, J. A., Limb, C. J., and Chatterjee, M. (2016). Deficits in the sensitivity to pitch sweeps by school-aged children wearing cochlear implants. *Front. Neurosci.* 10:73. doi: 10.3389/fnins.2016.00073

Fox, J., and Weisberg, S. (2011). *An {R} companion to applied regression.* 2nd Edn. Thousand Oaks CA: Sage. Available at: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion (Accessed July 1, 2019).

Gilbers, S., Fuller, C., Gilbers, D., Broersma, M., Goudbeek, M., Free, R., et al. (2015). Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. *i-Perception* 6:0301006615599139.

Green, T., Faulkner, A., Rosen, S., and Macherey, O. (2005). Enhancement of temporal periodicity cues in cochlear implants: effects on prosodic perception and vowel identification. *J. Acoust. Soc. Am.* 118, 375–385. doi: 10.1121/1.1925827

Hopyan-Misakyan, T. M., Gordon, K. A., Dennis, M., and Papsin, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children

with unilateral right cochlear implants. *Child Neuropsychol.* 15, 136–146. doi: 10.1080/09297040802403682

Jiam, N. T., Caldwell, M., Deroche, M. L., Chatterjee, M., and Limb, C. J. (2017). Voice emotion perception and production in cochlear implant users. *Hear. Res.* 352, 30–39.

Kuznetsova, A., Brockhoff, P. B., and Christensen, R. H. B. (2017). lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* 82, 1–26. doi: 10.18637/jss.v082.i13

Laan, G. P. M. (1997). The contribution of intonation, segmental durations, and spectral features. *Speech Comm.* 22, 43–65. doi: 10.1016/S0167-6393(97)00012-5

Luo, X., Fu, Q.-J., and Galvin, J. J. 3rd. (2007). Vocal emotion recognition by normal-hearing listeners and cochlear implant users. *Trends Amplif.* 11, 301–315. doi: 10.1177/1084713807305301

Luo, X., Kern, A., and Pulling, K. R. (2018). Vocal emotion recognition performance predicts the quality of life in adult cochlear implant users. *J. Acoust. Soc. Am.* 144, EL429–EL435. doi: 10.1121/1.5079575

Nakata, T., Trehub, S. E., and Kanda, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. *J. Acoust. Soc. Am.* 131, 1307–1314. doi: 10.1121/1.3672697

Nash, J. C., and Varadhan, R. (2011). Unifying optimization algorithms to aid software system users: optimx for R. *J. Stat. Softw.* 43, 1–14. doi: 10.18637/jss.v043.i09

Paquette, S., Ahmed, G. D., Goffi-Gomez, M. V., Hoshino, A. C. H., Peretz, I., and Lehmann, A. (2018). Musical and vocal emotion perception for cochlear implant users. *Hear. Res.* 370, 272–282. doi: 10.1016/j.heares.2018.08.009

Peng, S. C., Lu, N., and Chatterjee, M. (2009). Effects of cooperating and conflicting cues on speech intonation recognition by cochlear implant users and normal hearing listeners. *Audiol. Neurotol.* 14, 327–337. doi: 10.1159/000212112

Peng, S.-C., Lu, H.-P., Lu, N., Lin, Y.-S., Deroche, M. L. D., and Chatterjee, M. (2017). Processing of acoustic cues in lexical-tone identification by pediatric cochlear-implant recipients. *J. Speech Lang. Hear. Res.* 60, 1223–1235. doi: 10.1044/2016_JSLHR-S-16-0048

Peng, S.-C., Tomblin, J. B., Cheung, H., Lin, Y.-S., and Wang, L.-S. (2004). Perception and production of mandarin tones in prelingually deaf children with cochlear implants. *Ear Hear.* 25, 251–264. doi: 10.1097/01.AUD.0000130797.73809.40

Peng, S.-C., Tomblin, J. B., and Turner, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. *Ear Hear.* 29, 336–351. doi: 10.1097/AUD.0b013e318168d94d

Přibilová, A., and Přibil, J. (2009). "Spectrum modification for emotional speech synthesis" in *Multimodal signals: Cognitive and algorithmic issues*. Berlin, Heidelberg: Springer, 232–241.

R Core Team (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: https://www.R-project.org/ (Accessed January 5, 2019).

Scherer, K. R. (2003). Vocal communication of emotion: a review of research paradigms. *Speech Comm.* 40, 227–256. doi: 10.1016/S0167-6393(02)00084-5

Schorr, E. A., Roth, F. P., and Fox, N. A. (2009). Quality of life for children with cochlear implants: perceived benefits and problems and the perception of single words and emotional sounds. *J. Speech Lang. Hear. Res.* 52, 141–152. doi: 10.1044/1092-4388(2008/07-0213)

See, R. L., Driscoll, V. D., Gfeller, K., Kliethermes, S., and Oleson, J. (2013). Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing. *Otol. Neurotol.* 34, 490–498. doi: 10.1097/MAO.0b013e318287c985

Sueur, J. (2018). *Sound analysis and synthesis with R*. Culemborg, Netherlands: Springer, Use R! Series.

Sueur, J., Aubin, T., and Simonis, C. (2008). Seewave: a free modular tool for sound analysis and synthesis. *Bioacoustics* 18, 213–226. doi: 10.1080/09524622.2008.9753600

Tinnemore, A. R., Zion, D. J., Kulkarni, A. M., and Chatterjee, M. (2018). Children's recognition of emotional prosody in spectrally-degraded speech is predicted by their age and cognitive status. *Ear Hear.* 39, 874–880. doi: 10.1097/AUD.0000000000000546

Van De Velde, D. J., Schiller, N. O., Levelt, C. C., Van Hueven, V. J., Beers, M., Briaire, J. J., et al. (2019). Prosody perception and production by children with cochlear implants. *J. Child Lang.* 46, 111–141. doi: 10.1017/S0305000918000387

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. New York: Springer-Verlag.

Winn, M. B., Chatterjee, M., and Idsardi, W. J. (2012). The use of acoustic cues for phonetic identification: effects of spectral degradation and electric hearing. *J. Acoust. Soc. Am.* 131, 1465–1479. doi: 10.1121/1.3672705

Pragmatic Language Skills: A Comparison of Children With Cochlear Implants and Children Without Hearing Loss

Michaela Socher [1]*, Björn Lyxell [1,2], Rachel Ellis [1], Malin Gärskog [3], Ingrid Hedström [3] and Malin Wass [4]

[1] Swedish Institute of Disability Research, Linköping University, Linköping, Sweden, [2] Special Needs Education, University of Oslo, Oslo, Norway, [3] Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden, [4] Department of Business Administration, Technology and Social Sciences, Luleå University of Technology, Luleå, Sweden

Pragmatic language ability refers to the ability to use language in a social context. It has been found to be correlated with success in general education for deaf and hard of hearing children. It is therefore of great importance to study why deaf and hard of hearing children often perform more poorly than their hearing peers on tests measuring pragmatic language ability. In the current study the Pragmatics Profile questionnaire from the CELF-IV battery was used to measure pragmatic language ability in children using cochlear implants ($N = 14$) and children without a hearing loss ($N = 34$). No significant difference was found between the children with cochlear implants (CI) and the children without hearing loss (HL) for the sum score of the pragmatics language measure. However, 35.71% of the children with CI performed below age norm, while only 5.89% of the children without HL performed below age norm. In addition, when dividing the sum score into three sub-measures: Rituals and Conversational skills (RCS), Asking for, Giving, and Responding to Information (AGRI), and Nonverbal Communication skills (NCS), significant differences between the groups were found for the NCS measure and a tendency for a difference was found for the RCS measure. In addition, all three sub-measures (NCS, AGRI, RCS) were correlated to verbal fluency in the children with CI, but not the children without HL.

Keywords: pragmatic language ability, hearing loss, cochlear implant, verbal fluency, children

1. INTRODUCTION

Pragmatic language ability refers to the ability to use language in a social context. It has been shown to be related to core language ability, including language comprehension and vocabulary skills, and also to cognitive skills (Matthews et al., 2018) for example inhibition, shifting, working memory (Channon and Watts, 2003; Blain-Briére et al., 2014) and reasoning ability (Turkstra et al., 1996). Children with autism spectrum disorder (Norbury and Bishop, 2002; Volden et al., 2009), children with ADHD (Camarata and Gibson, 1999; Kim and Kaiser, 2000; Staikova et al., 2013), and deaf and hard of hearing children (Jeanes, 2000; Most et al., 2010; Goberis et al., 2012; Rinaldi et al., 2013) tend to show poorer performance on several pragmatic language measures compared to typically developing children. Pragmatic language ability seems to be associated with success in general

education for deaf and hard of hearing children (Thagard et al., 2011). Thagard et al. (2011) showed that children with higher pragmatic language ability also have higher scores on tests measuring preparedness for first-grade work, math, and reading. Furthermore these children spend more time in general education settings. However, the causal direction of the relationship is unclear. Other studies have suggested that pragmatic language ability is less developed in deaf and hard of hearing children as both the quality and quantity of their daily face-to-face discourses are reduced (Jeanes, 2000; Most et al., 2010). Most et al. (2010) argue that a delay in language development resulting in less flexible use of language structures, reduced audibility during interactions and difficulties with theory of mind might be reasons for the differences seen between children with normal hearing and deaf and hard of hearing children. However, only few and quite diverse studies have focused on children with cochlear implants (CI) and their pragmatic language ability (Jeanes, 2000; Toe et al., 2007; Most et al., 2010; Thagard et al., 2011; Dammeyer, 2012; Goberis et al., 2012; Rinaldi et al., 2013; Toe and Paatsch, 2013).

Pragmatic language skills is an umbrella term for a number of complex verbal and non-verbal skills needed for real-life conversations. These skills range from responding to utterances in an appropriate way, maintaining the topic of the conversation, initiating new, and relevant topics (Matthews, 2014), to not inappropriately interrupt the other speaker, turn-taking (Bonifacio et al., 2007; Longobardi et al., 2017), the ability to ask for clarification and adapting the language to the needs of the conversational partner (Longobardi et al., 2017). In order to be able to successfully use these skills it is important to be able to consider all or some of the following: the context of an utterance (Loukusa et al., 2007; Matthews et al., 2018), acoustic cues like intonation and stress (Paradis, 1998; Most et al., 2010), and non-verbal cues (Russell and Grizzle, 2008). Pragmatic language skills develop during childhood (Loukusa et al., 2007; Longobardi et al., 2017). Mastering these complex skills takes until adolescence or even early adulthood (Matthews, 2014). Pragmatic language skills have been linked to social competence (Conti-Ramsden and Botting, 2004), peer relationship, mental health (Helland et al., 2014), and collaborative-based learning (Murphy et al., 2014).

Children with CI have been found to perform more poorly on a number of pragmatic language abilities. Jeanes (2000) analyzed referential communication between children (four age groups: 8-, 11-, 14-, and 17-year old) and found that profoundly deaf high school students using oral communication use requests for clarification more often than their hearing peers. However, in comparison to the hearing group, the requests were more often nonspecific, which led Jeanes (2000) to suggest that the communicative competence of the deaf and hard of hearing children is not as mature. Ibertsson et al. (2009) as well found teenagers with CI to use more requests for clarification when communicating with a well-known peer without a hearing loss (HL). However, in contrast to Jeanes (2000) in the study by Ibertsson et al. (2009) the teenagers with CI mostly used specific requests for clarifications. Ibertsson et al. (2009) argue that this is a way to control the conversation more. In accordance to this a more recent study done by Toe and Paatsch (2013) indicates that

the pragmatic language skills of children with CI at age 9–12 are good enough to ensure a fluent conversation, but that they tend to control the conversation more than children without HL. Toe and Paatsch (2013) analyzed 10 min spontaneous conversations between children with CI and children without HL the same age. Toe and Paatsch (2013) suggest that children with CI try to control the conversation more in order to prevent conversation breakdown. In addition, the results found by Toe and Paatsch (2013) indicate that children with CI have problems with contingency, the ability to maintain the topic of the conversation and to add new and relevant information. This is in accordance with results found by Most et al. (2010). The authors evaluated spontaneous communication between children age 6 and 9 and a familiar adult. Most et al. (2010) found that both children with CI and children with hearing aids (HA) showed problems in the area of turn taking, the ability to have a conversation with smooth interchanges between the conversational partners. This was especially the case for contingency, a skill which none of the children with CI or HA used appropriately, and response and adjacency (no pause between the utterance of the conversational partner and the child's utterance), two skills which were only used appropriately by two of the children with CI or HA. In the studies by Jeanes (2000), Most et al. (2010), and Toe and Paatsch (2013) one instance of conversation in the lab was analyzed. One disadvantage with this approach is that it is not clear whether results translate to real-life, where children need to communicate with different partners in different settings. In order to capture how well children are doing in real-life, other studies have used questionnaires to measure pragmatic language skills in children with CI. Goberis et al. (2012) used a checklist with items covering six categories: states needs, gives commands, personal interaction, wants explanation, shares knowledge, and shares imagination. Parents were then asked to rate a number of skills in each category to be: not present, preverbal, uses on to three words, or uses more complex language. By age six children with CI only used complex language for 6.6% of the items and by age seven they used complex language for 69% of the items. In comparison, children without HL used complex language for 100% of the items by the time they were 6 years old. In contrast to that, results from Guerzoni et al. (2016) suggest that already toddlers with a CI have pragmatic language skills comparable to hearing toddlers. Guerzoni et al. (2016) used a questionnaire using two scales, one for assertiveness and one for responsiveness. The assertiveness scale included items covering the ability to ask questions, make requests, and make suggestions, while the responsiveness scale covered the ability to respond to questions and requests, and turn taking. However, in contrast to the study by Goberis et al. (2012) parents only rated how often a certain behavior occurred. As the children in the study by Guerzoni et al. (2016) were only around 2 years of age it might be that differences between the groups were not apparent because they are only observed for more complex skills and more complex conversations, which a toddler might not yet have. Overall it seems like children with CI have problems with some but not all domains of pragmatic language ability. It should be emphasized that there are only very few studies studying pragmatic language ability in deaf and hard of hearing children

and those existing are very diverse, using different age groups and measures. In addition, there is a large time gap between some of the studies. It is therefore unclear if technical improvement of cochlear implants, changes in rehabilitation programs, the use of different measures or the age of the participants have led to different results. The present study aims to get an insight into the current real-life pragmatic language skills of children with CI and to compare them to those of children without HL.

It has been suggested that pragmatic language ability is not only connected to other language skills but also to different cognitive abilities (Turkstra et al., 1996; Channon and Watts, 2003; Martin and McDonald, 2003; Douglas, 2010; Blain-Briére et al., 2014; Matthews et al., 2018). Matthews et al. (2018) point out that it is hard to distinguish between pragmatic language ability and the ability to understand words and sentences. The authors add that some children still mainly show language problems in a social context and that it is therefore important to try to separate these skills. It is not surprising that most studies reviewed by Matthews et al. (2018) found correlations between "formal language" (vocabulary and grammar) and pragmatic language ability. The ability to understand sentences and words do not, however, seem to be the only important skills. Other studies have also shown associations to reasoning ability (Turkstra et al., 1996), cognitive flexibility (Ketelaars et al., 2012; Bacso and Nilsen, 2017), working memory, inhibition, and shifting ability (Channon and Watts, 2003; Blain-Briére et al., 2014; Matthews et al., 2018). Children with CI have been found to perform more poorly than children without HL on a number of executive function skills, like working memory (Wass et al., 2008; Kronenberger et al., 2013), reasoning (Bandurski and Ga1kowski, 2004; Edwards et al., 2011), and cognitive flexibility (Kenett et al., 2013; Wechsler-Kashi et al., 2014). These abilities seem to be associated with pragmatic language ability in normally developing children (Turkstra et al., 1996; Ketelaars et al., 2012; Blain-Briére et al., 2014; Bacso and Nilsen, 2017; Matthews et al., 2018). A delay in these cognitive functions might therefore lead to a delay in pragmatic language skills. However, the association between these cognitive skills and pragmatic language ability in children with CI has not yet been studied. Previous research suggests that the development of certain pragmatic language skills is delayed in children with CI compared to children without HL even when being matched on language ability (Most et al., 2010). This indicates that other factors apart from language ability play a role. To our knowledge there is no study looking into the connection between language measures, cognitive measures and pragmatic language ability in children with CI in comparison to children without HL. This study aims to take the first step in filling this research gap.

2. METHODS

2.1. Participants

Fifty-five children participated in the study. Seventeen of them were fitted with cochlear implants (CI). The 17 children with CI were recruited from a special school as well as via the hearing clinic in Stockholm, Sweden. They attended pre-school class, first grade, and second grade, respectively. The hearing loss of one child was caused by Usher syndrome. This syndrome leads also to a visual disability. Unfortunately no data concerning the visual impairment was collected. However, it was not reported by the test leader that any visual problems occurred during testing. To our knowledge, none of the other children had any additional disability apart from their hearing loss. Three of the children with CI were excluded from the study. One because the parents did not fill in the Pragmatics Profile and two as data on three of the other measures were missing. The mean age of the remaining 14 children (10 girls) was 6.77 years with a standard deviation of 11.13 months. Three of the children were unilaterally implanted and 11 had bilateral CIs. Their deafness was detected at a mean age of 11.14 months, with a standard deviation of 13.84 months. They received their implants at a mean age of 24.07 months with a standard deviation of 19.55 months. Two of the children were bilingual (using sign language and oral language). Four children used only oral language. The remaining eight children used oral language as their main communication mode and signs for support. One of those eight children was reported to not sign him/herself, but the parents used signs as support. A detailed description of the children with CI is provided in **Table 1**. The 38 children without HL were recruited from schools in Linköping, Sweden and attended a pre-school class. Four of the children without HL were excluded from the analysis. One because the test session was interrupted several times, one because s/he was not able/willing to finish all tasks and two because the parents did not fill in the Pragmatics Profile. The mean age of the remaining 34 children (17 girls) was 6.52 years with a standard deviation of 4.01 months. Thirty of the children took part in an intervention study and the results reported here are their pre-test results. The children without HL were tested individually, either during school time in a separate room or at home. The children received stickers for their participation. A consent form was signed by the caregivers. Both children and caregivers were told that they could drop out of the study at any point without giving a reason. The study was approved by the Linköping Research Ethics Committee (dnr 2015/308-31).

2.2. Material
2.2.1. Pragmatic Language Ability

The pragmatic language ability of the children was tested using the Pragmatics Profile from the Swedish version of the clinical evaluation of language fundamentals 4 screening test battery—CELF-IV (Semel et al., 2013). This measure has a high reliability for the tested age group (0.96). The Pragmatics Profile is a questionnaire containing 50 statements which the caregiver has to rate on a four-point scale. The 50 statements cover three different areas: Rituals and Conversational Skills—RCS (e.g., makes/responds to greetings to/from others), Asking for, Giving and Responding to Information—AGRI (e.g., asks for help from others appropriately), and Non-Verbal Communication—NCS (e.g., knows how someone is feeling based on non-verbal cues) (Pearson Education Inc., 2008a,b). For the rating scale, the following verbal items are used: Never, Sometimes, Often, Always. In this study the sub-scores for the three sub-measures have been used as measures in addition to the standard sum score.

Age	Detection of deafness	First implant (CI)	Unilateral/Bilateral	Communication mode	Schooling
8 years 9 months	2 days	24 months	Bilateral (30 months old)	Oral (10% sign)	Special
8 years 7 months	12 months	24 months	Bilateral	Oral (10% sign)	Special
8 years 11.5 months	44 months	60 months	Unilateral	Only oral	Special
8 years 4 months	30 months	36 months	Bilateral	Only oral	Special
7 years 9 months	Newborn	12 months	Bilateral (18 months old)	Oral (sign support)	Special
7 years 4 months	1 month	66 months	Unilateral	Oral (10% sign)	Special
6 years 6 months	2 months	8 months	Bilateral (20 months old)	Bilingual (Oral + sign)	Special
6 years 3 months	newborn	36 months	Unilateral	Bilingual (oral + sign)	Special
5 years 8 months	3 months	7 months	Bilateral (12 months old)	Only oral	Mainstream
5 years 9 months	1 month	8 months	Bilateral (18 months old)	Only oral	Mainstream
7 years	6 months	12 months	Bilateral (20 months old)	Oral (sign as support; not signing self)	Mainstream
5 years 8 months	3 months	7 months	Bilateral (15 months old)	Oral (sign as support)	Mainstream
5 years 7 months	24 months	28 months	Bilateral (38 months old)	Oral (sign as support)	Mainstream
7 years	6 months	9 months	Bilateral (11 months old)	Oral (sign as support)	Mainstream

2.2.2. Core Language Measures

2.2.2.1. Language comprehension

The Swedish version of TROG-2—Test for Reception of Grammar version 2 (Bishop, 2003; Eldblom and Sandberg, 2009), was used to assess children's language comprehension ability. This test consisted of 20 blocks of four sentences. The child saw four pictures and listened to a recorded sentence (e.g., "The star is not red"). The sentences were spoken by a native female speaker. The child was then instructed to point to the image corresponding to the sentence. The child got one point for every correct answer. After four wrong blocks in a row the test was terminated. A block was counted as being wrong if the child gave at least one wrong answer within the block. If the child did not answer at all twice in a row the test was terminated as well. In order to explain the task, two practice trials were used. The child received feedback on those two trials. The task was first continued after they gave the correct answer to both practice trials.

2.2.2.2. Vocabulary skills

To test the children's vocabulary skills, the Expressive Vocabulary task from the CELF-IV battery (Semel et al., 2013) was used. This is a picture naming task. The child was shown pictures (e.g., of an elephant) and asked to name them/ a specific part of the picture (e.g., the elephant's trunk). The task started with a demonstration

trial and two practice trials, after that 20 test trials followed. If the child was not able to name four pictures in a row the task was terminated. For every correctly named picture the child received one point.

2.2.3. Verbal Cognitive Measures

2.2.3.1. Verbal reasoning

To test verbal reasoning ability the Spoken Analogies sub-test of the Swedish ITPA-3 battery (Hammill et al., 2013) was used. The child listened to sentences of the following kind: "A dad is big, a baby is...," and was asked to fill in the missing word. This test consisted of two practice trials and 25 test trials. The test was terminated after three consecutive incorrect answers. For every correct word, the child got one point.

2.2.3.2. Verbal fluency

To test verbal fluency a semantically based fluency task was used (Benton and Hamsher, 1976). The child was asked to name as many animals as possible within 1 min. The child received one point for every animal.

2.2.3.3. Verbal working memory

The sentence completion and recall task from the SIPS battery (Wass et al., 2008) was used as a measure for verbal working

memory. The children heard a recorded sentence, spoken by a female speaker, with the last word missing (e.g., "A car has tires. A plane has..."). The child was then asked to fill in the missing word. A standard instruction was used and the child could practice using two examples before the real test started. There were six different levels, for level 1 children listen to two sentences, for level 2 they listen to three, for level 3 they listen to four, for level 4 they listen to two, for level 5 they listen to three, and for level 6 they listen to four sentences. The child got points for every word they recalled correctly. The test leader gave the first phoneme of the words as a cue if the child was not able to give an answer in the recall phase. If a cue was given the child only got 0.5 points for the recalled word.

2.3. Procedure

The Pragmatics Profile was handed out to the caregivers via the school or by the test leader and filled in at home. The rest of the testing took place at the respective school or at home. All children within the current study were tested by a speech and language pathologist or by a speech and language pathologist student in the last university term. If available in the test room a microphone and/or amplifier was used during the testing in order to enhance the speech signal for the oral test material. If these resources were not available, the child was asked if s/he wanted to use headphones to listen to the oral test material. All children preferred to use the laptop loudspeakers. The order of the tests was randomized and the test session was recorded using a Dictaphone.

2.4. Statistical Analysis

We used R (R Core Team, 2016) with the packages effsize (Torchiano, 2018) and cocor (Diedenhofen and Musch, 2015) for our analyses. To sort and edit the data for analysis, the packages dplyr (Wickham et al., 2019), tidyr (Wickham and Henry, 2019), and purrr (Henry and Wickham, 2019) were used. The graphs were made using the package ggplot2 (Wickham, 2009).

The alpha value was set to 0.05. All data was checked for normality and homogeneity of variance. To analyse the differences between the groups for the sum score for pragmatic language ability as well as for the sub-measure RCS, a Welch's t-test was used as homogeneity of variance was not given. For the other sub-measures, AGRI and NCS, a Student's t-test was used. To analyse the association between the language and verbal cognitive measures and pragmatic language ability, correlations have been calculated for the children without HL as well as for the children with CI. As the pragmatic measure was split into its sub-measures for the group comparison, this was also done for the correlations. For normally distributed data, Pearson correlations were calculated. For non-normally distributed data, Spearman correlations were calculated. The Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995) was used to decrease the false discovery for multiple comparisons.

3. RESULTS

There was no significant difference between the two groups in terms of their age, $t_{(14.41)} = 1.00, p = 0.333, d = 0.448$.

Age of implantation was not significantly correlated with the pragmatic language skills of the children with cochlear implants (CI), $\rho = -0.08, p = 0.609$. Additionally, the groups did not differ in terms of their non-verbal cognitive skills [Matrix test from the Wechsler Nonverbal Scale of Ability (Wechsler and Naglieri, 2007)], $t_{(46.00)} = 0.58, p = 0.567, d = 0.183$.

3.1. Group Differences in Pragmatic Language Ability

The sum score of the pragmatics profile of the children without HL and the children with cochlear implants was not significantly different, $t_{(17.07)} = 1.50, p = 0.152, d = 0.581$. However 5 out of 14 children with CI had scores below the age-norm, while only 2 out of 34 children without HL performed below the age-norm. All of the children with CI who performed below the age norm attended special school. Three of them were implanted early (≤ 24 months), one received the implant at 36 month of age and one was implanted late (66 months).

After comparing the two groups on the sum score, sub-scores for the three measures included in the Pragmatics Profile have been calculated. For the RCS sub-measure there was no significant difference between the children without HL and the children with CI, $t_{(16.33)} = 1.79, p = 0.093, d = 0.717$. For the AGRI sub-measure no significant difference was found between the groups, $t_{(46.00)} = 0.18, p = 0.858, d = 0.057$. For the NCS sub-measure a significant difference, $t_{(46.00)} = 2.22, p = 0.032, d = 0.704$, was found between the groups with children without HL performing better than the children with CI. This difference was still significant when excluding the two items "using a variation of tone of voice" and "recognizing that other people use different tone of voice" which could be argued are influenced by hearing with a CI, $t_{(46.00)} = 2.19, p = 0.033, d = 0.696$). For a graphical representation of the results (see **Figure 1**), means, standard deviation, and range are reported in **Table 2**.

3.2. Association Between Language and Verbal Cognitive Measures and Pragmatic Language Ability

3.2.1. Children With CI

All three pragmatic sub-measures: RCS, $\rho = 0.64, p = 0.040$, AGRI, $\rho = 0.74, p = 0.021$, and NCS, $\rho = 0.66, p = 0.040$, were significantly positively correlated with verbal fluency but no other measure.

3.2.2. Children Without HL

The RCS score of the children without HL was significantly positively correlated with their language comprehension ability, $r = 0.40, p = 0.033$, and their verbal reasoning ability, $\rho = 0.46, p = 0.017$. The AGRI score of the children without HL was not significantly correlated with any of the measured skills. The NCS score of the children without HL was significantly positively correlated with their vocabulary skill, $\rho = 0.49, p = 0.013$, and with their verbal reasoning ability, $\rho = 0.53, p = 0.009$.

FIGURE 1 | Pragmatic language skills: Raw scores for the children with CI and the children without HL. For the children with CI, green represents those attending special education and red represents those attending mainstream education.
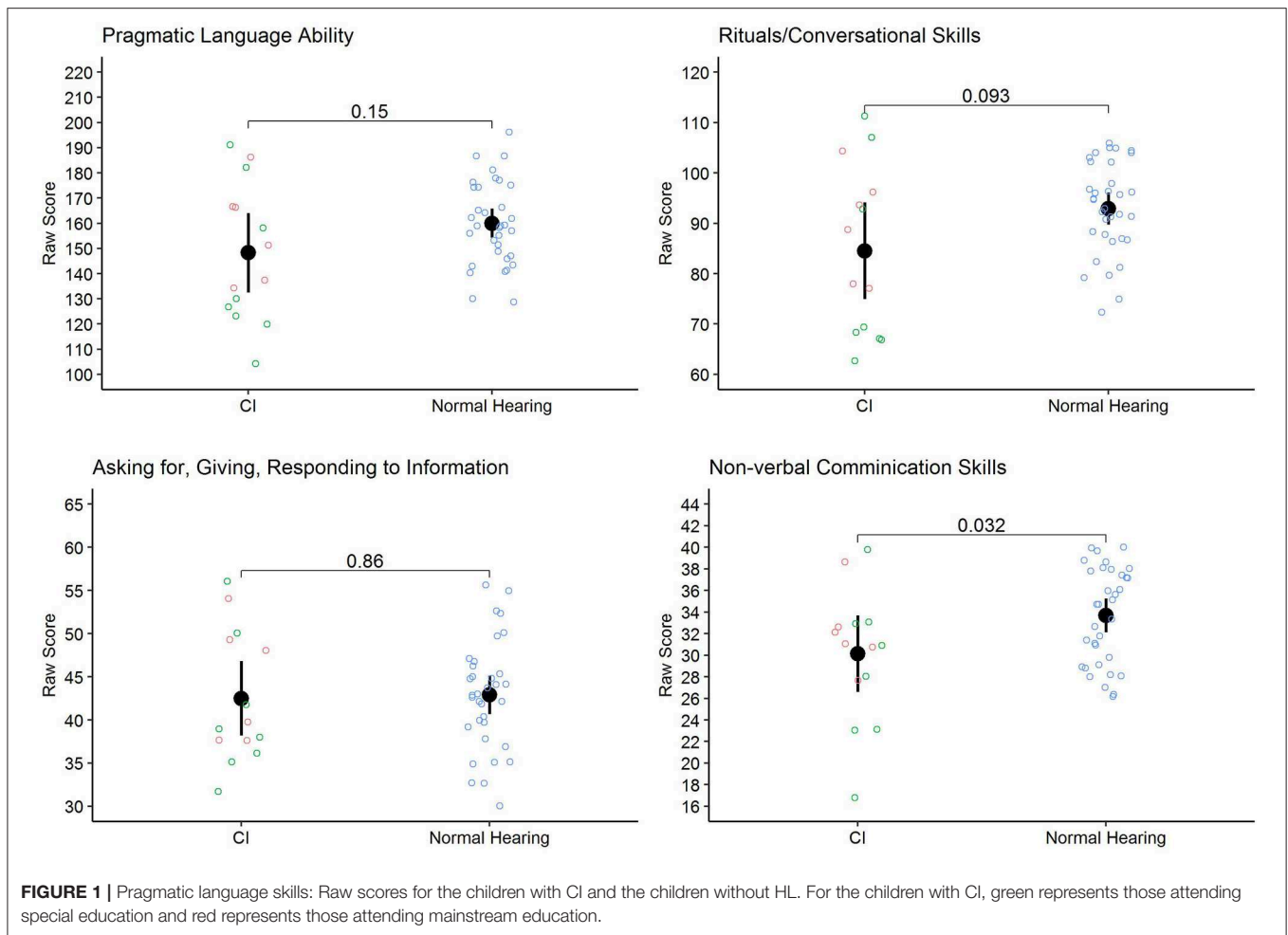
TABLE 2 | Descriptive data for the pragmatic language skills of children with CI and children without HL.

	Pragmatic profile sum score			RCS			AGRI			NCS		
	M	SD	Range	*M*	SD	Range	*M*	SD	Range	*M*	SD	Range
Children with CI	148.29	27.27	[104, 191]	84.5	16.62	[63, 111]	42.5	7.5	[32, 56]	30.14	6.14	[17, 40]
Children without HL	160	16.51	[129, 196]	92.91	9.12	[72, 106]	42.88	6.32	[30, 56]	33.67	4.5	[26, 40]

4. DISCUSSION

The aim of the current study was to analyse whether there is a difference between children using CI and children without HL in terms of their pragmatic language ability. In addition associations between pragmatic language ability and different verbal cognitive and language measures were analyzed, first to see which skills are possibly influencing the pragmatic language ability of children with CI and second to analyse differences in relationship patterns of children with CI and children without HL.

No significant difference was found between the children without HL and the children with CI for the sum score of

the pragmatics language measure. This is in accordance with the results found by Guerzoni et al. (2016). It differs, however, from results suggesting differences between children with CI and children without HL in terms of their pragmatic language ability (Jeanes, 2000; Most et al., 2010; Goberis et al., 2012; Rinaldi et al., 2013). The present study, as well as the study by Guerzoni et al. (2016), used parent rating, while in other studies the researches have rated the conversational skills of the children while communicating with a familiar adult (Most et al., 2010) or a peer (Jeanes, 2000; Ibertsson et al., 2009; Toe and Paatsch, 2013). It could be argued that parents tend to overestimate the competence of their child. However, the reliability of the measure

used in the present study has been reported to be high ($r^a = 0.96$) (Semel et al., 2013). In addition, other studies (Goberis et al., 2012; Rinaldi et al., 2013) have found poorer ratings for pragmatic language ability for deaf and hard of hearing children compared to children without HL even when being judged by their parents. Certainly, as the use of different measures of pragmatic language ability as well as different age groups, make it hard to directly compare results between studies. Different pragmatic language measures often include different domains of pragmatic language ability (Russell and Grizzle, 2008), which means that even if differences have been found in one specific measure, this does not necessarily mean that the two groups differ on all aspects of pragmatic language ability. In addition, although no significant differences could be seen on group level, 5 out of 14 children with CI (35.71%) performed below the age norm, while only 2 out of 34 children without HL (5.89%) performed below the age norm. All children with CI performing below the age-norm were attending special school. This is in accordance with the results of Thagard et al. (2011), who found a correlation between time spent in general education and pragmatic language ability. However, for both the results of the current study and the results found by Thagard et al. (2011) it is unclear if children having problems in the pragmatic language domain are the ones in need of special education or if being in special education leads to a delay in pragmatic language skills. Most et al. (2010) argue that one reason for the poorer pragmatic language ability of deaf and hard of hearing children might be that they have fewer possibilities to practice. This might especially be the case for children attending special education as they may have even fewer possibilities to engage in discourse with hearing peers or hearing adults who are not trained to talk to deaf and hard of hearing children, or to use sign support in comparison to children with CI attending mainstream education. Further studies should evaluate whether and how communication behavior in school and at home influences the pragmatic language ability of children with CI. Increased knowledge about this topic would benefit the development of intervention programs to improve the development of pragmatic language.

In the present study a significant difference between children with CI and children without HL was found for the Nonverbal Communication skills (NCS) sub-measure. Intuitively, non-verbal communication skills should not be influenced by having a hearing loss. In addition, Most et al. (2010) found no difference between children without HL and deaf and hard of hearing children neither on a non-verbal communication nor on a para-linguistic scale for pragmatic language ability. This is therefore a surprising result. Two items included in the NCS measure are: "varying tone of voice" and "recognizing varying tone of voice." It could be argued that those two skills are influenced by hearing with a CI. Comparing the two groups on the NCS scale without including those two items, however, still led to a significant group difference. This means hearing ability does not seem to be the main issue. A number of the items used in the sub-measure NCS for example "being able to recognize how somebody is feeling" or "understanding facial expressions" could be related to Theory of Mind development. The term "Theory of Mind" (ToM) refers to the ability to know about your own and other people's mental states. A child who can attribute beliefs, knowledge, emotions, desires, and intentions to other people and understands that those may differ from his/her own beliefs, knowledge, emotions, desires, and intentions has mastered ToM. This is usually the case around age five to six (Liu et al., 2018). A child with fully developed ToM skills should be able to recognize how somebody is feeling as well as understand facial expressions. Even the ability to recognize varying tone of voice is important, as distinct emotional states or intentions might be indicated by differences in tone of voice. Children who have fully developed ToM skills should therefore have higher scores on the NCS sub-measure. Studies have found that the development of ToM is often delayed in deaf and hard of hearing children (Peterson and Siegal, 2000; Lundy, 2002; Peterson, 2004; Ketelaar et al., 2012; Liu et al., 2018). In a meta-analysis done by Milligan et al. (2007), significant relations between language ability and ToM have been found. As children with CI are often delayed in terms of their language development, their delayed development of ToM is no surprise. Peterson (2004) found that children with CI perform on par with age-matched children with autism on tasks measuring ToM. The authors argue that restricted discourse between deaf and hard of hearing children and their hearing parents could be a reason for the delayed development. This is in accordance with the suggestion by Most et al. (2010) that pragmatic language development could be influenced by the opportunities to practice conversations.

No significant group difference was found for the Rituals and Conversational skills (RCS) and Asking for, Giving, and Responding to Information (AGRI) sub-measures. This is a promising result. Children with CI seem to be able to master these important parts of pragmatic language ability. For the AGRI sub-measure the result is in accordance with previous studies. This measure involves the abilities to ask for clarification, reacting to requests for clarification, explaining, and asking why things are like they are and why people do what they do, as well as a number of social skills, like asking for help, accepting apologies etc. Jeanes (2000) found that deaf and hard of hearing children using oral language used even more requests for clarifications than did children without HL and that they responded appropriately to requests for clarification. Furthermore, Antia et al. (2011) reported that the social skills of deaf and hard of hearing children are within their age norm. For the RCS sub-measure, the results are in accordance with studies suggesting that children with CI have conversational skills that are good enough to ensure a fluent conversation with a hearing peer (Toe and Paatsch, 2013). It differs, however, from other results suggesting difficulties of children with CI with verbal turn taking (Most et al., 2010; Paatsch and Toe, 2014). It should be mentioned that although the difference for the RCS sub-measure was not significant, there was a tendency for the children without HL to obtain higher scores than the children with CI, and the accompanying effect size was as high as it was for the NCS sub-measure. As the sub-measure included not only conversational skills but also the use of rituals, like saying hello or goodbye, it might be the case that some but not all of the abilities measured differed between the groups.

The correlation patterns between pragmatic language ability and language and verbal cognitive ability was different for

children with CI and the children without HL. For the children without HL language comprehension as well as verbal reasoning were positively correlated with the RCS scale. Furthermore, vocabulary skills and verbal reasoning were positively correlated with the NCS scale. As Matthews et al. (2018) point out it is often hard to distinguish between language skills, like language comprehension and vocabulary skills, and pragmatic language ability, a correlation between those skills was therefore expected. In addition the significant correlation between pragmatic language skills and verbal reasoning is in accordance to results from a study by Turkstra et al. (1996). Turkstra et al. (1996) suggest that inferential reasoning is important for pragmatic language ability and these two abilities are therefore associated.

For the children with CI, verbal fluency was the only skill correlated with all three sub-measures of pragmatic language ability. Previous studies (Kenett et al., 2013; Wechsler-Kashi et al., 2014) found that children with CI have a less developed semantic network. Semantic network here refers to the organization of words and different word meanings within the mental lexicon. Wechsler-Kashi et al. (2014) evaluated the organization of the semantic network of children with CI using verbal fluency tasks. The authors suggest that the children performed more poorly than children without HL as lexical organization is underdeveloped. The results from a computational analysis done by Kenett et al. (2013) support this view. Children with CI seem to have less strong connections between different words. Because of that, the activation of one word in their semantic network does not spread as much as it does for children without HL. The better the semantic network is developed, the better the performance on a verbal fluency task as more words are activated and their retrieval is therefore eased. The results from the current study suggest that children with CI who have a better developed semantic network have higher pragmatic language ability. A reason for these findings might be that a more structured network enables language to be used in a more flexible way. However, as only correlations have been used in the current study the causal direction is not clear. It could be that the quality and quantity of face-to-face interactions influence both the structure of the semantic network as well as pragmatic language ability. In addition, no correlations between verbal fluency and pragmatic language ability have been found for the children without HL. It might be that children with CI use different strategies for social communication that are more influenced by their semantic network. It might also be that the semantic network of children without HL is developed to a degree where more improvement does not influence pragmatic language ability anymore. More studies are needed to untangle the relationship pattern between hearing loss, verbal fluency, and pragmatic language ability.

4.1. Limitation of the Study

In the present study a small sample of children with CI was tested. It is possible that significant differences were therefore not detected for some of the variables. There was a tendency for a difference on the RCS measure and the accompanying effect size was fairly high. It is likely that a large sample size would have been needed to detect the significant difference of

the groups on the RCS scale. In the present study we have used a parent rating to measure pragmatic language skills. While this offers the possibility to get more insight into real-life pragmatic language skill compared to when analyzing conversations in the lab it also leads to some disadvantages. First it is a subjective measure. Further studies should aim to combine subjective and objective measures to get a better insight into the pragmatic language skills of children with CI. Second it is an inclusive measure. This makes it possible to get a broad overview over the current status of pragmatic language skill development but makes it hard to analyse which specific sub-skill might be causing differences. Differences between sub-skills might even go unnoticed if they are only measured by one or two items and differences therefore don't lead to significant differences on the sum measure or on the sub-measure level. Further research should aim to get data for a bigger group of children to be able to do a more precise item analysis to evaluate differences on a more detailed level. A further limitation of the study is the heterogeneity in terms of age of implantation of the children with CI. However, age of implantation was not correlated with pragmatic language skills and removing the two children with the oldest implantation age (60 and 66 months, leading to a SD of 11 months) did not change the results. Further studies should aim to collect more data concerning the pragmatic language skill of children with CI to be able to analyze the influence of age of implantation in more detail. A further limitation of the study is the missing information about the pre-implant hearing thresholds of the children. It might be the case that the degree of hearing loss influences the pragmatic language development. It is important to conduct more research on this topic to evaluate which other factors apart from verbal fluency are of importance for the pragmatic language development of children with CI. One additional factor might be the socioeconomic status of the parents. Rowe (2008) has reported a relation between child-directed speech and socioeconomic status of the parents. As child-directed speech is likely to influence pragmatic language development it is important for further studies to take the influence of this variable into consideration.

5. CONCLUSION

The results of the current study suggest that many children with CI show pragmatic language ability comparable to their hearing peers and in accordance to their age-norm. In the present study, significant differences were found on a measure connected to theory of mind, a skill found to be delayed in deaf and hard of hearing children. It has been suggested that the quality and quantity of face-to-face interactions influence both theory of mind and pragmatic language ability. Further studies are needed to analyse the influence of communication styles of care givers, teachers and peers on the development of pragmatic language ability in children. Results from the current study show that the development of the semantic network is associated with pragmatic language ability of children with CI. Verbal fluency was correlated with all three sub-measures of pragmatic language ability. The causal direction is unclear. It might be that children

with a better developed semantic network are able to use language in a more flexible way. Alternatively, quality and quantity of oral interaction might influence both the development of the semantic network and of pragmatic language ability. To be able to develop interventions for children with CI showing problems in the pragmatic language domain it is important to get more insight into the connection between conversation, verbal fluency, and pragmatic language ability.

DATA AVAILABILITY STATEMENT

The datasets generated for this study will not be made publicly available. It was ensured to the parents in the information letter that no data will be send to anyone not part of the research team. This was also included in the ethics application.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Regionala etikprövningsnämnden i Linköping. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

The study was prepared and designed by MS, BL, MG, IH, and MW. Acquisition of data was done by MG, IH together with Linn Hellgren and Elias Larsson (acknowledged). Analysis and interpretation of result was carried out by mainly MS, RE, and MW. The first draft of the manuscript was written by MS. All authors took part in critical revision of the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Antia, S. D., Jones, P., Luckner, J., Kreimeyer, K. H., and Reed, S. (2011). Social outcomes of students who are deaf and hard of hearing in general education classrooms. *Except. Child.* 77, 489–504. doi: 10.1177/001440291107700407

Bacso, S. A., and Nilsen, E. S. (2017). What's that you're saying? Children with better executive functioning produce and repair communication more effectively. *J. Cogn. Dev.* 18, 441–464. doi: 10.1080/15248372.2017.1336438

Bandurski, M., and Ga1kowski, T. (2004). The development of analogical reasoning in deaf children and their parents' communication mode. *J. Deaf Stud. Deaf Educ.* 9, 153–175. doi: 10.1093/deafed/enh018

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Benton, A., and Hamsher, K. (1976). *Multilingual Aphasia Examinantion.* Iowa City, IA: University of Iowa Press.

Bishop, D. V. M. (2003). *Test for Reception of Grammar: Version 2; TROG-2.* San Antonio, TX: Harcourt Assessment.

Blain-Briére, B., Bouchard, C., and Bigras, N. (2014). The role of executive functions in the pragmatic skills of children age 4-5. *Front. Psychol.* 5:240. doi: 10.3389/fpsyg.2014.00240

Bonifacio, S., Girolametto, L., Bulligan, M., Callegari, M., Vignola, S., and Zocconi, E. (2007). Assertive and responsive conversational skills of Italian–speaking late talkers. *Int. J. Lang. Commun. Dis.* 42, 607–623. doi: 10.1080/13682820601084386

Camarata, S. M., and Gibson, T. (1999). Pragmatic language deficits in attention-deficit hyperactivity disorder (ADHD). *Mental Retard. Dev. Disabil. Res. Rev.* 5, 207–214. doi: 10.1002/(SICI)1098-2779(1999)5:3<207::AID-MRDD7>3.3.CO;2-F

Channon, S., and Watts, M. (2003). Pragmatic language interpretation after closed head injury: relationship to executive functioning. *Cogn. Neuropsychiatry* 8, 243–260. doi: 10.1080/135468000344000002

Conti-Ramsden, G., and Botting, N. (2004). Social difficulties and victimization in children with SLI at 11 years of age. *J. Speech Lang. Hear. Res.* 47, 145–161. doi: 10.1044/1092-4388(2004/013)

Dammeyer, J. (2012). A longitudinal study of pragmatic language development in three children with cochlear implants. *Deafness Educ. Int.* 14, 217–232. doi: 10.1179/1464315412Z.00000000024

Diedenhofen, B., and Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE* 10:e0121945. doi: 10.1371/journal.pone.0121945

Douglas, J. M. (2010). Relation of executive functioning to pragmatic outcome following severe traumatic brain injury. *J. Speech Lang. Hear. Res.* 53, 365–382. doi: 10.1044/1092-4388(2009/08-0205)

Edwards, L., Figueras, B., Mellanby, J., and Langdon, D. (2011). Verbal and spatial analogical reasoning in deaf and hearing children: the role of grammar and vocabulary. *J. Deaf Stud. Deaf Educ.* 16, 189–197. doi: 10.1093/deafed/enq051

Eldblom, J., and Sandberg, K. (2009). *Svensk normering av TROG-2 för åldrarna 6–8 år samt undersökning av sambandet mellan grammatisk språkförståelse och narrativförståelse* (Degree project Bachelor's thesis). University of Gothenburg, Gothenburg, Sweden. Retrieved from: https://gupea.ub.gu.se/bitstream/2077/26897/1/gupea_2077_26897_1.pdf

Goberis, D., Beams, D., Dalpes, M., Abrisch, A., Baca, R., and Yoshinaga-Itano, C. (2012). The missing link in language development of deaf and hard of hearing children: pragmatic language development. *Semin. Speech Lang.* 33, 297–309. doi: 10.1055/s-0032-1326916

Guerzoni, L., Murri, A., Fabrizi, E., Nicastri, M., Mancini, P., and Cuda, D. (2016). Social conversational skills development in early implanted children. *Laryngoscope* 126, 2098–2105. doi: 10.1002/lary.25809

Hammill, D. D., Mather, N., and Roberts, R. (2013). *ITPA-3: Illinois Test of Psycholinguistic Abilities. Manual* (B. Holmgren, Trans.). Stockholm; Austin, TX: Hogrefe Psykologiförlaget AB, Pro-ED, Inc.

Helland, W. A., Lundervold, A. J., Heimann, M., and Posserud, M.-B. (2014). Stable associations between behavioral problems and language impairments across childhood–the importance of pragmatic language problems. *Res. Dev. Disabil.* 35, 943–951. doi: 10.1016/j.ridd.2014.02.016

Henry, L., and Wickham, H. (2019). *purrr: Functional Programming Tools.* R package version 0.3.2. Available online at: https://CRAN.R-project.org/package=purrr

Ibertsson, T., Hansson, K., Maki-Torkko, E., Willstedt-Svensson, U., and Sahlen, B. (2009). Deaf teenagers with cochlear implants in conversation with hearing peers. *Int. J. Lang. Commun. Dis.* 44, 319–337. doi: 10.1080/13682820802052067

Jeanes, R. C. (2000). The pragmatic skills of profoundly deaf children. *J. Deaf Stud. Deaf Educ.* 5, 237–247. doi: 10.1093/deafed/5.3.237

Kenett, Y. N., Wechsler-Kashi, D., Kenett, D. Y., Schwartz, R. G., Ben-Jacob, E., and Faust, M. (2013). Semantic organization in children with cochlear implants: computational analysis of verbal fluency. *Front. Psychol.* 4:543. doi: 10.3389/fpsyg.2013.00543

Ketelaar, L., Rieffe, C., Wiefferink, C. H., and Frijns, J. H. M. (2012). Does hearing lead to understanding? Theory of mind in toddlers and preschoolers with cochlear implants. *J. Pediatr. Psychol.* 37, 1041–1050. doi: 10.1093/jpepsy/jss086

Ketelaars, M. P., Jansonius, K., Cuperus, J., and Verhoeven, L. (2012). Narrative competence and underlying mechanisms in children with pragmatic language impairment. *Appl. Psycholinguist.* 33, 281–303. doi: 10.1017/S014271641100035X

Kim, O. H., and Kaiser, A. P. (2000). Language characteristics of children with ADHD. *Commun. Disord. Q.* 21, 154–165. doi: 10.1177/152574010002100304

Kronenberger, W. G., Pisoni, D. B., Henning, S. C., and Colson, B. G. (2013). Executive functioning skills in long-term users of cochlear implants: a case control study. *J. Pediatr. Psychol.* 38, 902–914. doi: 10.1093/jpepsy/jst034

Liu, M., Wu, L., Wu, W., Li, G., Cai, T., and Liu, J. (2018). The relationships among verbal ability, executive function, and theory of mind in young children with cochlear implants. *Int. J. Audiol.* 57, 881–888. doi: 10.1080/14992027.2018.1498982

Longobardi, E., Lonigro, A., Laghi, F., and O'Neill, D. K. (2017). Pragmatic language development in 18- to 47-month-old italian children: a study with the language use inventory. *First Lang.* 37, 252–266. doi: 10.1177/0142723716689273

Loukusa, S., Leinonen, E., and Ryder, N. (2007). Development of pragmatic language comprehension in finnish-speaking children. *First Lang.* 27, 279–296. doi: 10.1177/0142723707076568

Lundy, J. E. B. (2002). Age and language skills of deaf children in relation to theory of mind development. *J. Deaf Stud. Deaf Educ.* 7, 41–56. doi: 10.1093/deafed/7.1.41

Martin, I., and McDonald, S. (2003). Weak coherence, no theory of mind, or executive dysfunction? Solving the puzzle of pragmatic language disorders. *Brain Lang.* 85, 451–466. doi: 10.1016/S0093-934X(03)00070-1

Matthews, D. (2014). *Pragmatic Development in First Language Acquisition, Vol. 10.* Amsterdam: John Benjamins Publishing Company.

Matthews, D., Biney, H., and Abbot-Smith, K. (2018). Individual differences in children's pragmatic ability: a review of associations with formal language, social cognition, and executive functions. *Lang. Learn. Dev.* 14, 186–223. doi: 10.1080/15475441.2018.1455584

Milligan, K., Astington, J. W., and Dack, L. A. (2007). Language and theory of mind: meta-analysis of the relation between language ability and false-belief understanding. *Child Dev.* 78, 622–646. doi: 10.1111/j.1467-8624.2007.01018.x

Most, T., Shina-August, E., and Meilijson, S. (2010). Pragmatic abilities of children with hearing loss using cochlear implants or hearing aids compared to hearing children. *J. Deaf Stud. Deaf Educ.* 15, 422–437. doi: 10.1093/deafed/enq032

Murphy, S. M., Faulkner, D. M., and Farley, L. R. (2014). The behaviour of young children with social communication disorders during dyadic interaction with peers. *J. Abnorm. Child Psychol.* 42, 277–289. doi: 10.1007/s10802-013-9772-6

Norbury, C. F., and Bishop, D. V. M. (2002). Inferential processing and story recall in children with communication problems: a comparison of specific language impairment, pragmatic language impairment and high-functioning autism. *Int. J. Lang. Commun. Dis.* 37, 227–251. doi: 10.1080/13682820210136269

Paatsch, L. E., and Toe, D. M. (2014). A comparison of pragmatic abilities of children who are deaf or hard of hearing and their hearing peers. *J. Deaf Stud. Deaf Educ.* 19, 1–19. doi: 10.1093/deafed/ent030

Paradis, M. (1998). The other side of language: pragmatic competence. *J. Neurolinguist.* 11, 1–10. doi: 10.1016/S0911-6044(98)00001-3

Pearson Education, Inc. (2008a). *CELF-4: Scoring Assistant Sample Reports.* Retrieved from: https://www.pearsonclinical.co.uk/Psychology/ChildCognitionNeuropsychologyandLanguage/ChildLanguage/ClinicalEvaluationofLanguageFundamentals-FourthEditionUK(CELF-4UK)/Resources/CELF4SampleReport.pdf

Pearson Education, Inc. (2008b). *CELF-4: Technical Report.* Retrieved from: http://images.pearsonassessments.com/images/tmrs/tmrs_rg/CELF_4_Tech_Report.pdf

Peterson, C. C. (2004). Theory-of-mind development in oral deaf children with cochlear implants or conventional hearing aids. *J. Child Psychol. Psychiatry* 45, 1096–1106. doi: 10.1111/j.1469-7610.2004.t01-1-00302.x

Peterson, C. C., and Siegal, M. (2000). Insights into theory of mind from deafness and autism. *Mind Lang.* 15, 123–145. doi: 10.1111/1468-0017.00126

R Core Team (2016). *R: A Language and Environment for Statistical Computing.* Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Rinaldi, P., Baruffaldi, F., Burdo, S., and Caselli, M. C. (2013). Linguistic and pragmatic skills in toddlers with cochlear implant: linguistic and pragmatic skills in CI. *Int. J. Lang. Commun. Dis.* 48, 715–725. doi: 10.1111/1460-6984.12046

Rowe, M. L. (2008). Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. *J. Child Lang.* 35, 185–205. doi: 10.1017/S0305000907008343

Russell, R. L., and Grizzle, K. L. (2008). Assessing child and adolescent pragmatic language competencies: toward evidence-based assessments. *Clin. Child Fam. Psychol. Rev.* 11, 59–73. doi: 10.1007/s10567-008-0032-1

Semel, E., Wiig, E. H., and Secord, W. A. (2013). *CELF-4, Clinical Evaluatuon of Language Fundamentals - Fourth Edition. Svensk Version* (M. Garsell, Trans.). Bloomington, IN: NCS Pearson, Inc.

Staikova, E., Gomes, H., Tartter, V., McCabe, A., and Halperin, J. M. (2013). Pragmatic deficits and social impairment in children with ADHD. *J. Child Psychol. Psychiatry* 54, 1275–1283. doi: 10.1111/jcpp.12082

Thagard, E. K., Hilsmier, A. S., and Easterbrooks, S. R. (2011). Pragmatic language in deaf and hard of hearing students: correlation with success in general education. *Am. Ann. Deaf* 155, 526–534. doi: 10.1353/aad.2011.0008

Toe, D., Beattie, R., and Barr, M. (2007). The development of pragmatic skills in children who are severely and profoundly deaf. *Deafness Educ. Int.* 9, 101–117. doi: 10.1179/146431507790560011

Toe, D. M., and Paatsch, L. E. (2013). The conversational skills of school-aged children with cochlear implants. *Cochlear Implants Int.* 14, 67–79. doi: 10.1179/1754762812Y.0000000002

Torchiano, M. (2018). *effsize: Efficient Effect Size Computation.* R package version 0.7.4. Available online at: https://CRAN.R-project.org/package=effsize

Turkstra, L. S., McDonald, S., and Kaufmann, P. M. (1996). Assessment of pragmatic communication skills in adolescents after traumatic brain injury. *Brain Injury* 10, 329–346. doi: 10.1080/026990596124359

Volden, J., Coolican, J., Garon, N., White, J., and Bryson, S. (2009). Brief report: pragmatic language in autism spectrum disorder: relationships to measures of ability and disability. *J. Autism Dev. Dis.* 39, 388–393. doi: 10.1007/s10803-008-0618-y

Wass, M., Ibertsson, T., Lyxell, B., SahléN, B., HäLlgren, M., Larsby, B., et al. (2008). Cognitive and linguistic skills in swedish children with cochlear implants - measures of accuracy and latency as indicators of development. *Scand. J. Psychol.* 49, 559–576. doi: 10.1111/j.1467-9450.2008.00680.x

Wechsler, D., and Naglieri, J. A. (2007). *WNV: Svensk Version: Manual* (M. Garsell, Trans.). San Antonio, TX: Harcourt Assessment, Inc.

Wechsler-Kashi, D., Schwartz, R. G., and Cleary, M. (2014). Picture naming and verbal fluency in children with cochlear implants. *J. Speech Lang. Hear. Res.* 57, 1870–1872. doi: 10.1044/2014_JSLHR-L-13-0321

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis.* New York, NY: Springer-Verlag New York.

Wickham, H., François, R., Henry, L., and Müller, K. (2019). *dplyr: A Grammar of Data Manipulation.* R package version 0.8.1. Available online at: https://CRAN.R-project.org/package=dplyr

Wickham, H., and Henry, L. (2019). *tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions.* R package version 0.8.3. Available online at: https://CRAN.R-project.org/package=tidyr

Auditory, Cognitive, and Linguistic Factors Predict Speech Recognition in Adverse Listening Conditions for Children With Hearing Loss

Ryan W. McCreery[1]*, *Elizabeth A. Walker*[2], *Meredith Spratford*[1], *Dawna Lewis*[1] *and Marc Brennan*[3]

[1] *The Audibility Perception and Cognition Laboratory, Boys Town National Research Hospital, Omaha, NE, United States,* [2] *Pediatric Audiology Laboratory, Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, United States,* [3] *Amplification and Perception Laboratory, Department of Special Education and Communication Disorders, University of Nebraska, Lincoln, NE, United States*

Objectives: Children with hearing loss listen and learn in environments with noise and reverberation, but perform more poorly in noise and reverberation than children with normal hearing. Even with amplification, individual differences in speech recognition are observed among children with hearing loss. Few studies have examined the factors that support speech understanding in noise and reverberation for this population. This study applied the theoretical framework of the Ease of Language Understanding (ELU) model to examine the influence of auditory, cognitive, and linguistic factors on speech recognition in noise and reverberation for children with hearing loss.

Design: Fifty-six children with hearing loss and 50 age-matched children with normal hearing who were 7–10 years-old participated in this study. Aided sentence recognition was measured using an adaptive procedure to determine the signal-to-noise ratio for 50% correct (SNR50) recognition in steady-state speech-shaped noise. SNR50 was also measured with noise plus a simulation of 600 ms reverberation time. Receptive vocabulary, auditory attention, and visuospatial working memory were measured. Aided speech audibility indexed by the Speech Intelligibility Index was measured through the hearing aids of children with hearing loss.

Results: Children with hearing loss had poorer aided speech recognition in noise and reverberation than children with typical hearing. Children with higher receptive vocabulary and working memory skills had better speech recognition in noise and noise plus reverberation than peers with poorer skills in these domains. Children with hearing loss with higher aided audibility had better speech recognition in noise and reverberation than peers with poorer audibility. Better audibility was also associated with stronger language skills.

Conclusions: Children with hearing loss are at considerable risk for poor speech understanding in noise and in conditions with noise and reverberation. Consistent with the predictions of the ELU model, children with stronger vocabulary and working memory abilities performed better than peers with poorer skills in these domains. Better

aided speech audibility was associated with better recognition in noise and noise plus reverberation conditions for children with hearing loss. Speech audibility had direct effects on speech recognition in noise and reverberation and cumulative effects on speech recognition in noise through a positive association with language development over time.

Keywords: children, hearing loss, noise, reverberation, speech recognition, hearing aids

INTRODUCTION

Children spend a considerable amount of time listening in environments with suboptimal acoustics, including high levels of background noise and reverberation (Knecht et al., 2002; Crukley et al., 2011). Because noise and reverberation are ubiquitous, auditory learning, and socialization frequently occur in conditions with an acoustically degraded speech signal. The ability to understand degraded speech is an important developmental skill that does not reach full maturity until adolescence in children with typical hearing (Johnson, 2000; Corbin et al., 2016). The protracted developmental time course for speech recognition in adverse listening conditions in typically developing children has been attributed to the parallel maturation of cognitive and linguistic skills (Sullivan et al., 2015; McCreery et al., 2017; MacCutcheon et al., 2019).

Children with hearing loss face even greater challenges than their peers with typical hearing for understanding speech in adverse acoustic environments. Noise and reverberation frequently co-occur in classrooms and other listening environments experienced by children (Klatte et al., 2010). Whereas, some children with hearing loss who have well-fitted hearing aids or cochlear implants can understand speech in quiet as well as their peers with normal hearing (McCreery et al., 2015), very few children with hearing loss reach comparable levels of performance as their peers with normal hearing in noise (Goldsworthy and Markle, 2019) or reverberation (Neuman et al., 2012). The persistence of speech recognition deficits for children with hearing loss even after access to the signal has been restored with amplification raises questions about the mechanisms that affect the ability to understand degraded speech in everyday listening environments. The main goal of this study was to examine the factors that predicted individual differences in speech recognition in noise and in noise with reverberation by children with hearing loss.

The loss of audibility associated with hearing loss is a primary contributor to difficulties understanding speech in noise or reverberation among children with hearing loss. Blamey et al. (2001) demonstrated that speech recognition for children with hearing loss was strongly related to the child's pure tone average threshold with poorer recognition for children with greater degrees of hearing loss. Children with hearing loss who have better aided detection thresholds for pure tones also had better open-set word recognition (Davidson and Skinner, 2006). However, detection thresholds in quiet may not reflect individual differences in speech recognition in noise, so more recent studies have attempted to use measures of aided speech audibility at conversational levels as a predictor. Speech audibility

is often quantified using the Speech Intelligibility Index (SII; ANSI S3.5-1997), which estimates the proportion of the long-term average speech spectrum that is audible. Because the SII directly measures the audibility of speech for an individual, it can be considered a more accurate measure for predicting speech recognition than thresholds from the audiogram. The degree to which hearing loss limits speech audibility has been explored as a predictor of unaided (Scollie, 2008) and aided (Stiles et al., 2012; McCreery et al., 2015, 2017) speech recognition for children. In general, studies have found that children with hearing loss who have greater aided audibility for speech have better aided speech recognition in quiet (McCreery et al., 2015) and in noise (McCreery et al., 2017; Walker et al., 2019).

Mixed findings from the adult literature make predicting the effects of amplification on speech recognition in noise and reverberation for children with hearing loss more challenging. In one study, adults performed better in reverberation with wide-dynamic range compression (WDRC) amplification compared to linear amplification on a sentence recognition task (Shi and Doherty, 2008), suggesting that the increased audibility that occurs with WDRC compared to linear amplification may have enhanced listeners' speech recognition in reverberation. In contrast, Reinhart and Souza (2016) found that recognition for adults with hearing loss in reverberation was best when the release times for the amplitude compression were slower and more linear, as faster compression release times resulted in greater distortion of the temporal envelope of the speech signal. Children's hearing aids are often fitted to optimize speech audibility using WDRC (Scollie et al., 2005), but children are also likely to be more susceptible than adults to distortions of temporal and spectral cues in the speech signal (Hall et al., 2012). Thus, the effects of maximizing audibility with amplification on speech recognition in noise and reverberation for children remain difficult to predict without being directly examined.

Multiple studies have demonstrated that the addition of reverberation negatively affects speech recognition in school-age children with normal hearing (Neuman and Hochberg, 1983; Bradley and Sato, 2008; Wróblewski et al., 2012). In a study of children with normal hearing, Nabelek and Robinson (1982) reported that children required up to 20 dB higher sensation level to reach similar levels of performance as adults when listening in reverberation. A later study by Johnson (2000) found that while the developmental trajectory for speech recognition in noise and reverberation for children with normal hearing often did not reach adult levels of performance until the teenage years, the sensation level did not affect performance across age. Thus, the effects of increasing audibility on speech recognition in noise and reverberation for children have been

mixed. However, speech recognition in listening conditions with noise and reverberation for children with hearing loss has not been widely studied. Forty years ago, Finitzo-Hieber and Tillman (1978) conducted what remains one of the only examinations of the effects of noise and reverberation on children with hearing loss using hearing aids. While all children had poorer speech recognition in degraded conditions, the combined effects of noise and reverberation created disproportionate difficulty understanding speech for the children with hearing loss. However, the children with hearing loss in the Finitzo-Hieber and Tillman study used monaural, linear amplification. Also, because the study was conducted prior to the development of hearing aid verification methods, the level of audibility provided by the hearing aids for the children in that study was not specified. The implications of the results for children with hearing loss who use bilateral hearing aids with WDRC are difficult to generalize from this study.

Cognitive and linguistic skills are also likely to support speech recognition in noise and reverberation for children with hearing loss. The Ease of Language Understanding model (ELU; Rönnberg et al., 2013) is a model of language processing that suggests that listeners rely on their knowledge of language and cognitive skills like working memory and attention to understand speech in degraded conditions. The predictions of the ELU model that children with greater working memory capacities have better speech in noise than peers with reduced working memory capacities have been confirmed by some previous studies of speech recognition in noise for children with normal hearing (Stiles et al., 2012; Sullivan et al., 2015; McCreery et al., 2017; MacCutcheon et al., 2019). However, work by Magimairaj et al. (2018) did not find an association between language or working memory and sentence recognition in babble noise for children with normal hearing. The ELU model has also been extended to predict speech recognition in children with hearing loss (McCreery et al., 2015). Children with better working memory and language skills often have better speech understanding in noise than children with poorer working memory and language. Based on the predictions of the ELU model, it is reasonable to predict that cognitive and linguistic skills would be helpful for listening to speech degraded by noise and reverberation.

One potential mechanism that explains the link between cognitive and linguistic skills and listening in noise and reverberation is related to children's increased susceptibility to informational masking (Brungart et al., 2001; see Leibold and Buss, 2019 for a review). Reverberation is a particularly challenging masking signal because the reverberant signal can cause energetic masking, where the energy of the reverberant signal overlaps with the target signal in the auditory system. In addition, reverberation contains speech-like spectral and temporal cues that are similar to speech and create uncertainty about the stimuli, which are both characteristics of informational masking (Durlach et al., 2003). Children are also less likely to benefit from temporal fluctuations in a masker than are adults (Hall et al., 2012), which has been attributed to the development the development of temporal processing. To date, few studies have examined the cognitive and linguistic contributions to the development of informational masking.

However, some evidence from adults suggests that listeners with better working memory skills may be less susceptible to distortions of the auditory signal from amplification or hearing loss (see Akeroyd, 2008, for a review). Other researchers have suggested that increased susceptibility to informational masking among children may be related to difficulty segregating the target signal from the masking signal (Leibold et al., 2016). Greater susceptibility to informational masking in children has been attributed to deficits in auditory attention (Allen and Wightman, 1995; Corbin et al., 2016), but the effects of individual attention skills on conditions that produce informational masking have not been directly studied in children to our knowledge.

Potential interactions may exist between amplification and linguistic and cognitive skills that could influence the relationship of these factors with speech recognition in noise and reverberation. Recent evidence suggests that children who have better audibility through their hearing aids not only have better speech recognition (McCreery et al., 2015), but also have stronger language skills (Tomblin et al., 2014, 2015). This relationship between audibility and language development suggests that audibility has immediate effects related to a listener's access to the acoustic signal and cumulative effects related to its long-term influence on language. These relationships could make it difficult to disambiguate the effects of audibility related to access to the speech signal from the cumulative effects of audibility on language skills that are likely to support listening in noise and reverberation. Mediation models (Baron and Kenny, 1986) have been used to examine the pattern of associations between outcomes and predictors that are inter-related. A recent study by Walker et al. (2019) used a mediation analysis to determine if the effects of audibility on speech recognition for a gated word recognition task were direct or mediated by the relationship between audibility and language skills. The results suggested that audibility was related to both language and gated speech recognition, supporting both immediate, and cumulative influences of audibility. A similar approach was used to attempt to disambiguate that complex relationship in the current study.

The overall goal of this study was to examine factors that predicted individual differences in aided speech recognition for children with hearing loss and a group of children with typical hearing matched for age and socioeconomic status (indexed by maternal education level). Three research questions were examined:

- Does listening in noise and reverberation present additional challenges to children with hearing loss compared to peers with normal hearing? Based on previous research, we predicted greater difficulty listening in noise and reverberation for children with hearing loss than peers with normal hearing.
- Consistent with the ELU model, do linguistic and cognitive abilities predict individual differences in speech recognition in noise and reverberation for children with normal hearing and children with hearing loss? We predicted that our results would be consistent with the ELU model in that children with stronger working memory, language, and auditory attention skills would have better speech recognition in

noise and reverberation than children with poorer skills in these domains.

- Does aided speech audibility have a direct relationship with speech recognition in noise and reverberation or is the relationship mediated by the child's language skills? Based on previous research, we anticipated that the relationship between audibility and aided speech recognition in noise and reverberation would include direct and mediated paths.

METHOD

Participants

Children with normal hearing ($n = 50$) and children with mild-to-severe hearing loss ($n = 56$) participated in the experiment. Children were recruited from research centers at Boys Town National Research Hospital (Omaha, Nebraska) and the University of Iowa (Iowa City, Iowa) and the surrounding areas as part of a longitudinal study of developmental outcomes for children with bilateral mild-to-severe hearing loss. **Table 1** shows the demographic characteristics of the children in the sample. Data collection occurred during the summer following either 1st or 3rd grade when children were 7–9 years-old. All children were from homes where spoken English was the primary language and did not have other diagnosed developmental conditions at the time of enrollment in the study. All 56 of children with hearing loss wore bilateral hearing aids. The study was approved by the Institutional Review Board of Boys Town National Research Hospital.

Procedure

Audiometry and Hearing Aid Assessment

Hearing sensitivity was assessed for all children using age-appropriate behavioral audiometric assessment techniques. Children with normal hearing were screened via air conduction with headphones at 20 dB HL at 500, 1,000, 2,000, and 4,000 Hz. For children with hearing loss, air- and bone-conduction audiometric thresholds were measured at octave frequencies from 250 to 8,000 Hz using either Etymotic ER-3A insert or TDH-49 circumaural earphones. The thresholds at 500, 1,000, 2,000, and 4,000 Hz were averaged to calculate the pure-tone average (PTA) for each ear, and the PTA for the better-ear was used to represent degree of hearing loss in the statistical analyses.

For children with hearing loss, audibility for the long-term average speech spectrum for a 65 dB SPL input was measured at their daily use settings with the Audioscan Verifit probe microphone system (Dorchester, Ontario). The Verifit calculated the aided Speech Intelligibility Index (SII; ANSI S3.5-1997) for the 65 dB SPL speech signal for each ear as an estimate of speech audibility in quiet. The sensation level of the hearing aid output in 1/3 octave was measured and multiplied by an importance weight specified in the ANSI standard to represent the amount of speech information in each band. The weighted audibility across bands was added together to calculate the weighted proportion of speech that was audible through the child's hearing aids. The SII was expressed as a value between 0 and 1, where 0 indicates that none of the speech spectrum was audible, and 1 indicates

TABLE 1 | Participant characteristics.

Group	Children with normal hearing	Children with hearing loss	Statistical tests
Number	50	56	
Sex	Female = 23, Male =27	Female = 23, Male = 33	$X^2 = 1.33$, $p = 0.62$
Maternal education level	15.5 years of education	15.3 years of education	$t = 1.12$, $p = 0.75$
Grade	1st grade = 32; 3rd grade = 18	1st grade = 28; 3rd grade = 28	
Age (Years)	1st grade—M = 7.5 (SD = 0.4); 3rd grade–M = 9.0 (SD = 0.35)	1st grade—M = 7.5 (SD = 0.6); 3rd grade M = 9.0 (SD = 0.4)	$t = 0.07$, $p = 0.96$ $t = 0.01$, $p = 0.98$
PPVT standard score	Mean = 112.7	Mean = 111.6	$t = 1.08$, $p = 0.35$
NEPSY attention scaled score	Mean = 9.1	Mean = 9.0	$t = 0.06$, $p = 0.93$
AWMA OOO standard score	Mean = 110.9	Mean = 111.7	$t = 1.24$, $p = 0.18$
Age of confirmation of hearing loss	N/A	Mean= 12.7 months, Median = 3 months	
Age of HA fitting	N/A	Mean = 16.4 months Median = 6 months	
Better-ear PTA	N/A	Mean = 45.6 dB HL Median = 44 dB HL	
Better-ear aided SII	N/A	Mean = 0.77 Median = 0.83	
Average hours of HA use per day	N/A	Mean = 9.88 Median = 10.5	
RMS error		Mean left ear =5.09; Mean right ear = 5.27	

PPVT, Peabody Picture Vocabulary Test; NEPSY Attention, NEPSY Auditory Attention subtest; AWMA OOO, Automated Working Memory Assessment Odd-One-Out subtest; HL, Hearing Level; PTA, Pure-tone average hearing thresholds at 500, 1,000, 2,000, and 4,000 Hz; SII, Speech-intelligibility index for average speech (60 or 65 dB SPL) at one meter with hearing aids; Hearing aid use based on hearing aid data logging or parent report. RMS error is the geometric mean of the deviations of the hearing aid output from prescriptive target at 500, 1,000, 2,000, and 4,000 Hz. Listener sex was compared using a X^2 test and all other group comparisons were made using an independent samples t-test.

that the entire speech spectrum was audible. Aided SII data for the children who wore hearing aids are shown in **Table 1**. The output of the hearing aid was measured in the child's ear canal whenever possible. If the child was uncooperative, the child's real-ear-to-coupler-difference (RECD) was measured for each ear. The RECD was then applied to measures of hearing aid output in the 2 cm^3 coupler on the Verifit system to simulate the output of the hearing aid in the child's ear canal. The proximity of each child's fitting to Desired Sensation Level (DSL; Scollie et al., 2005) prescriptive targets at 500, 1,000, 2,000, and 4,000 Hz was measured and the geometric mean of the errors was taken to estimate the root-mean-square (RMS) error for each hearing aid. The average RMS error for each ear is reported in **Table 1**. Average hours of daily hearing aid use were assessed to describe the consistency of hearing aid use for children in the sample. Hearing aid use was estimated by either parent report or the automatic data logging system in the hearing aids. Because the hearing aids used by the children in the study were fitted by audiologists who were not associated with the study, information about specific signal processing features activated in the hearing aids from the fitting software were not available with one exception. Information about frequency lowering was collected. Frequency lowering (Phonak Sound Recover or Oticon Speech Rescue) was activated in 40% of the 1st grade fittings and 45% of the 3rd grade fittings.

Language, Working Memory, and Auditory Attention

Each child completed standardized measures of language, working memory, and executive function. Children with hearing aids used their hearing aids during these assessments. Receptive vocabulary skills were assessed using the Peabody Picture Vocabulary Test (PPVT-IV; Dunn and Dunn, 2007). Children were presented with target words and then pointed to a picture in a set of four possible pictures that best corresponds to the target word. Visuospatial working memory was assessed using the Odd-One-Out subtest of the Automated Working Memory Assessment (AWMA; Alloway et al., 2008). This visuospatial working memory task was selected to minimize the effects of differences in hearing and language abilities on working memory performance between children with normal hearing and children with hearing loss. For the Odd-One-Out, children are visually presented with sets of three complex shapes. One of the shapes is different than the other two shapes. The child points to the shape that does not match the other shapes. The child is then asked to remember the position of the different shape on a screen with three blank boxes. The number of sets of shapes increases throughout the task until the child misses a specific number of sets across consecutive blocks of trials. The PPVT and AWMA yielded raw scores and standard scores with a normative mean = 100 and a standard deviation = 15. Each child also completed the Auditory Attention subtest of the Developmental Neuropsychological Assessment (NEPSY-II; Brooks et al., 2009), which measured the ability to sustain auditory attention. During the Auditory Attention subtest of the NEPSY, children listen to a recorded series of words that are presented at a rate of one per second. The child must attend to the words and touch a red circle each time that they hear the word "red," but not for other

words. The total score is based on a combination of accuracy for "red" trials, where an incorrect response would be not touching the red circle when "red" was presented, and errors where the child touched the red button for any other word. The Auditory Attention subtest yields a scaled score with a normative mean = 10 and a standard deviation = 3.

Adaptive Speech Recognition Task

The stimuli for the speech recognition task were 250 low-predictability sentences described in a previous study (McCreery et al., 2017). Each sentence included four key words that were within the lexicon of 5-year-old children based on a child lexical database (Storkel and Hoover, 2010). The sentences were constructed with a simple, subject-verb-adjective-object syntactic structure. The sentences were recorded at 44,100 Hz sampling rate with 32-bit resolution as spoken by a female, native-English talker. An unmodulated speech-spectrum noise was created by taking the Fast Fourier Transform (FFT) of the concatenated set of sentences, randomizing the phase at each time point, and taking the inverse-FFT of the resulting signal to generate a noise that matched the long-term average spectrum of the talker. The noise had a cosine-squared 100 ms ramp-up before and ramp-down after each stimulus. For the simulated reverberation conditions, the target sentences and masker were convolved with the binaural room impulse response for a small classroom (20" × 20") with a reverberation time of 600 ms (RT60), which was the modal reverberation time for a sample of classrooms in a study of classroom acoustics that included children in the age range of this experiment (Dockrell and Shield, 2006).

Children were seated in a sound-attenuating audiometric test room or mobile van with the examiner. The speech and noise were presented from two speakers co-located in the front of the child at a position where the speech was calibrated at 65 dB SPL using a 1,000 Hz calibration tone with the same RMS level as the speech stimuli. Presentation in sound field was used so children could listen through their hearing aids. Children with hearing loss listened to conditions at their normal hearing aid use settings. Sentences were chosen at random without replacement for each trial. The level of the masker was adapted using a one-down, one-up procedure (Levitt, 1971) with custom software to estimate the signal-to-noise ratio where each child got 50% of the sentences correct (SNR50). The starting SNR for each track was 20 dB, the initial step-size was 5 dB, and after two reversals, the step size decreased to 3 dB for the final 6 reversals. Because the stopping rule for the adaptive track was based on the number of reversals, the number of sentences presented for each track varied across children from 20 to 42 with an average of 25 sentences per condition. The examiner scored responses during the task. Noise and noise + reverberation conditions were completed by each child in random order.

Statistical Analyses

All statistical analyses and data visualization were completed using R Statistical Software (R Development Core Team, 2018). Data visualization was completed using the ggplot2 (Wickham and Chang, 2016) and sjPlot (Lüdecke and Schwemmer, 2018) packages for R. Descriptive statistics were generated for each

predictor and outcome measure. For language and cognitive measures, standard scores were used to compare children in the experiment to the normative sample for each test, whereas raw scores were used to represent each construct in statistical analyses with age or grade as a covariate. Pearson correlations were calculated between predictors and outcomes to show the pattern of bivariate relationships between the predictors and outcomes for the study to support the inclusion of predictors in the multivariate models. For all the children, a linear mixed model was conducted to test the effects of linguistic and cognitive skills on speech recognition in noise and noise + reverberation using the lme4 package (Bates et al., 2015) for R. All possible interaction terms were assessed for each model, but only interactions that met the criterion for statistical significance ($p < 0.05$) are reported for simplicity with the exception of the subject type (NH vs. HL) interaction with reverberation, which was specifically hypothesized. The effects of aided audibility on language, working memory, and SNR50 in reverberation were also assessed for children with hearing loss with linear regression using a mediation analysis approach. The normality of each model's residuals was assessed to identify potential violations of statistical assumptions. To control for Type I error rate for statistical tests involving multiple comparisons, the p-values were adjusted using the False Discovery Rate procedure (Benjamini and Hochberg, 1995).

RESULTS

Figures 1–**3** compare the standard scores between children with hearing loss and children with normal hearing on the PPVT, NEPSY Auditory Attention, and AWMA Odd-One-Out tasks, respectively. There were no significant differences between children with hearing loss and children with normal hearing

on these measures based on two-sample t-tests (see **Table 1**). **Table 2** shows the Pearson correlations between the linguistic and cognitive standard scores and SNR50 for noise and the SNR50 for noise plus reverberation conditions for all participants. All of the predictor variables were significantly correlated with SNR50 for both speech recognition conditions. The strength of the significant correlations was medium (0.28) to large (0.80) for each bivariate relationship (Cohen, 1988).

Figure 4 shows the SNR50 for both groups of children in noise and noise plus reverberation conditions. The effects of reverberation condition (noise and noise plus reverberation), grade (1st vs. 3rd), language (PPVT), auditory attention
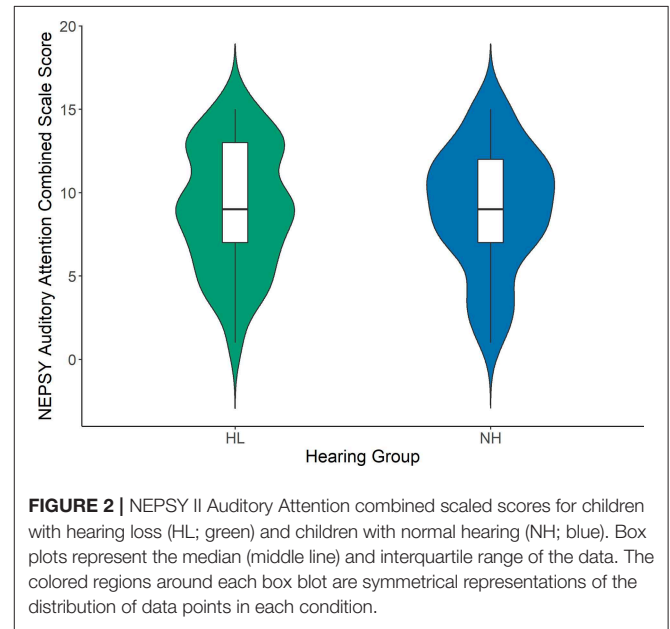
FIGURE 2 | NEPSY II Auditory Attention combined scaled scores for children with hearing loss (HL; green) and children with normal hearing (NH; blue). Box plots represent the median (middle line) and interquartile range of the data. The colored regions around each box blot are symmetrical representations of the distribution of data points in each condition.
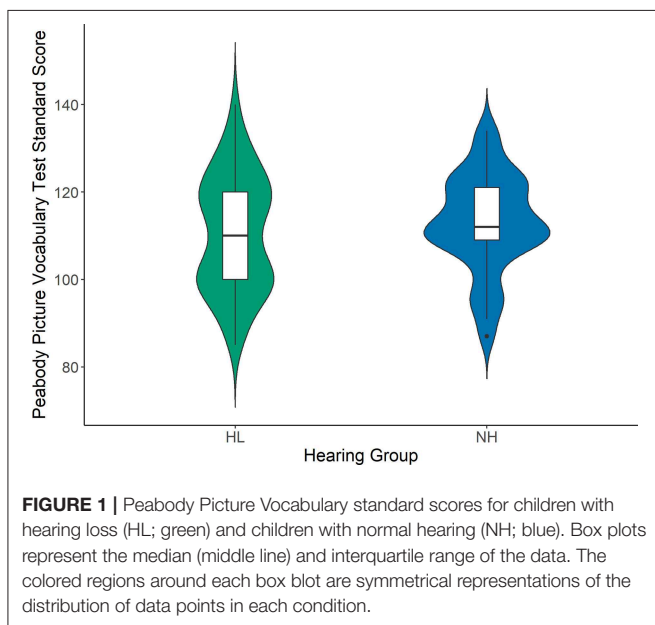
FIGURE 1 | Peabody Picture Vocabulary standard scores for children with hearing loss (HL; green) and children with normal hearing (NH; blue). Box plots represent the median (middle line) and interquartile range of the data. The colored regions around each box blot are symmetrical representations of the distribution of data points in each condition.
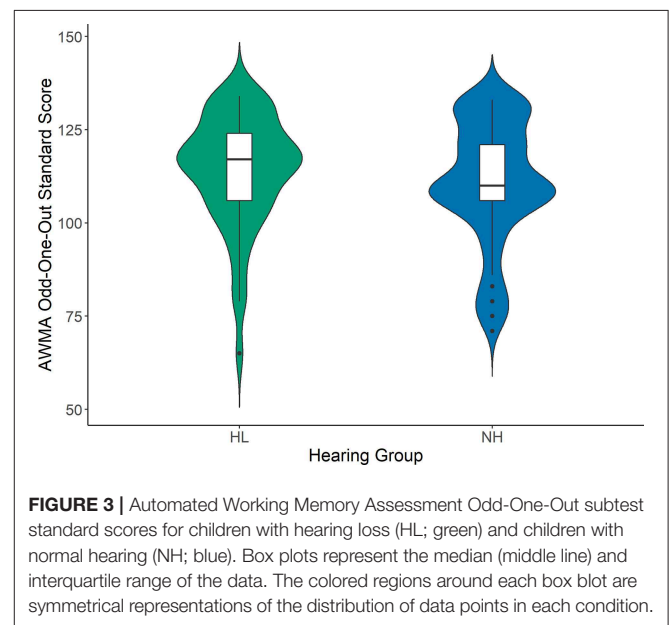
FIGURE 3 | Automated Working Memory Assessment Odd-One-Out subtest standard scores for children with hearing loss (HL; green) and children with normal hearing (NH; blue). Box plots represent the median (middle line) and interquartile range of the data. The colored regions around each box blot are symmetrical representations of the distribution of data points in each condition.

(NEPSY), visuospatial working memory (AWMA Odd-One-Out) on SNR50 for sentence recognition for children with normal hearing and children with hearing loss were examined using a linear mixed model. **Table 3** shows the statistical results of that model. Children with normal hearing had an SNR50 (Mean = 7.7 dB) that was significantly lower, by 8.1 dB, than children with hearing loss (Mean = 15.8 dB). Reverberation significantly increased the SNR50 by 5.5 dB. The lack of a significant group by reverberation condition interaction indicates that the magnitude of the group differences did not vary significantly between noise and noise + reverberation conditions. Children in 1st grade had higher (3.1 dB) SNR50 than children in 3rd grade, but this difference was not significant after controlling for other factors. Children with better visuospatial working memory and receptive vocabulary had significantly lower SNR50 than children with lower scores in these domains. There was no statistically

significant difference in SNR50 based on individual differences in auditory attention after controlling for other factors.

Table 4 includes the Pearson correlations between speech recognition, auditory variables (better-ear aided SII and better-ear pure tone average), and cognitive and linguistic standard scores for the children with hearing loss. Similar to the combined correlations for children with normal hearing and children with hearing loss (**Table 2**), the SNR50 for each condition was correlated with receptive vocabulary and working memory. Additionally, the NEPSY Auditory Attention score was correlated with the noise plus reverberation condition, but not the noise condition. Better-ear pure-tone average was not significantly associated with any predictor or outcome. The better-ear aided SII was correlated with receptive vocabulary, but not other predictors. The strength of the significant correlations was medium (0.32) to large (0.82) for each bivariate relationship (Cohen, 1988).

To examine the relationship between aided audibility, language, and speech recognition for children with hearing loss, a series of linear regression models were conducted with the children with mild-to-severe hearing loss to test whether the relationship between aided speech audibility and speech recognition in noise and reverberation was mediated by language skills. **Table 5** includes the results from the models. Individual regression models with better-ear aided SII and PPVT as predictors of aided SNR50 in noise and reverberation were completed. Each model indicated that audibility and language were significant individual predictors of SNR50 for children with hearing loss. A combined model that included both language and audibility was conducted and yielded the same pattern of results as the individual models. This pattern of results suggests that aided audibility has a direct positive effect on the SNR50 for noise and reverberation and an indirect positive effect on

TABLE 2 | Pearson correlations between speech recognition, cognition, and linguistic factors for all children.

	SNR50—N	SNR50—N+R	PPVT	NEPSY attention	AWMA OOO
SNR50—N					
SNR50—N + R	0.801*				
PPVT	−0.460*	−0.452*			
NEPSY attention	−0.258*	−0.264*	0.243*		
AWMA	−0.324*	−0.309*	0.445*	0.332*	

*SNR50, Signal-to-noise ratio for 50% correct; N, Noise; R, Reverberation; PPVT, Peabody Picture Vocabulary Test; NEPSY Attention, NEPSY Auditory Attention subtest; AWMA OOO, Automated Working Memory Assessment Odd-One-Out subtest *p < 0.05.*
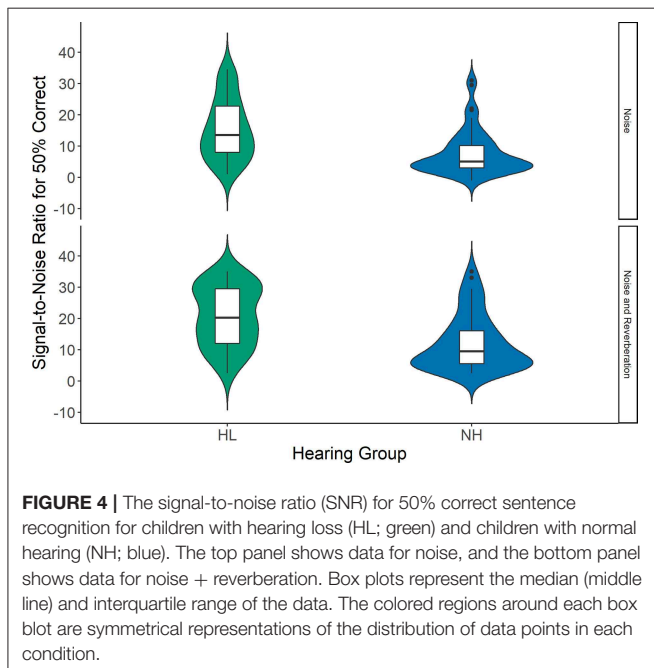
FIGURE 4 | The signal-to-noise ratio (SNR) for 50% correct sentence recognition for children with hearing loss (HL; green) and children with normal hearing (NH; blue). The top panel shows data for noise, and the bottom panel shows data for noise + reverberation. Box plots represent the median (middle line) and interquartile range of the data. The colored regions around each box blot are symmetrical representations of the distribution of data points in each condition.

TABLE 3 | Linear mixed model for predictors of SNR50 for all children.

Predictors	Estimates	Confidence interval	*p*
Subject type (NH vs. HL)	−8.1	−10.5 to −6.34	**<0.001**
Grade (1st vs. 3rd)	3.11	0.25 to 6.4	0.064
Reverberation (None vs. RT60 = 600 ms)	5.5	1.25 to 6.5	**<0.001**
NEPSY attention	−0.24	−0.65 to 0.30	0.226
AWMA OOO	−0.33	−0.59 to −0.07	**0.032**
PPVT	−0.25	−0.45 to −0.05	**<0.001**
Subject type x reverberation	−1.19	−1.98 to −0.24	0.41
Random effects			
Residual variance (σ²)	25.17		
Subject variance	31.24		

PPVT, Peabody Picture Vocabulary Test; NEPSY Attention, NEPSY Auditory Attention subtest; AWMA OOO, Automated Working Memory Assessment Odd-One-Out subtest; Estimates represent the coefficients for each variable in the model. For categorical predictors, the estimate represents the mean difference. For continuous predictors, the estimate represents the change in SNR50 for a one unit change in the predictor. All p-values for significant effects are bolded.

TABLE 4 | Pearson correlation between speech recognition and auditory variables, cognition, and linguistic factors for children with hearing loss.

	SNR50-Noise	SNR50–N+R	PPVT	NEPSY attention	AWMA OOO	Better-ear SII	Better-ear PTA
SNR50 – Noise							
SNR50 – N + R	0.737*						
PPVT	−0.457*	−0.351*					
NEPSY attention	−0.222*	−0.245*	0.121				
AWMA OOO	−0.420*	−0.321*	0.323*	0.245*			
Better-ear aided SII	−0.210	−0.305*	0.325*	0.111	0.032		
Better-ear PTA	0.073	0.032	−0.206	−0.162	−0.092	−0.846*	

*SNR50, Signal-to-noise ratio for 50% correct; N+R, Noise plus reverberation; PPVT, Peabody Picture Vocabulary Test; NEPSY Attention, NEPSY Auditory Attention subtest; AWMA OOO, Automated Working Memory Assessment Odd-One-Out subtest; Better-ear SII, Aided SII for 60 or 65 dB SPL speech signal; Better-ear PTA, Audiometric pure-tone average of thresholds at 500–4,000 Hz in the better ear *p < 0.05 (after adjustment for False Discovery Rate).*

TABLE 5 | Linear regression models for the mediation effects of language and audibility on SNR50 in reverberation for children with hearing loss.

Predictors	Estimates	Confidence interval	p
FULL MODEL			
PPVT	−0.32	−0.15 to −0.55	**<0.001**
Better-Ear aided SII	−4.85	−3.21 to −5.34	**0.04**
PATH MODELS			
Language predicting SNR 50-R			
PPVT	−0.28	−0.10 to 0.42	**<0.01**
Better-Ear aided SII predicting SNR 50-R			
Better-Ear aided SII	−5.43	−3.4 to −7.1	**0.01**

SNR50-R, SNR-50 for the reverberation condition; Better-Ear SII; PPVT, Peabody Picture Vocabulary Test; Estimates represent the coefficients for each variable in the model. For continuous predictors, the estimate represents the change in SNR50 for a one unit change in the predictor. All p-values for significant effects are bolded.

SNR50 through language ability based on the linear regression model between better-ear aided SII and language for children with hearing loss who wear hearing aids.

DISCUSSION

The goal of this study was to measure aided speech recognition in noise and reverberation for children with hearing loss and a group of children with typical hearing matched for age and socioeconomic status. Children with hearing loss completed speech recognition in noise and noise + reverberation with their hearing aids. Auditory, cognitive, and linguistic factors were analyzed to determine if they predicted individual differences in speech recognition in noise and reverberation. For children with hearing loss, the inter-relationships between speech audibility, language, and speech recognition in noise and in noise + reverberation were also examined separately. As expected, children with hearing loss performed more poorly than children with normal hearing in noise and in noise plus reverberation conditions. Individual differences in speech recognition for children with normal hearing and children with hearing loss in all adverse conditions were partially predicted by language, working memory, and auditory attention. For children with hearing loss,

the better-ear aided audibility for speech was a positive predictor of language and the aided SNR50 for noise + reverberation. Language also significantly predicted the aided SNR50 even after controlling for audibility.

Children with hearing loss were at a significant disadvantage when listening in adverse conditions compared to peers with typical hearing, even with amplification. Overall, children with normal hearing had an SNR50 that was more than 8 dB better than children with hearing loss. Reverberation (RT60) of 600 ms reduced SNR50 for both groups by an additional 5 dB. Although children with hearing loss performed more poorly than peers with normal hearing, the performance difference between groups was similar for both noise and noise plus reverberation conditions. The finding of poorer performance for children with hearing loss is consistent with the previous literature (Finitzo-Hieber and Tillman, 1978), but the previously observed pattern where reverberation disproportionately affected children with hearing loss was not replicated. We speculate that the interaction between hearing status and reverberation condition in the Finitzo-Hieber and Tillman study may have been driven by the fact that children with hearing loss in that study only listened monaurally. However, there are numerous other differences between the participants, amplification conditions, and experimental design of the studies that make it difficult to pinpoint why the interaction between hearing status and reverberation for speech recognition was not observed in this study. Future research could focus on further elucidating these factors.

As predicted by the ELU model (Rönnberg et al., 2008), individual cognitive and linguistic abilities were associated with speech recognition in noise and noise plus reverberation for children with hearing loss and children with normal hearing. Children with better vocabulary and working memory have better speech recognition in noise and noise plus reverberation conditions than peers with poorer skills in these areas. There were no interactions between these effects and condition, suggesting that the relationship between cognitive and linguistic abilities was similar for noise and noise plus reverberation conditions. These results extend previous research based on the ELU model to include children with hearing loss in conditions of noise and reverberation, which had not been examined previously. Further, the inclusion of auditory attention is consistent with

predictions that listening in adverse conditions may be related to susceptibility to informational masking that can occur with reverberation (Durlach et al., 2003) and the disruption of temporal cues in the speech signal by noise and reverberation, as age-related changes in the ability to use temporal cues have been posited to be associated with attention (Hall et al., 2012).

Our previous research has also demonstrated that children with better aided speech audibility have better speech recognition under degraded conditions because of the direct effects of audibility on speech recognition (McCreery et al., 2015, 2017), as well as indirect effects due to the cumulative influence of audibility on language development (Tomblin et al., 2015). Separate linear regression analyses of children with hearing loss in the current study indicated that audibility was positively associated with language and speech recognition in noise and reverberation, but that language also had a unique contribution to speech recognition in degraded conditions. This pattern confirms the pattern from previous research for both direct and indirect associations between audibility and speech recognition in noise for children with hearing loss. Audibility not only benefits the child through signal audibility, but also through an accumulation of auditory experience over time that fuels the language skills needed to understand speech in adverse conditions. This finding highlights the importance of consistent hearing aid use for children with hearing loss to promote access to sound for speech recognition and for long-term development of the linguistic skills that support degraded speech recognition (Tomblin et al., 2015).

Despite the fact that this study was one of the first to examine the auditory, cognitive, and linguistic factors that predict speech recognition in noise and reverberation for children who wear hearing aids, there are several limitations that could be addressed in future research on this topic. The reverberation simulation used in this study was implemented to be completed with minimal equipment requirements so that children could be tested at multiple sites. More sophisticated methods of reverberation simulation have been developed and used in recent studies with adults with hearing loss (Zahorik, 2009; Reinhart and Souza, 2016) and children with normal hearing (Wróblewski et al., 2012) or cochlear implants (Neuman et al., 2012). Future studies should take advantage of these methods to complete a more realistic assessment of speech recognition in noise and reverberation than was possible in the current study. Measures of working memory and auditory attention were chosen that would be appropriate for children with hearing loss and minimize potential confounds related to differences in audibility and language skills across subjects. A visuospatial working memory task was used, but the auditory presentation of the attention task may have been affected by auditory or linguistic abilities. The measure of auditory attention showed weak relationships with language and audibility, but future research could include visuospatial attention tasks to further minimize potential confounds. The study design also did not include realistic masking or spatial conditions that children might encounter in their everyday listening environments, which have been examined in other studies of children with normal

hearing (MacCutcheon et al., 2019) and children with hearing aids (Ching et al., 2011) or cochlear implants (Misurelli and Litovsky, 2012). Thus, we expect that children with hearing loss may have performed more favorably if the target and masker were spatially separated than in the current study where target and masker were co-located. However, previous research with children with hearing loss demonstrates large individual differences in spatial release from masking with hearing aids (Ching et al., 2011; Browning et al., 2019). Thus, the effects of noise and reverberation on aided spatial release from masking in children with hearing loss would need to be directly examined in future research.

The current study was conducted as part of a longitudinal study of children with mild to severe hearing loss, and therefore, did not include children with cochlear implants. Children with cochlear implants are likely to have significant challenges in noise and reverberation (Neuman et al., 2012). The factors that predict individual differences in speech recognition in adverse conditions in that population should be examined in future studies. Previous studies with adults with hearing loss suggest that amplification parameters may influence speech recognition in noise and reverberation (Shi and Doherty, 2008; Reinhart and Souza, 2016); however, this study was conducted with children using their hearing aids at their personal use settings, and so amplification parameters were not manipulated. Individual differences in amplitude compression settings or other hearing aid signal processing features among children in the study may have contributed to individual variability in speech recognition scores. However, this study was not designed to assess the influence of amplification parameters or hearing aid signal processing features other than audibility on speech recognition in degraded conditions. Future studies could include manipulation of children's amplification parameters to address this question.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Boys Town National Research Hospital Institutional Review Board. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. All children completed written assent or consent to participate in the study.

AUTHOR CONTRIBUTIONS

RM, EW, MS, MB, and DL planned the proposed experiment. EW, MS, and MB were involved in data collection. RM wrote the first draft of the article. RM, EW, MS, MB, and DL edited and approved the final version of the article.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Allen, P., and Wightman, F. (1995). Effects of signal and masker uncertainty on children's detection. *J. Speech Lang. Hear. Res.* 38, 503–511. doi: 10.1044/jshr.3802.503

Alloway, T. P., Gathercole, S. E., Kirkwood, H., and Elliott, J. (2008). Evaluating the validity of the automated working memory assessment. *Edu. Psychol.* 28, 725–734. doi: 10.1080/01443410802243828

ANSI (1997). *Methods for Calculation of the Speech Intelligibility Index. ANSI S3.5-1997*. New York, NY: American National Standards Institute.

Baron, R. M., and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J. Pers. Soc. Psychol.* 51:1173. doi: 10.1037//0022-3514.51.6.1173

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *arXiv:1406.5823*.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

Blamey, P., Sarant, J., Paatsch, L., Barry, J., Bow, C., Wales, R., et al. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *J. Speech Lang. Hear. Res.* 44, 264-285. doi: 10.1044/1092-4388(2001/022)

Bradley, J. S., and Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *J. Acoust. Soc. Am.* 123, 2078–2086. doi: 10.1121/1.2839285

Brooks, B. L., Sherman, E. M., and Strauss, E. (2009). NEPSY-II: A developmental neuropsychological assessment. *Child Neuropsychol.* 16, 80–101. doi: 10.1080/09297040903146966

Browning, J. M., Buss, E., Flaherty, M., Vallier, T., and Leibold, L. J. (2019). Effects of adaptive hearing aid directionality and noise reduction on masked speech recognition for children who are hard of hearing. *Am. J. Audiol.* 28, 101–113. doi: 10.1044/2018_AJA-18-0045

Brungart, D. S., Simpson, B. D., Ericson, M. A., and Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *J. Acoust. Soc. Am.* 110, 2527–2538. doi: 10.1121/1.1408946

Ching, T. Y., Van Wanrooy, E., Dillon, H., and Carter, L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *J. Acoust. Soc. Am.* 129, 368–375. doi: 10.1121/1.3523295

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd Edn.* Hillsdale, NJ: Erlbaum.

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear Hear.* 37:55. doi: 10.1097/AUD.0000000000000201

Crukley, J., Scollie, S., and Parsa, V. (2011). An exploration of non-quiet listening at school. *J. Edu. Audiol.* 17, 23–35.

Davidson, L., and Skinner, M. (2006). Audibility and speech perception of children using wide dynamic range compression hearing aids. *Am. J. Audiol.* 15, 141–53. doi: 10.1044/1059-0889(2006/018)

Dockrell, J. E., and Shield, B. M. (2006). Acoustical barriers in classrooms: the impact of noise on performance in the classroom. *Br. Educ. Res. J.* 32, 509–525. doi: 10.1080/01411920600635494

Dunn, L. M., and Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*. Bloomington, MN: Pearson Assessments. doi: 10.1037/t15144-000

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., and Kidd, G. Jr. (2003). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *J. Acoust. Soc. Am.* 114, 368–379. doi: 10.1121/1.1577562

Finitzo-Hieber, T., and Tillman, T. (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *J. Speech Hear. Res.* 21, 440–458. doi: 10.1044/jshr.2103.440

Goldsworthy, R. L., and Markle, K. L. (2019). Pediatric hearing loss and speech recognition in quiet and in different types of background noise. *J. Speech Lang. Hear. Res.* 62, 758–767. doi: 10.1044/2018_JSLHR-H-17-0389

Hall, J. W., Buss, E., Grose, J. H., and Roush, P. A. (2012). Effects of age and hearing impairment on the ability to benefit from temporal and spectral modulation. *Ear Hear.* 33, 340–348. doi: 10.1097/AUD.0b013e31823fa4c3

Johnson, C. E. (2000). Childrens' phoneme identification in reverberation and noise. *J. Speech Lang. Hear. Res.* 43, 144–157. doi: 10.1044/jslhr.4301.144

Klatte, M., Hellbrück, J., Seidel, J., and Leistner, P. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: a field study. *Environ. Behav.* 42, 659–692. doi: 10.1177/0013916509336813

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., and Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms. *Am. J. Audiol.* 11, 65–71. doi: 10.1044/1059-0889(2002/009)

Leibold, L. J., and Buss, E. (2019). Masked speech recognition in school-age children. *Front. Psychol.* 10:1981. doi: 10.3389/fpsyg.2019.01981

Leibold, L. J., Yarnell, A. B., and Buss, E. (2016). Masked speech perception thresholds in infants, children, and adults. *Ear Hear.* 37, 345–353. doi: 10.1097/AUD.0000000000000270

Levitt, H. C. C. H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Lüdecke, D., and Schwemmer, C. (2018). *sjPlot: Data Visualization for Statistics in Social Science*. Version 2.6.2.

MacCutcheon, D., Pausch, F., Füllgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., et al. (2019). The contribution of individual differences in memory span and language ability to spatial release from masking in young children. *J. Speech Lang. Hear. Res.*

Magimairaj, B. M., Nagaraj, N. K., and Benafield, N. J. (2018). Children's speech perception in noise: evidence for dissociation from language and working memory. *J. Speech Lang Hear. Res.* 61, 1294–1305. doi: 10.1044/2018_JSLHR-H-17-0312

McCreery, R., Walker, E., Spratford, M., Oleson, J., Bentler, R., Holte, L., et al. (2015). Speech recognition and parent ratings from auditory development questionnaires in children who are hard of hearing. *Ear Hear.* 36(Suppl. 1), 60S–75S. doi: 10.1097/AUD.0000000000000213

McCreery, R. W., Spratford, M., Kirby, B., and Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *Int. J. Audiol.* 56, 306–315. doi: 10.1080/14992027.2016.1266703

Misurelli, S. M., and Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry. *J. Acoust. Soc. Am.* 132, 380–391. doi: 10.1121/1.4725760

Nabelek, A. K., and Robinson, P. K. (1982). Monaural and binaural speech perception in reverberation for listeners of various ages. *J. Acoust. Soc. Am.* 71, 1242–1248. doi: 10.1121/1.387773

Neuman, A. C., and Hochberg, I. (1983). Children's perception of speech in reverberation. *J. Acoust. Soc. Am.* 73, 2145–2149. doi: 10.1121/1.389538

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2012). Measuring speech recognition in children with cochlear implants in a virtual classroom. *J. Speech Lang. Hear. Res.* 55, 532–540. doi: 10.1044/1092-4388(2011/11-0058)

R Development Core Team (2018). *R: A Language and Environment for Statistical Computing*. Version 3.4.4.

Reinhart, P. N., and Souza, P. E. (2016). Intelligibility and clarity of reverberant speech: effects of wide dynamic range compression release time and working memory. *J. Speech Lang. Hear. Res.* 59, 1543–1554. doi: 10.1044/2016_JSLHR-H-15-0371

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The Ease of Language Understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7:31. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47 (Suppl. 2), S99–S105.

Scollie, S. (2008). Children's speech recognition scores: the Speech Intelligibility Index and proficiency factors for age and hearing level. *Ear Hear.* 29, 543–556. doi: 10.1080/14992020802301167

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., et al. (2005). The desired sensation level multistage input/output algorithm. *Trends Amplif.* 9, 159–197. doi: 10.1177/108471380500900403

Shi, L.-F., and Doherty, K. A. (2008). Subjective and objective effects of fast and slow compression on the perception of reverberant speech in listeners with hearing loss. *J. Speech Lang. Hear. Res.* 51, 1328–1340. doi: 10.1044/1092-4388(2008/07-0196)

Stiles, D., Bentler, R., and McGregor, K. (2012). The Speech intelligibility index and the pure-tone average as predictors of lexical ability in children fit with hearing aids. *J. Speech Lang. Hear. Res.* 55, 764–778. doi: 10.1044/1092-4388(2011/10-0264)

Storkel, H. L., and Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behav. Res. Methods* 42, 497–506. doi: 10.3758/BRM.42.2.497

Sullivan, J. R., Osman, H., and Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *J. Speech Lang. Hear. Res.* 58, 1043–1051. doi: 10.1044/2015_JSLHR-H-14-0204

Tomblin, J. B., Harrison, M., Ambrose, S. E., Oleson, J., and Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss. *Ear Hear.* 36(Suppl. 1), 76S−91S. doi: 10.1097/AUD.0000000000000219

Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., and Moeller, M. P. (2014). The influence of hearing aids on speech and language development of children with hearing loss. *JAMA Otolaryngol. Head Neck Surg.* 140, 403–409. doi: 10.1001/jamaoto.2014.267

Walker, E. A., Redfern, A., and Oleson, J. J. (2019). Linear mixed-model analysis to examine longitudinal trajectories in vocabulary depth and breadth in children who are hard of hearing. *J. Speech Lang. Hear. Res.* 62, 525–542. doi: 10.1044/2018_JSLHR-L-ASTM-18-0250

Wickham, H., and Chang, W. (2016). *ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics.* Version 2.2.1.

Wróblewski, M., Lewis, D. E., Valente, D. L., and Stelmachowicz, P. G. (2012). Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. *Ear Hear.* 33:731. doi: 10.1097/AUD.0b013e31825aecad

Zahorik, P. (2009). Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J. Acoust. Soc. Am.* 126, 776–791. doi: 10.1121/1.3167842

Comorbidity of Auditory Processing, Attention, and Memory in Children With Word Reading Difficulties

Rakshita Gokula[1,2]*, Mridula Sharma[1,2,3], Linda Cupples[1,3] and Joaquin T. Valderrama[1,2,4]

[1] Department of Linguistics, Macquarie University, Sydney, NSW, Australia, [2] HEARing Cooperative Research Centre, Melbourne, VIC, Australia, [3] Centre for Language Sciences, Macquarie University, Sydney, NSW, Australia, [4] National Acoustic Laboratories, Sydney, NSW, Australia

Objectives: To document the auditory processing, visual attention, digit memory, phonological processing, and receptive language abilities of individual children with identified word reading difficulties.

Design: Twenty-five children with word reading difficulties and 28 control children with good word reading skills participated. All children were aged between 8 and 11 years, with normal hearing sensitivity and typical non-verbal intelligence. Both groups of children completed a test battery designed to assess their auditory processing, visual attention, digit memory, phonological processing, and receptive language.

Results: When compared to children who were good readers, children with word reading difficulties obtained significantly lower average scores on tests of auditory processing, including the frequency pattern test, gaps in noise, frequency discrimination, Dichotic Digit difference Test, and Listening in Spatialized Noise. The two groups did not differ on the discrimination measures of sinusoidal amplitude modulation or iterated rippled noise. The results from children with word reading difficulties showed that 5 children (20%) had comorbid deficits in auditory processing, visual attention, and backward digit memory; whereas 12 children (48%) had comorbid auditory processing and visual attention deficits only, and 2 children (8%) had comorbid deficits in auditory processing and digit memory; the remaining children had only auditory processing, visual attention, or digit memory deficits.

Conclusion: The current study highlights the general co-existence of auditory processing, memory, and visual attention deficits in children with word reading difficulties. It is also noteworthy, however, that only one fifth of the current cohort had deficits across all measured tasks. Hence, our results also show the significant individual variability inherent in children with word reading difficulties.

Keywords: word reading difficulty, auditory processing, cognition, digit memory, receptive language

INTRODUCTION

Co-morbidities in children with developmental disorders are more the norm than the exception. Several studies have reported that children with auditory processing difficulties have coexisting deficits in language skills (Benasich et al., 2002; McArthur and Bishop, 2004; Wible et al., 2005; Sharma et al., 2009), attention skills (Dhamani et al., 2013; Allen and Allan, 2014; Gyldenkærne et al., 2014; Sharma et al., 2014; Tomlin et al., 2015), and/or memory skills (Allen and Allan, 2014; Sharma et al., 2014). Other studies suggest that children with reading difficulties exhibit coexisting deficits in auditory processing (Goswami et al., 2002; Banai and Ahissar, 2004; Fischer and Hartnegg, 2004; Halliday and Bishop, 2006a; Sharma et al., 2006, 2009; Iliadou et al., 2009; Reid et al., 2010; Hämäläinen et al., 2012), language skills (Scarborough, 1990; Wise et al., 2007; Scarborough et al., 2009) attention skills (Willcutt and Pennington, 2000; Willcutt et al., 2005) and/or working memory abilities (Swanson et al., 1989, 2009; Swanson, 1993). Considering the weight of evidence to date, which shows that children are more likely to have deficits across multiple skills than deficits in isolated skills, this research was designed to investigate the range and frequency of different co-morbidities evident in children with word reading difficulties. Word reading difficulties were defined as scores that fell 1.5 standard deviation (SD) or more below the typical mean in oral reading of non-words (Badian, 1996; Compton et al., 2001; Banai and Ahissar, 2006; Paul et al., 2006; Baird et al., 2011).

The variables of auditory processing, attention, working memory, phonological processing, and language were measured in the search for co-morbidities, due to their frequently observed association with reading ability. The clinical motivation for this study lay in the belief that targeting multiple functional areas in the assessment process would significantly help a clinician to collect sufficient information to guide a multi-disciplinary approach in order to manage the full range of deficits that a given child may exhibit.

Auditory Processing and Reading

Auditory processing has been studied extensively in relation to children's reading ability, with the earliest theory proposing an association between auditory processing and phonic decoding skills in particular; that is, the ability to read by mapping letters onto sounds (Tallal, 1980, 1984). Tallal (1980) demonstrated this association using an auditory temporal-order judgment task in which children with reading difficulty made significantly more errors than children with typical reading skill at fast but not slow presentation rates.

According to Ramus (2001), however, sensory theories of dyslexia, such as that proposed by Tallal and colleagues, suffered from a number of potential weaknesses, including: the failure of some studies to find an association between reading and auditory processing; the finding that only a subgroup of participants is often responsible for reported group differences; and the possibility that apparent sensory deficits might instead reflect differences in the strategies used for task completion. Furthermore, Ramus questioned whether there was sufficient evidence to support claims of a causal association between perceptual processing and word reading indirectly through a potential phonological relationship.

Despite such criticisms, there is continued interest in the possible role of basic auditory processing deficits as they relate to reading difficulty (e.g., Goswami et al., 2002; Leppänen et al., 2010; Goswami, 2011; Huss et al., 2011; Casini et al., 2018). Rise-time theory, described originally by Goswami et al. (2002), proposed that children with dyslexia experience a basic auditory processing deficit that interferes with their perception of the rhythmic timing of speech. In support of this proposal, they used a beat detection task to show that children with reading difficulties performed significantly more poorly than control children who were matched on chronological age. More recently, Casini et al. (2018) assessed the temporal and intensity discrimination skills of children with poor word reading. They reported that children with word reading deficits performed more poorly on a temporal discrimination task than a group of peers with typical reading skills who were matched on chronological age. The groups did not differ significantly, however, in their judgments of intensity. In one of the few longitudinal studies in this area, Leppänen et al. (2010) reported that auditory ERPs in neonates were correlated with measures of phonological awareness at 3.5 years of age and letter knowledge at 5 years of age, and were significant predictors of 9-year-old measures of reading speed and reading accuracy after controlling for a range of other potentially important variables.

These various findings support a role for basic auditory processes in the development of typical reading skills, but other researchers have questioned such a role. For example, Snowling et al. (2018) found no evidence in their longitudinal data for an association between early frequency discrimination (measured at 4.5 and 5.5 years of age) and later reading outcomes (measured at 5.5 and 8 years old). Furthermore, because executive function at 4.5 predicted frequency discrimination at 5.5, they suggested that poor performance on auditory processing tasks might be due to comorbid attentional difficulties in some children. This suggestion accords with the hypothesis offered by Ramus (2001, p. 395) that auditory processing deficits might be found only in people with reading difficulty who also have some other developmental disorder, which he refers to as a "hidden factor." It is also consistent with Leppänen et al.'s (2010) suggestion that an early auditory processing deficit may not be sufficient, on its own, to cause a reading difficulty. In sum, despite decades of work in the field of auditory processing and reading, the evidence of specific auditory processing skills and their contribution to reading is not well understood. A question of direct relevance to research endeavors in this field is how auditory processing should be measured.

Relevant Auditory Processing Skills

Auditory processing is an umbrella term that encapsulates abilities such as auditory discrimination (e.g., frequency discrimination), spectral resolution and discrimination (e.g., amplitude and frequency modulation), temporal ordering (e.g., frequency patterning), and performance in degraded listening conditions (e.g., listening in noise) (American Speech-Language-Hearing Association [ASHA], 1996). While there are more

recent definitions of auditory processing offered in the literature, none define the specific skills and co-morbidities of auditory processing as explicitly as the (American Speech-Language-Hearing Association [ASHA], 1996, 2005) documentation (Jerger and Musiek, 2000).

Ramus (2001) and Goswami (2015) in their respective reviews of the literature noted that most theories of reading difficulty that attribute an important role to auditory processing deficits assume an intervening association with phonology; that is, auditory processing deficits result in phonological impairment, which in turn leads to a reading difficulty. As noted above, however, auditory processing can be measured in a multitude of ways. Missing from the literature is a detailed understanding of how and why different measures of auditory processing may be associated with particular components of phonological processing and/or reading subskills.

Another challenge for the field is the presence of ambivalent results across a number of auditory processing tasks. For instance, Hämäläinen et al. (2012) in their review reported that children with reading difficulties performed significantly worse than children with no reading difficulties in frequency discrimination (FD; Banai and Ahissar, 2006; Goswami et al., 2010). Conversely, a study by Adlard and Hazan (1998) reported that FD was unaffected in children with reading difficulties. In the review (Hämäläinen et al., 2012), there were studies that reported significantly worse thresholds on frequency modulation (FM) at slow rates of 2 Hz in children with reading difficulties compared to children with no reading difficulties (Gibson et al., 2006; Dawes et al., 2009; Wright and Conlon, 2009). In contrast, other researchers found no differences in FM thresholds of children with reading difficulties and their age-matched peers at modulation rates of 2 Hz (Halliday and Bishop, 2006b; Dawes et al., 2009), 20 Hz (Halliday and Bishop, 2006b), or 240 Hz (Adlard and Hazan, 1998). Similar dichotomous reports have been noted for amplitude modulated (AM) thresholds as well. For instance, Rocheron et al. (2002) reported significant group differences for very low and high modulation rates of 4 and 128 Hz; whereas Hämäläinen et al. (2009) reported no significant differences on the same task, at a modulation rate of 20 Hz.

Findings from studies that assessed children's performance on the more commonly used clinical tests such as Frequency Pattern Test (FPT), Dichotic Digit Test (DDT), Gaps in Noise (GIN), and speech in noise (Sharma et al., 2006; Iliadou et al., 2009) are more consistent in showing that children with reading difficulties have poorer responses than their age-matched peers with typical reading skills. Barker et al. (2017) used an iPad-based app to assess FPT and DDT and found that children with poor reading comprehension were significantly worse on both measures compared to children with good reading comprehension skills.

The different patterns of results obtained across various studies that involve similar tasks and children of a similar age are of interest, because they raise questions about the reliability of tests used (e.g., test–retest), heterogeneous characteristics of the population, and variability in performance (e.g., intrinsic attention during assessment). The theoretical basis for the contribution of auditory processing to reading will remain a challenge while these three aspects remain unanswered. Therefore, a secondary aim of the current study was to evaluate the individual profiles of a sample of children with word reading difficulties on well-established auditory processing tasks.

Phonological Processing, Vocabulary, Visual Attention, Digit Memory and Reading

The relationship between phonological processing and word reading is well established, and therefore the current study included assessment of phonological processes to confirm the presence of individual variability, if any, in the current cohort (Wagner and Torgesen, 1987). For the same reason, assessments of receptive vocabulary, visual attention, and digit memory were included in the test battery. Notably, however, the aim of the current study was not to determine *whether* these skills were worse in our cohort of poor word readers than in a peer group of typical readers, but rather to document significant group differences and to discover the extent to which the current sample of children with word reading deficits exhibited individual variability in these skills.

Reading involves not only the conversion of print to speech, but also the assignment of meaning to words and larger units of language, with comprehension being the ultimate intention (Ouellette, 2006). Regardless of whether the relationship between reading and receptive vocabulary is direct (Scarborough et al., 2009), or one that is mediated by phonological awareness (Whitehurst and Lonigan, 1998), children's vocabulary is an essential component of oral language that is crucial for skilled reading (Perfetti et al., 1996; Muter et al., 2004; Ouellette, 2006). Vocabulary growth appears to play a role in the development of phoneme awareness (Metsala, 1999; Goswami, 2001; Walley et al., 2003), which in turn is associated with word decoding. Felton et al. (1987) found that children with reading difficulties were significantly poorer on measures of receptive vocabulary than age-matched controls with typical reading skills. Ouellette (2006), in a study of 60 children, found vocabulary to be the sole measure to concurrently predict decoding ability (measured using oral reading of non-words) when variables of age and non-verbal intelligence were controlled.

The cognitive measures of particular relevance to this study are those of visual attention and digit memory. Visual attention, especially in the spatial domain, is employed for reading (Stevens and Bavelier, 2012). Casco et al. (1998) studied the association between visual selective attention and reading rate in children. Visual selective attention was assessed using a task that required children to identify target alphabets that were interleaved amidst visually similar symbols. Children who made more errors on the visual attention task were significantly slower readers. Rabiner and Coie (2000) reported that attention difficulties predicted the concurrent reading achievement of typically developing children, after measures such as IQ, and prior reading achievement were controlled within the group.

Working memory is another cognitive domain that has been studied extensively to assess its association with children's reading skill. Working memory, as described by Baddeley (1986), includes

a central executive component which monitors the phonological loop and the visuo-spatial sketchpad which are responsible for sound based input and visual input respectively. Working memory tasks assess an individual's ability to maintain task-related information while processing that information further or performing another cognitive task. Swanson et al. (2009) in a meta-analysis of 88 studies noted poorer working memory, as measured using reading/listening span, digit span, and digits backwards, in children with reading disabilities in support of Baddeley's theoretical framework.

The Current Study: Reading, Auditory Processing, Phonological Processing, Visual Attention, Digit Memory and Language Skills

The current research had two aims. The first aim was to determine whether children with word reading difficulties have poorer auditory processing skills than age-matched control children with typical reading skills. The second aim was to identify individual profiles and commonly occurring co-morbidities within the group of poor non-word readers. In order to determine the individual profiles, all 25 children were tested on auditory processing, phonological processing, visual attention, digit memory, and language tasks. Before determining the profiles, the question of performance criteria was considered.

Performance Criteria

A crucial methodological question in any research that involves the identification of impaired performance is: what constitutes a deficit. A common strategy is to identify a deficit in terms of how far below the typical mean an individual score falls in standard deviation (SD) units. Wilson and Arnott (2013) discussed the impact of choosing a criterion for identification of auditory processing disorder, where the use of 1 SD below or 2 SD below the mean can lead to disparity in the numbers of children diagnosed. Notably, Wilson and Arnott chose to use minus 2 SD as the diagnostic criterion in line with ASHA guidelines. In accordance with this approach, a deficit in auditory processing was identified in the current study when individual scores fell 2 SD or more below the mean. A similar criterion was not deemed appropriate for all measures, however. Much of the cognitive literature is consistent in using 1 SD below the mean as the demarcation for deficits in attention (Dawes et al., 2009; Landerl and Willburger, 2010; Franceschini et al., 2012) and memory (Swanson et al., 1989, 2009; Swanson, 1993). The reading literature is different again, typically using 1.5 SD below the mean to indicate a deficit (Badian, 1996; Compton et al., 2001; Banai and Ahissar, 2006; Paul et al., 2006; Baird et al., 2011). The current paper is not designed to determine the most appropriate criteria for identification of atypically poor performance across the range of skills measured, and is therefore aligned with the published literature in defining deficits as follows:

- ≥2 SD below the mean for auditory processing tasks (American Speech-Language-Hearing Association [ASHA], 1996);
- ≥1.5 SD below the mean for reading;
- ≥1 SD below the mean for attention and memory.

We hypothesized that a majority of the children with word reading difficulties in the current study would have concurrent comorbidities on all of the measured skills.

MATERIALS AND METHODS

Participants

Fifty-three children aged 8–11 years (Mean age in years ± SD = 9.7 ± 1.17) participated in the current study. Of the 53 children, 25 (9.5 ± 1.16 years) were identified as having reading concerns (henceforth referred to as word reading difficulty) because they scored at least 1.5 SD below the mean on the Castles and Coltheart 2 (CC2) test of non-word reading. Twenty-eight were typically developing control children (10.0 ± 1.09 years) who scored within 1 SD of the mean or better on the same test of non-word reading. The participants in the study did not report with any other developmental concerns. All participants spoke English as their first language and attended schools that used English as the medium of instruction. Participants were recruited through advertisements in the Macquarie University Speech and Hearing Clinic (Sydney, NSW, Australia), on social media sites, and in children's magazines available freely to families across Sydney. Parents provided written informed consent for their children to participate in the study, and each child gave verbal assent as per the requirement of the Human Research Ethics Committee at Macquarie University (Reference No: 5201600441). Families received a token of appreciation for participating in the study. The study conformed, at all times, to the guidelines of the Australian Government: National Health and Medical Research Council (2018).

Inclusion Criteria

All participants were tested for normal hearing sensitivity using clinical tests: otoscopy, tympanometry, pure tone audiometry, acoustic reflex thresholds, and distortion product otoacoustic emissions (DPOAEs). Acoustic reflex thresholds and DPOAEs were used to identify any underlying hearing loss that may not be picked up by audiometry. Non-verbal cognitive ability was also assessed to ensure that all children had an age-appropriate Non-Verbal Intelligence Quotient (NVIQ) of 85 or greater.

Otoscopy was conducted to determine the general health of the ear canal and identify any visible signs of infection. Tympanometry was carried out with a 226 Hz probe tone to test middle ear status. Pure tone audiometry (PTA) was conducted using the modified Hughson-Westlake procedure (5 dB steps). During PTA, participants were instructed to indicate whenever they heard a sound and were asked to pay close attention to soft sounds. PTA and tympanometry were carried out in a sound-treated booth. Children were included in the study if their hearing thresholds were ≤15 dB HL at octave frequencies between 250 Hz and 8000 Hz (American National Standards Institute [ANSI], 1996)

and showed normal ear compliance and ear canal volume (Medwetsky et al., 2009).

Acoustic reflex thresholds were obtained through ipsilateral and contralateral stimulation at octave frequencies from 500 to 4000 Hz. Children were included in the study if ipsilateral and contralateral acoustic reflexes were detected at 500, 1000, and 2000 Hz consistent with their audiometric thresholds. DPOAE testing was conducted for both ears between 1000 and 6000 Hz. Children were included if they had present DPOAEs on three consecutive frequencies with at least a signal to noise ratio (SNR) of +6 dB (Medwetsky et al., 2009) consistent with PTA and immittance results.

The Wechsler Non-Verbal Scale of Ability (WNV) was used to assess non-verbal cognitive ability (Wechsler and Naglieri, 2006). This test includes a non-verbal mode of instruction using pictures from the test manual. The matrices and spatial span subtests of the WNV assessed the children's non-verbal reasoning and spatial memory skills.

In the matrices subtest, children finished an incomplete matrix of images by pointing to the correct image from within a list of options. The test items gradually increased in difficulty. Testing stopped once a child responded incorrectly to four out of five consecutive matrices. In the spatial span task, children were presented with blocks on a board. Each block had a number visible only to the examiner. The examiner pointed to a specific number of blocks in a given sequence that the child had to imitate either in the same order or in reverse order. The test began with three blocks. For each length of block number, two sequences were presented. If the child was able to imitate one or both sequences for a given number of blocks, the number of blocks presented increased by one. Testing stopped once a child responded incorrectly to both sequences at a given block length. A standard score, equivalent to NVIQ, was assigned on the basis of the raw scores for matrices and spatial span subtests (Wechsler and Naglieri, 2006). A standard score of ≥85 (Massa and Rivera, 2009) was required for a child to be included in the study.

Table 1 presents the means and SDs of children's age, PTA, and WNV scores according to group. This table shows no significant difference for the audiometry thresholds obtained by children in
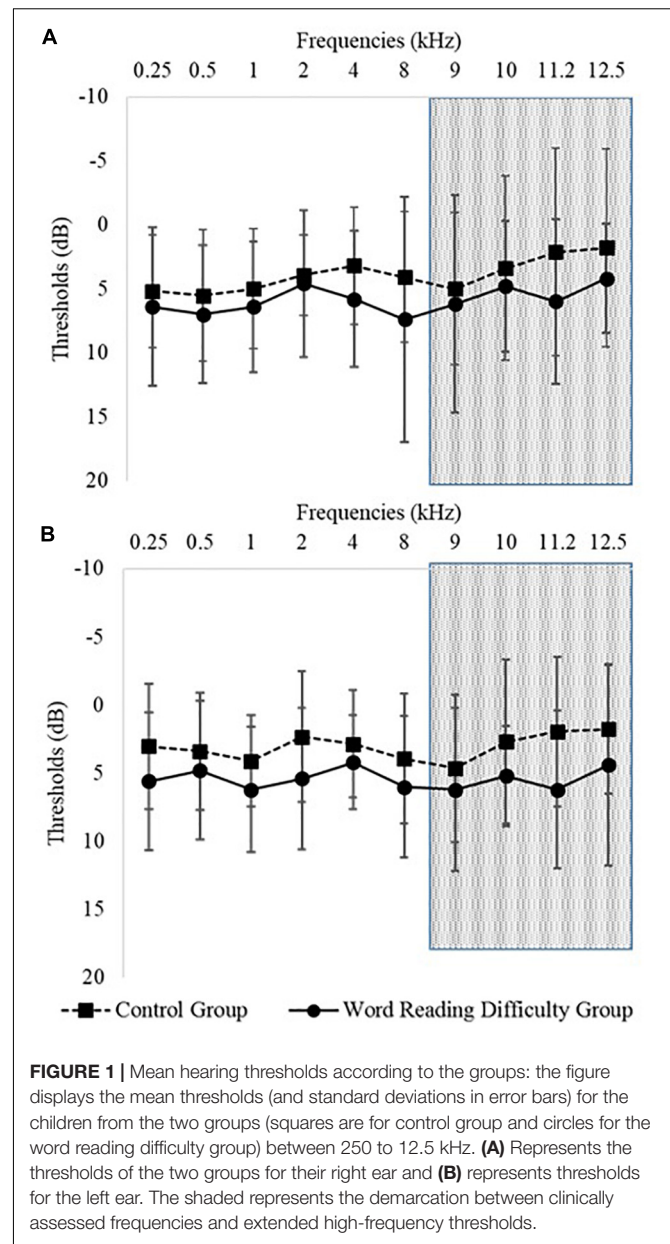
FIGURE 1 | Mean hearing thresholds according to the groups: the figure displays the mean thresholds (and standard deviations in error bars) for the children from the two groups (squares are for control group and circles for the word reading difficulty group) between 250 to 12.5 kHz. **(A)** Represents the thresholds of the two groups for their right ear and **(B)** represents thresholds for the left ear. The shaded represents the demarcation between clinically assessed frequencies and extended high-frequency thresholds.

TABLE 1 | Means (and SDs in parentheses) for age, PTA (0.5 – 4 kHz), and WNV scores for children in the two groups.

Groups	Age* (SD)	PTA** Right (SD)	PTA** Left (SD)	WNV (SD)
Control, N = 28	10.0 (11)	4.4 (3.4)	3.2 (3.0)	118.7 (10.4)
Females, N = 11	9.3 (0.9)	4.2 (3.9)	3.4 (3.3)	119.2 (10.7)
Males, N = 17	10.5 (1.0)	4.4 (3.2)	3.0 (3.0)	118.4 (10.5)
Word Reading Difficulty, N = 25	9.4 (1.2)	5.6 (4.7)	4.8 (4.0)	112.4 (9.29)
Females, N = 9	9.8 (1.2)	5.3 (4.3)	4.6 (4.8)	113.1 (4.96)
Males, N = 16	9.4 (1.2)	6.3 (4.6)	5.4 (3.3)	112.1 (11.2)

*Age is presented in years. **PTA is presented in dB HL.*

the two groups from 500 to 4k Hz for both ears [$F(3,153) = 1.64$, $p = 0.18$]. The group mean audiometry result for the left and right ear is presented in **Figure 1** that shows no significant difference for the extended high frequencies 8k to 12.5k Hz [$F(4,204) = 0.35$, $p = 0.82$]. The control group showed a small but significant advantage on the WNV when compared to the group with word reading difficulties [$F(1,51) = 5.28$, $p = 0.03$], but both groups scored around 1 SD above the mean on average (see **Table 1**).

Tests

Testing for each participant occurred over two separate days within a 7- to 10-day period. Each testing session lasted 2.5–3 h with regular breaks. Testing was conducted in a distraction-free Macquarie University Speech and Hearing Clinic room (Sydney,

NSW, Australia). To minimize any effect of procedural bias, testing order was randomized.

The Maximum Likelihood Procedure (MLP) toolbox (Grassi and Soranzo, 2009) was used to develop the FD (Peter et al., 2014), sinusoidal amplitude modulation (SAM; Peter et al., 2014), and iterated ripple noise (IRN; Peter et al., 2014) tests for the study. The stimuli for the behavioral hearing tests (FD, SAM, and IRN threshold) were created at a sampling rate of 44100 Hz. Staircase method was used for threshold estimation. The clinically available test stimuli for FPT (Musiek, 2002) and GIN (Shinn et al., 2009) were routed through a clinical audiometer (AC 40) and played through HDA 200 headphones (Sennheiser Electronic Corporation, Old Lyme, CT, United States) at a level of 50 dB HL (American National Standards Institute [ANSI], 1996). Listening in Spatialized Noise-Sentences (LiSN-S; Cameron and Dillon, 2007) and Dichotic Digit difference Test (DDdT; Cameron et al., 2016) were played through the computer via commercially available software, through headphones accompanying the LiSN-S test (HD 215, Sennheiser Electronic Corporation, Old Lyme, CT, United States).

Calibration of the auditory stimuli (FPT, SAM, IRN, FD, and GIN) was carried out using a Type 2231 sound level meter (SLM), a Type 4152 artificial ear, a Type 4144 1-inch pressure microphone, and an AC 40 Audiometer (Brüel & Kjaer Sound & Vibration Measurements A/S, Naerum, Denmark). The stimuli were calibrated to ensure that the output from the Audiometer was 50 dB HL.

Auditory processing assessment for all children included FPT – right and left ear (Musiek, 2002), GIN – right and left ear (Shinn et al., 2009), FD (Peter et al., 2014), SAM – 4 and 40 Hz (Peter et al., 2014), IRN – 32 and 4 iterations (Peter et al., 2014), LiSN-S (Cameron and Dillon, 2007) and Dichotic Digit difference Test – dichotic and diotic listening (DDdT; Cameron et al., 2016). Phonological processing was assessed using the elision subtest from the Comprehensive Test of Phonological Processing (Wagner et al., 1999). Visual attention was assessed using the sky search, map mission, creature counting, and same and opposite world subtests from the Test of Everyday Attention for Children (Manly et al., 2001). Digit memory was assessed using the digit forward and backward subtask from Clinical Evaluation of Language Fundamentals – Fourth edition (CELF-4; Semel et al., 2006). Receptive vocabulary was assessed using the Peabody Picture Vocabulary Test – Fourth edition (PPVT-4; Dunn and Dunn, 2007). The current study used tests that have extensively been employed in previous research. The details of the tests are provided in **Supplementary Tables 1, 2** to enable reduplication of the current research. The current study also used three auditory processing measures: FD, IRN, and SAM, the stimulus and methodology for which, tend to vary across publications. Details of these tests are presented in **Table 2**.

Statistical Analysis

One-way Analyses of Variance (ANOVAs) were used to compare the groups' performance on each task. A conservative p-value of 0.01 was used to reduce the likelihood of type I errors associated with multiple comparisons. Standard scores were not available for all auditory processing measures; in these cases, Analyses of Covariance (ANCOVAs) were conducted on raw scores, with age as a covariate to account for any age-related differences in performance.

Criteria for Individual Sub-Profiles

Word and non-word reading abilities were assessed using the CC2 test. This test includes three word lists containing regular words (whose correct pronunciation is in line with letter-sound rules; e.g., take), irregular words (whose correct pronunciation conflicts with letter-sound rules; e.g., eye), and non-words (e.g., norf). The children in the current study were assessed on their ability to read all three types of words, and those whose non-word reading scores fell 1.5 SD or more below the normative mean were classified as having word reading difficulties. In accordance with the performance criteria outlined earlier (see section Performance Criteria), children in the study were identified as having an auditory processing deficit if they scored 2 SD or more below the mean on one or more of the auditory processing tests (American Speech-Language-Hearing Association [ASHA], 1996). They were identified as having poor attention if they scored at least 1 SD below the mean on either of the visual attention tasks – selective or switching, and they were identified as having a working memory problem if their performance on the backwards digit task was at least 1 SD below the typical mean. Correlations were also conducted across the tests used in the current study to observe the linear relationships between the variables across which the groups were compared.

RESULTS

Tests of Reading and Phonological Processing

Children's raw scores for regular, irregular and non-word reading on the CC2 test were compared to the age-based norms to derive z-scores. The control group achieved mean z-scores of 1.7 (SD = 0.85), 0.97 (SD = 0.78), and 1.1 (SD = 0.91) on regular word, irregular word, and non-word reading, respectively. By contrast, the group with word reading difficulties had mean z-scores of -1.9 (SD = 0.40), -1.4 (SD = 0.66), and -1.9 (SD = 0.40) respectively. Children in the control group achieved standard scores of 13.8 (SD = 1.20) on the phonological awareness test of elision while the children with word reading difficulties had an average standard score of 10.3 (SD = 2.65) with a statistically significant difference between the groups ($F[1,51] = 45.0$, $p < 0.001$).

Auditory Processing Tests

Scores on most of the individual auditory processing tasks were significantly worse, on average, for children with word reading difficulty than for the control group (see results from univariate analyses in **Table 3**). The group differences were significant for FD, FPT, GIN, DDdT scores (dichotic and diotic), and LiSN-S measures (low cue; high cue); but there were no significant group effects for SAM thresholds or IRN thresholds.

TABLE 2 | Details of the auditory processing tests employed in the current study.

Measures	Tests	Procedure
Auditory processing	Frequency Discrimination (FD)	*Stimuli:* A 1000 Hz tone was used as the stimulus. The target tone was varied between 1100 and 1001 Hz by multiplicative step size factors of 2 for the first two reversals and 1.41 thereafter (Moore et al., 2011). The stimuli were 250 ms in duration. There was a 50 ms gap between the two tones. *Procedure:* FD is a measure affected by learning, and changes with repetition of the task (McArthur and Hogben, 2012). Therefore, the current study incorporated testing across two runs. Thresholds were estimated using a 3 Alternative Forced Choice (AFC) task which uses the 2-down 1-up tracking method which helps calculate the 70.7% correct point on a psychometric function (Levitt, 1971). The target tone was provided in the first, second, or third interval, alongside two comparison tones. Feedback was provided after each trial to inform the participant whether their response was correct or incorrect. *Response and Scoring:* The participants were asked to identify the sound that was different amongst the three stimuli. Every correct response resulted in a reduction in the difference between the tones. Every wrong response led to an increase in the difference between the tones. The arithmetic mean of the last six reversals in a block of 12 was taken as the threshold. A logarithmic transformation was applied to the threshold. The norms for the test were taken from Halliday and Bishop (2006a). Two such responses were obtained from the participants across separate blocks. As per Moore et al. (2008), assessment of track widths across two response blocks of the participants ensured "genuine good" and "genuine poor" responses compared to non-compliant responses from children who lost their attention during the task. Visual assessment of the track widths for each participant ensured that the study did not include any participants with non-compliant responses due to poor attention. A lower score (in Hz) represented better frequency discrimination ability in a participant.
	Threshold for detection of Sinusoidal Amplitude Modulation (SAM)	*Stimuli:* For the SAM task white noise modulated at 4 and 40 Hz was used. Over the first two reversals, the modulation depth was reduced by 2 dB after which the reversals were in 1 dB step size till the threshold was estimated. *Procedure:* Similar to the procedure employed to obtain the FD score. *Response and Scoring:* Similar to the response and scoring in the FD test. No logarithmic transformation was applied to the thresholds. The norms for the test were taken from Peter et al. (2014). Two such responses were obtained from the participants across separate blocks. A lower score (in dB) represented better thresholds for detection of AM in a sinusoidal signal.
	Threshold for discrimination of Iterated Rippled Noise (IRN)	*Stimuli:* IRN with 4 and 32 iterations were created with a delay of 10 ms. Four iterations correspond to a weak pitch percept, and 32 iterations have a strong pitch percept (Krishnan et al., 2014). A 500 ms Gaussian noise was low-pass filtered at 3000 Hz and added back to itself to create the IRN stimulus. The starting level for the gain (in dB) was set at 9.8 dB, which corresponds to a gain of 0.32 (Peter et al., 2014). From the starting level, the gain was reduced in step sizes of −3 dB (for the first two responses) and −1.5 dB (over the course of the rest of the test). *Procedure:* Similar to the procedure for the FD task. *Response and Scoring:* Similar to the response and scoring in the FD task. No logarithmic transformation was applied to the thresholds. The norms for the test were taken from Peter et al. (2014). Two such responses were obtained from the participants across separate blocks. A lower score (in dB) represented better thresholds for detection of pitch embedded in the IRN signal.

TABLE 3 | ANOVA results alongside the means (and standard deviations in parentheses) across the two groups for the individual auditory processing tests.

Test		Control Mean (*SD*)	Word Reading Difficulty Mean (*SD*)	*F*-value	*p*-value	Effect size
FD	Run 1 (log)	1.09 (0.31)	1.72 (0.53)	24.6	<0.001*	0.330
	Run 2 (log)	1.00 (0.30)	1.59 (0.59)	18.24	<0.001*	0.267
IRN	32 iterations	19.20 (2.81)	18.67 (2.73)	0.34	0.561	0.007
	04 iterations	13.28 (3.10)	11.62 (2.43)	5.13	0.028	0.093
SAM	40 Hz	−15.78 (1.59)	−14.71 (3.50)	1.13	0.291	0.022
	4 Hz	−11.96 (2.57)	−9.55 (3.79)	5.31	0.025	0.096
FPT	Right (%)	93.78 (6.95)	71.82 (21.26)	23.06	<0.001*	0.316
	Left (%)	92.23 (11.12)	70.10 (22.51)	18.10	<0.001*	0.266
GIN	Right (ms)	4.96 (0.88)	6.36 (1.93)	9.14	0.004	0.155
	Left (ms)	5.14 (0.80)	6.36 (1.93)	7.04	0.010	0.124
DDdT	Dichotic (*z*-score)	0.62 (1.14)	−1.02 (1.17)	26.86	<0.001*	0.345
	Diotic (*z*-score)	0.59 (0.88)	−1.29 (1.25)	40.50	<0.001*	0.443
LiSN– S	Low cue Score (*z*-score)	−0.43 (0.81)	−1.25 (0.98)	11.13	0.002	0.179
	High cue Score (*z*-score)	0.42 (0.74)	−0.29 (0.89)	10.36	0.002	0.169
	Talker advantage (*z*-score)	−0.18 (0.88)	−0.30 (0.96)	0.23	0.634	0.004
	Spatial advantage (*z*-score)	0.02 (1.27)	−0.44 (0.99)	2.19	0.148	0.041
	Total advantage (*z*-score)	0.73 (0.79)	0.29 (0.83)	3.91	0.053	0.071

*In these analyses, the df, error df for the F-values is (1,50) for tests with raw scores and (1,51) for tests with z-scores. *p-value set to 0.01 to account for multiple comparisons.*

TABLE 4 | Visual attention and digit memory skills across the groups.

Cognitive tests	Control Mean (*SD*)	Word Reading Difficulty Mean (*SD*)	*F*-value	*p*-value	Effect size
Visual Attention					
Selective Attention	9.6 (2.5)	8.1 (1.5)	6.84	0.012*	0.118
Switching Attention	12.1 (2.7)	6.6 (2.2)	63.2	<0.001*	0.554
Digit memory					
Digit forward	12.8 (2.3)	8.6 (2.4)	41.9	<0.001*	0.451
Digit backward	12.4 (1.6)	8.0 (2.1)	73.9	<0.001*	0.592

*Univariate ANOVA F-values for scaled scores from the subtests of the working memory test. The df, error df for the F-values is (1,51). *p-value set to 0.01 to account for multiple comparisons.*

FIGURE 2 | Venn diagram displaying the co-morbidities observed in the children (*n* = 25) with word reading difficulties in the current study.
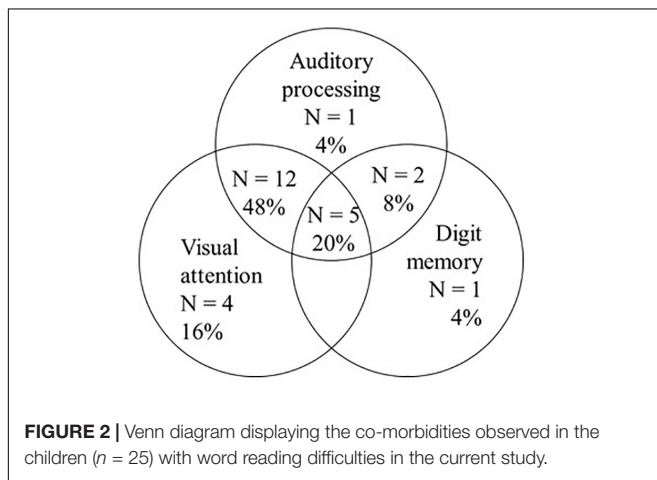
Vocabulary, Visual Attention, Digit Memory

A univariate analysis of variance conducted on children's PPVT-4 standard scores showed that participants with word reading difficulty knew significantly fewer spoken word meanings (99.4 ± 7.9) than children in the control group (110.2 ± 9.3; $F[1,51] = 19.3$, $p < 0.001$).

Overall scores for selective attention and attention switching were determined from the standard scores of eight measures of visual attention (sky search accuracy and time, attention score, map mission, creature counting accuracy and time, same and opposite world). Mean results for selective and switching attention are presented in **Table 4**. Univariate ANOVAs revealed a significant group difference which favored the control children on switching attention, and a smaller yet significant difference on selective attention. An additional two-way ANOVA confirmed the presence of a significant interaction between group and subtest [$F(1,51) = 24.15$, $p < 0.001$]. Children from the control group also achieved significantly higher scores than the children with word reading difficulty on digit memory forward and backward (see **Table 4**).

Subgroup Profiles

In accordance with the second hypothesis, children were allocated to subgroups according to their pattern of comorbid deficits. To determine the individual profiles, we considered

only those tasks for which norms were available. Thus, for auditory processing, FPT, DDdT, GIN and LiSN were considered. Also utilized to define the individual profiles were: the phonological processing task of elision, receptive vocabulary (PPVT), attention, and memory tasks, all of which have published norms.

In general, children attained age appropriate scores on the phonological processing task of elision, with all but a single child scoring within 1 SD of the mean or better. Similarly, all children scored within 1 SD of the mean or better on the PPVT-4 measure of receptive vocabulary. In light of this consistently good performance, the variables of phonological processing and receptive vocabulary were not included for subgrouping purposes.

Figure 2 and **Table 5** show that, of the 25 children with word reading difficulties, 20% (*n* = 5) had comorbid deficits in three variables: auditory processing, visual attention, and digit memory. A larger percentage of children (56%, *n* = 14) had comorbid deficits in two variables: 12 children had auditory processing deficits and visual attention difficulties, and 2 had deficits in auditory processing and digit memory. No child experienced comorbid deficits in only visual attention and digit memory. Finally, six (24%) of the children with word reading difficulties displayed a comorbid deficit in just one other variable: four children had visual attention difficulties, one an auditory processing deficit, and one a deficit in digit memory. An alternative way of thinking about these subgrouping data is that 84% (*n* = 21) of this cohort of children with word reading difficulties had comorbid visual attention problems, and 80% had auditory processing deficits. Further detail regarding the specific deficits displayed by each child with non-word reading difficulties is presented in **Table 5**. This table presents the profiles of the 25 children with non-word reading deficits on word reading, auditory processing, attention, and digit memory. This table shows that, within the total cohort, children presented a tendency to have deficits on multiple auditory processing tasks (*n* = 13) or on both visual attention tasks of switching and selective (*n* = 11).

Correlations Across Auditory Processing Tasks in Children With Word Reading Difficulties

Table 6 presents the Pearson's correlation coefficients between the auditory processing measures. With age taken as covariate,

TABLE 5 | Profiles of the 25 children with non-word reading deficits on word reading, auditory processing, attention, digit memory.

	Age/ Gender	Word reading [Regular/ Irregular]	Auditory processing [FPT, DDdT (Dichotic/Diotic), GIN, LiSN-S (High and Low cue)]	Visual attention [Switching/ Selective]	Backward digit memory
EXP17	8/M	1.5 SD below Regular, Irregular	2 SD below FPT, DDdT and GIN	1 SD below on selective and switching	1 SD below
EXP05	8/M	1.5 SD below Regular	2 SD below FPT, DDdT, GIN and Low cue	2 SD below on selective and switching	1 SD below
EXP06	8/M	1.5 SD below Regular	2 SD below FPT, GIN, and Low cue		
EXP25	8.3/M	1.5 SD below Regular	2 SD below FPT and Low cue	2 SD below on selective and switching	1 SD below
EXP88	8.6/F	1.5 SD below Irregular, 1 SD below Regular	2 SD below on FPT and Low Cue	1 SD below on selective	
EXP60	8.6/F	1.5 SD below Regular, Irregular	2 SD below FPT, DDdT, and Low cue	1 SD below on selective; 2 SD below on switching	
EXP64	8.6/M	1.5 SD below Regular, Irregular	2 SD below FPT and GIN	1 SD below on selective and switching	
EXP09	8.7/F	1.5 SD below Irregular; 1 SD below Regular	2 SD below DDdT and Low cue	1 SD below on switching	
EXP07	8.7/M	1.5 SD below Regular, Irregular	2 SD below FPT, DDdT, and GIN	1 SD below on switching	1 SD below
EXP16	8.8/M	1.5 SD below Regular	2 SD below on FPT	1 SD below on switching	
EXP91	9.1/F	1.5 SD below Regular; 1 SD below Irregular		2 SD below on selective and switching	
EXP33	9.3/M	1.5 SD below Regular; 1 SD below Irregular		2 SD below on selective and switching	
EXP73	9.5/M	1.5 SD below Regular	2 SD below DDdT	2 SD below on selective	
EXP46	9.6/M	1.5 SD below Regular, Irregular	2 SD below on FPT and GIN	2 SD below on selective	
EXP81	9.6/F		2 SD below FPT	2 SD below on selective and switching	
EXP29	9.8/M	1.5 SD below Regular, Irregular	2 SD below FPT	2 SD below on selective and switching	
EXP43	9.8/M	1.5 SD below Regular, Irregular	2 SD below FPT;	1 SD below on selective; 2 SD below on switching	1 SD below
EXP15	10/M	1.5 SD below Regular, Irregular	2 SD below FPT and DDdT		1 SD below
EXP53	10.3/F	1.5 SD below Regular, Irregular	2 SD below DDdT	1 SD below on selective	
EXP39	11/M*	1.5 SD below Regular, Irregular	2 SD below FPT and DDdT	2 SD below on selective and switching	
EXP96	11/F	1 SD below Regular			1 SD below
EXP69	11.1/F	1 SD below Regular	2 SD below on DDdT		1 SD below
EXP54	11.3/M	1.5 SD below Regular, Irregular		2 SD below on switching	
EXP26	11.5/F	1.5 SD below Regular, Irregular	2 SD below FPT, DDdT, GIN, and Low cue	2 SD below on selective and switching	
EXP62	11.7/M	1.5 SD below Regular		2 SD below on selective; 1 SD below on switching	

*All but one child had phonological processing skills to be within 1 SD while all children had vocabulary scores to be within 1 SD. *1 SD below on phonological processing task of elision.*

Pearson correlations showed that FPT was highly correlated to GIN ($r = -0.70$, $p < 0.001$) and FD ($r = -0.79$, $p < 0.001$) but not DDdT ($r = 0.33$, $p = 0.10$). The dichotic score was correlated to the diotic score though ($r = 0.77$, $p < 0.001$). There were no more significant correlations between any of the other auditory processing measures. Furthermore, digit backwards scores were also not significantly correlated with any auditory processing measures (p's > 0.05).

Table 7 presents the Pearson's correlation coefficients for word reading, visual attention, receptive vocabulary, and the phonological processing measure of elision. This table shows no significant correlation between selective attention and attention

switching ($r = 0.29$, $p = 0.21$). This table also shows that none of the word reading measures were correlated with visual attention, receptive vocabulary, or elision (p's > 0.05) Irregular word reading was, however, significantly correlated with regular word reading ($r = 0.655$, $p < 0.001$).

DISCUSSION

The first aim of this research was to confirm the existence of average group differences between children with word reading difficulties and their peers with typical reading skills on a

TABLE 6 | Pearson's correlation between auditory processing measures and the digit backward task.

	Dichotic	Diotic	FPT	GIN	Low cue	High cue	FD	IRN	SAM	Digit backward
Dichotic	1	0.777*	0.339	−0.246	−0.335	−0.445	−0.323	0.115	−0.341	0.345
Diotic		1	0.352	−0.448	−0.245	−0.399	−0.229	0.308	−0.282	0.211
FPT			1	−0.700*	−0.487	−0.424	−0.798*	0.342	−0.395	0.226
GIN				1	0.386	0.371	0.521	−0.449	0.390	−0.083
Low Cue					1	0.524	0.453	−0.300	0.409	0.043
High Cue						1	0.178	−0.260	0.437	−0.139
FD							1	−0.331	0.525	−0.260
IRN								1	−0.455	0.036
SAM									1	−0.332
Digit Backward										1

All the measures in the analysis were raw scores since standardized scores were not available for all auditory processing tasks. Therefore, age was taken as covariate for the analysis.

TABLE 7 | Pearson correlation for word reading, visual attention, receptive vocabulary, and the phonological processing measure of elision.

	Regular word	Irregular word	Non-word	Selective attention	Switching attention	PPVT	Elision
Regular word reading	1	0.655*	0.386	0.006	0.055	−0.192	0.302
Irregular word reading		1	0.264	−0.172	−0.130	−0.045	0.074
Non-word reading			1	−0.177	−0.071	−0.131	0.244
Selective Attention				1	0.293	−0.052	−0.306
Switching Attention					1	−0.195	0.131
PPVT						1	0.077
Elision							1

All standardized scores were used in the analysis. Bonferroni correction was applied for the multiple comparisons in the analysis, and no variables were found to be significantly correlated to each other.

range of auditory processing measures. The second aim was to determine the subgroup profiles of children with word reading difficulties across a set of variables including auditory processing, visual attention, phonological processes, receptive language (vocabulary), and working memory.

Consistent with the previous literature, children with word reading difficulties performed significantly more poorly as a group on auditory processing, phonological processing (elision) receptive language (vocabulary), visual attention, and digit memory compared to children with age appropriate word reading skills. At an individual level, however, the picture is more complicated. For instance, despite significant group mean differences in receptive vocabulary and phonological awareness (elision), no individual child with a word reading difficulty scored more than 1 SD below the typical mean on vocabulary, and just one child with a reading difficulty achieved a score more than 1 SD below the mean on phonological awareness (elision). Clearly, in this case, group mean findings do not provide a reliable indicator of individual outcomes. Furthermore, individual outcomes vary markedly across variables: Some children in the current sample have relatively isolated reading problems (n = 6 with just one comorbid deficit), whereas others have multiple comorbid deficits of varying combinations (n = 19 with two or more comorbidities; see **Figure 2**). It is important to understand the nature of the various comorbidities to advance theoretical and clinical knowledge relating to word reading difficulties, their assessment and possible intervention.

Before considering the results further, it is important to acknowledge that the patterns of comorbidity described here depend critically on the criteria used to identify deficits. As outlined in some detail earlier in the paper (see section Performance Criteria), we adopted performance criteria that were used previously in related research literature. Regarding auditory processing, this approach meant that deficits were identified when scores fell 2 SD or more below the mean. Although the use of minus 2 SD as a criterion for atypical performance has been deemed arbitrary in the auditory processing literature (Dillon et al., 2012; Wilson and Arnott, 2013), it has the advantage of providing a conservative estimate of the occurrence of auditory processing deficits in the current cohort, which seems appropriate given our primary interest in discovering an association between auditory processing deficits and word reading difficulties. Based on previous literature, the current study employed different criteria to identify reading difficulties (−1.5 SD or more below the typical mean) and attention and memory deficits (1 SD or more below the typical mean). If these criteria were modified to −2 SD, the profiles presented here would change substantially in some respects (**Table 5**). For instance, none of the children would be regarded as having a digit memory deficit, and only 14 (56%) children would be considered to have visual attention deficits compared to the current 21 (84%). While this study was not designed to examine the appropriateness of the subgrouping criteria, the issue is central to how one

defines co-morbidity or heterogeneity in the population of children with word reading difficulties. Future studies will need to assess the test–retest reliability of the various tasks used to measure underlying abilities and evaluate the line of demarcation between a "typical" score and an "atypical" score in standard deviation units. In the meantime, it is critical that, as researchers, we are transparent in our choices and provide clear justifications.

Comorbidities in Children With Word Reading Difficulties

As the current results show, group effects can be misleading when they conceal marked variation in the comorbidities seen in a cohort of children with word reading difficulties. Until now, we have considered the pattern of performance across different tasks, but individual scores could also show evidence of variability according to the type of auditory processing deficits seen across individuals (Iliadou et al., 2009; Sharma et al., 2009), or the type of attention deficits, selective and/or switching. All these potential sources of individual variation could clearly be important in accounting for variability in findings across studies.

The high co-morbidities observed in the current cohort can be explained in two ways. First, there may be a causal relationship between one of the measured skills and the remaining associated skills, including word reading ability. This relationship might influence children's competence, such that they do not perform well on one aspect, and as a result also do poorly on some or most other aspects. Some researchers have suggested that attention is the "global" deficit that guides performance across the skills (Moore et al., 2010; Snowling et al., 2018). However, this suggestion does not appear to hold true for the current cohort in which 4 out of 21 children with an attention deficit showed evidence of *no other deficit*, and a further 4 children showed evidence of deficits in auditory processing and/or digit memory, despite having *no attention deficit*. Furthermore, all except one of the 25 children with reading difficulties, including those with visual attention deficits, performed within 1 SD of the typical mean on both phonological processing (elision) and receptive vocabulary. As regards the latter finding, it remains possible that a different pattern of results might have emerged had we used a different measure of phonological processing and/or a more global measure of language ability that did not reflect vocabulary knowledge alone.

A second possible explanation for the high co-morbidities seen in the current study is that non-word reading difficulties co-exist alongside deficits in auditory processing, visual attention, and digit memory, and an altogether different skill, not measured in the current study, underpins these multiple deficits. Moore et al. (2003) raised the possibility of a perceptual learning deficit guiding performance on cognitive, reading, and auditory processing skills. Other possibilities are the effectiveness of reading instruction that children receive, and/or the amount of time that they spend engaged in reading activities. It is a limitation of the current study that time spent reading was not included

in our assessment battery, which was already extensive and time-consuming. It would be useful in future studies to evaluate the association between this more practice-based variable and children's outcomes across the range of measures used here.

Auditory Processing Skills in Children With Word Reading Difficulty

In this study, children with word reading difficulties performed significantly more poorly, on average, on the FD, FPT, DDdT, and LiSN-S tasks compared to children with typical reading skills. These results are consistent with previous research conducted in children with word reading difficulties for FD (Halliday and Bishop, 2006a; McArthur et al., 2008; Goswami et al., 2010), FPT (Sharma et al., 2006, 2009), GIN (Zaidan and Baran, 2013), dichotic listening (Moncrieff and Black, 2007), and speech in noise percept (Bradlow et al., 2003; Ziegler et al., 2009) (which has been measured differentially across the literature).

SAM and IRN did not differ significantly across groups in the current cohort of children. This finding contrasts with that of Rocheron et al. (2002) who found a positive relationship between modulation detection, phonological processing, and reading abilities in typically developing children (Rocheron et al., 2002). However, the study included children who were severely reading-impaired (some performing 5 years below their reading age); and the deficit was measured on a passage reading task and not word reading. Therefore, it is possible that any differences in the types of auditory processing affected are a result of differences in the types of reading disorders considered in the two studies.

Table 5 provides detailed information about the various auditory processing skills affected in children with non-word reading deficits in the current study. Individual profiles show that most children had difficulty on FPT and DDdT consistent with previous research (Sharma et al., 2009). FPT is a complex task that requires children to attend to three tones that differ in frequency and are presented in a particular sequence. The children have to recognize the patterns, and label them in the correct order. The complexity of the task may be one reason for generally poor performance. In a recent study, FPT was found to be a unique predictor of word reading skills in children, which may explain why FPT is generally impacted in the current cohort of children with non-word reading difficulties (Sharma et al., 2019).

The Dichotic Digit difference task requires repetition of four numbers, and is therefore a relatively simpler task than FPT, and yet it was impacted in a similarly large number of children. It would appear, therefore, that children's poor performance is not due solely to the complexity of the task. DDdT is a relatively new measure (Cameron et al., 2016), which includes a dichotic and a diotic listening task. The dichotic-diotic difference is able to provide a measure of dichotic advantage while accounting for cognitive contributors such as attention and memory (Cameron et al., 2016). Dichotic listening ability has been assessed previously using DDT in children with reading deficits and it was observed that children with reading disorders had deficits on DDT

(Moncrieff and Black, 2007; Skarzynski et al., 2015). What is not clear is why DDT should be impacted in children with word reading deficits. Furthermore, Sharma et al. (2019) found that DDT did not contribute to word reading. Hashimoto et al. (2000) used a dichotic and diotic listening task involving phrases, and found, using functional magnetic resonance imaging, that areas such as planum temporale and superior temporal gyrus were activated more during the dichotic listening task than during the diotic task. The authors concluded that auditory areas associated with dichotic listening played a role in speech recognition. Thus, there might be an indirect relationship between dichotic listening and reading ability, with language mediating the link (Hashimoto et al., 2000). This relationship needs to be explored further, especially in light of the current study's finding that DDdT did not correlate with FPT, thus implying that they may be measuring different aspects of auditory processing. At the same time, most of the children with word reading difficulties ($n = 7$, 32%) had both FPT and DDdT deficits.

GIN was another task where the gap threshold of seven children with word reading difficulties was higher than the expected norm. In one previous research study, a link was reported between gap detection and reading skills in children (Walker et al., 2006). More importantly, it is interesting that none of the children in the current cohort had difficulty only on GIN; they had difficulty on FPT as well. Correlations showed that FPT and GIN were highly correlated. While correlations are not indicative of a causal relationship, the associations are informative. FPT does include some level of temporal processing that GIN is also assessing. However, the control group showed only a weak to moderate correlation between FPT and GIN ($r = -0.49$, $p = 0.01$). Perhaps the association between FPT and GIN is driven by other skills along with temporal processing that would account for differences in the associations between the two tasks in the two groups. Alternatively, perhaps the weak to moderate correlation was observed because of the control group's performance being close to ceiling on the FPT and GIN tasks (see **Table 3**).

The LiSN-S low cue listening situation represents the most difficult scenario wherein the target and distractors are acoustically similar and arrive from the same location. Seven children had difficulties on the low cue condition of this task. While LiSN-S has not been used previously to assess speech perception in noise in children with word reading difficulties, other similar tasks have been utilized in this population. For instance, Bradlow et al. (2003) assessed the ability of children with learning difficulties to perceive sentences (similar to LiSN-S) in noise. Consistent with the current results, the study reported that children with learning difficulties performed more poorly than their age-matched peers with typical development on the sentence perception in noise task. In another study, stepwise regression analyses showed that speech perception in noise was a unique predictor of composite word reading score (total performance across regular, irregular and non-word reading) even after phonological processing, attention,

and memory were accounted for in the model (Ziegler et al., 2009). The research also reported that removal of the fine structure of speech resulted in poor speech perception similar to when noise is introduced. The authors concluded that the core deficit in children with dyslexia (reading disorder) was the lack of speech clarity that often occurs in the presence of classroom noise. This finding explains the group results in the current study and provides support for the argument that children with reading difficulties require a higher signal to noise ratio compared to their peers with typical reading skills. However, only seven children (28%) showed deficits on the low cue condition of the LiSN-S task and always with FPT deficits, yet no correlation between the tasks was observed.

In the current study, a cohort of 25 children with non-word reading difficulties participated. It is apparent that a large number of variations exist at the individual level. For instance, while all children had poor non-word reading skills, not every child had regular and irregular word reading problems. Most children with non-word reading difficulties displayed deficits on FPT, but about half showed deficits on either or both low cue of LiSN-S and GIN. It was difficult to determine which of these children would have selective and/or attention switching deficits. It was also unclear why only a third of the cohort had backward digit memory deficits. A bigger dataset collected from children who have specific non-word reading difficulties (i.e., in the absence of real word difficulties) would be useful in attempting to evaluate whether there are common subgroup profiles that can explain variability and assist in designing clinical management programs.

CONCLUSION

The findings from the current study support the hypothesis that children with word reading difficulties have comorbidities across a range of skills, including auditory processing, visual attention, and digit memory. On the standardized tests of auditory processing (FPT, DDdT, LiSN-S, GIN), 80% of children with non-word reading difficulties showed a significant deficit. Although it is difficult to establish a clear link between performance on different tests, it is evident that identifying the presence of multiple deficits in individual children with reading difficulties is key to better management. One cannot assume that children with reading difficulties have a single problem. Equally important, the assumption that individual children with reading difficulties have deficits across *all* areas of functioning is also incorrect. Therefore, a multi-dimensional test battery encompassing a minimum of auditory processing, attention and memory in children with non-word reading difficulties will enable identification of important strengths and weaknesses. From a clinical perspective, the results suggest that the approach to assessment and management of children with word reading difficulties should be multidisciplinary and incorporate assessments of all relevant abilities.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/**Supplementary Files**.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Human Research Ethics Committee Macquarie University. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

RG carried out the data collection. RG conducted the data analysis with inputs from MS and LC. All authors made a substantial, direct and intellectual contribution to the work, and approved it for publication.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02383/full#supplementary-material

REFERENCES

Adlard, A., and Hazan, V. (1998). Speech perception in children with specific reading difficulties (Dyslexia). *Q. J. Exp. Psychol.* 51, 153–177. doi: 10.1080/713755750

Allen, P., and Allan, C. (2014). Auditory processing disorders: relationship to cognitive processes and underlying auditory neural integrity. *Int. J. Pediatr. Otorhinolaryngol.* 78, 198–208. doi: 10.1016/j.ijporl.2013.10.048

American National Standards Institute [ANSI] (1996). *Specifications for Audiometers (ANSI S3.6-1996)*. New York, NY: American National Standards Institute.

American Speech-Language-Hearing Association (2005). *Central Auditory Processing Disorders: The Role of the Audiologist*, Vol. 25. Rockville, MD: American Speech-Language-Hearing Association.

American Speech-Language-Hearing Association [ASHA], (1996). Central auditory processing: current status of research and implications for clinical practice. *Am. J. Audiol.* 5, 41–54.

Australian Government: National Health and Medical Research Council, (2018). *National Statement on Ethical Conduct in Human Research*. Canberra, ACT: Australian Government: National Health and Medical Research Council.

Baddeley, A. (1986). Working memory, reading and dyslexia. *Adv. Psychol.* 34, 141–152. doi: 10.1016/s0166-4115(08)61202-9

Badian, N. A. (1996). Reading disability defined as a discrepancy between listening and reading comprehension?: a longitudinal study of stability, gender differences, and prevalence. *J. Learn. Disabil.* 32, 138–148. doi: 10.1177/002221949903200204

Baird, G., Slonims, V., Simonoff, E., and Dworzynski, K. (2011). Impairment in non-word repetition: a marker for language impairment or reading impairment? *Dev. Med. Child Neurol.* 53, 711–716. doi: 10.1111/j.1469-8749.2011.03936.x

Banai, K., and Ahissar, M. (2004). Poor frequency discrimination probes dyslexics with particularly impaired working memory. *Audiol. Neurootol.* 9, 328–340. doi: 10.1159/000081282

Banai, K., and Ahissar, M. (2006). Auditory processing deficits in dyslexia: task or stimulus related? *Cereb. Cortex* 16, 1718–1728. doi: 10.1093/cercor/bhj107

Barker, M. D., Kuruvilla-Mathew, A., and Purdy, S. C. (2017). Cortical auditory-evoked potential and behavioral evidence for differences in auditory processing between good and poor readers. *J. Am. Acad. Audiol.* 28, 534–545. doi: 10.3766/jaaa.16054

Benasich, A. A., Thomas, J. J., Choudhury, N., and Leppänen, P. H. T. (2002). The importance of rapid auditory processing abilities to early language development. *Brain Behav. Immun.* 40, 278–292. doi: 10.1016/j.bbi.2008.05.010

Bradlow, A. R., Kraus, N., and Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *J. Speech Lang. Hear Res.* 46, 80–97. doi: 10.1044/1092-4388(2003/007)

Cameron, S., and Dillon, H. (2007). Development of the listening in spatialized noise-sentences test (LISN-S). *Int. J. Audiol.* 46, 145–153. doi: 10.1080/14992020601164170

Cameron, S., Glyde, H., Dillon, H., Whitfield, J., and Seymour, J. (2016). The dichotic digits difference test (DDdT): development, normative data, and test-retest reliability studies part 1. *J. Am. Acad. Audiol.* 26, 532–539. doi: 10.3766/jaaa

Casco, C., Tressoldi, P. E., and Dellantonio, A. (1998). Visual selective attention and reading efficiency are related in children. *Cortex* 34, 531–546. doi: 10.1016/S0010-9452(08)70512-4

Casini, L., Pech-Georgel, C., and Ziegler, J. C. (2018). It's about time: revisiting temporal processing deficits in dyslexia. *Dev. Sci.* 21, 1–14. doi: 10.1111/desc.12530

Compton, D. L., Defries, J. C., and Olson, R. K. (2001). Are RAN- and phonological awareness-deficits additive in children with reading disabilities? *Dyslexia* 7, 125–149. doi: 10.1002/dys.198

Dawes, P., Sirimanna, T., Burton, M., Vanniasegaram, I., Tweedy, F., and Bishop, D. V. M. (2009). Temporal auditory and visual motion processing of children diagnosed with auditory processing disorder and dyslexia. *Ear Hear.* 30, 675–686. doi: 10.1097/AUD.0b013e3181b34cc5

Dhamani, I., Leung, J., Carlile, S., and Sharma, M. (2013). Switch attention to listen. *Sci. Rep.* 3:1297. doi: 10.1038/srep01297

Dillon, H., Cameron, S., Glyde, H., Wilson, W., and Tomlin, D. (2012). An opinion on the assessment of people who may have an auditory processing disorder. *J. Am. Acad. Audiol.* 23, 97–105. doi: 10.3766/jaaa.23.2.4

Dunn, L. M., and Dunn, D. M. (2007). *PPVT-4: Peabody Picture Vocabulary Test*, 4th Edn. Bloomington, MN: Pearson Assessments.

Felton, R., Wood, F., Brown, I., Campbell, S., and Harter, M. R. (1987). Separate verbal memory and naming deficits in attention deficit disorder and reading disability. *Brain Lang.* 31, 171–184. doi: 10.1016/0093-934X(87)90067-8

Fischer, B., and Hartnegg, K. (2004). On the development of low-level auditory discrimination and deficits in dyslexia. *Dyslexia* 10, 105–118. doi: 10.1002/dys.268

Franceschini, S., Gori, S., Ruffino, M., Pedrolli, K., and Facoetti, A. (2012). A causal link between visual spatial attention and reading acquisition. *Curr. Biol.* 22, 814–819. doi: 10.1016/j.cub.2012.03.013

Gibson, L. Y., Hogben, J. H., and Fletcher, J. (2006). Visual and auditory processing and component reading skills in developmental dyslexia. *Cogn. Neuropsychol.* 23, 621–642. doi: 10.1080/02643290500412545

Goswami, U. (2001). Early phonological development and the acquisition of literacy. in *Handbook of Early Literacy Research,* eds D. K. Dickinson, and S. B. Neuman, Vol 1, (New York, NY: Guilford Press), 111–125.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001

Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nat. Rev. Neurosci.* 16, 43–54. doi: 10.1038/nrn3836

Goswami, U., Thomson, J., Richardson, U., Stainthorp, R., Hughes, D., Rosen, S., et al. (2002). Amplitude envelope onsets and developmental dyslexia: a new hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 10911–10916. doi: 10.1073/pnas.122368599

Goswami, U., Wang, H.-L. S., Cruz, A., Fosker, T., Mead, N., and Huss, M. (2010). Language-universal sensory deficits in developmental dyslexia: english, Spanish, and Chinese. *J. Cogn. Neurosci.* 23, 325–337. doi: 10.1162/jocn.2010.21453

Grassi, M., and Soranzo, A. (2009). MLP: a MATLAB toolbox for rapid and reliable auditory threshold estimation. *Behav. Res. Methods* 41, 20–28. doi: 10.3758/BRM.41.1.20

Gyldenkærne, P., Dillon, H., Sharma, M., and Purdy, S. C. (2014). The relationship between auditory processing disorders and attention deficits. *J. Am. Acad. Audiol.* 25, 676–687. doi: 10.3766/jaaa.25.7.6

Halliday, L. F., and Bishop, D. V. M. (2006a). Auditory frequency discrimination in children with dyslexia. *J. Res. Read.* 29, 213–228. doi: 10.1111/j.1467-9817.2006.00286.x

Halliday, L. F., and Bishop, D. V. M. (2006b). Is poor frequency modulation detection linked to literacy problems? A comparison of specific reading disability and mild to moderate sensorineural hearing loss. *Brain Lang.* 97, 200–213. doi: 10.1016/j.bandl.2005.10.007

Hämäläinen, J. A., Leppänen, P. H. T., Eklund, K., Thomson, J., Richardson, U., Guttorm, T. K., et al. (2009). Common variance in amplitude envelope perception tasks and their impact on phoneme duration perception and reading and spelling in Finnish children with reading disabilities. *Appl. Psycholinguist.* 30, 511–530. doi: 10.1017/S0142716409090250

Hämäläinen, J. A., Salminen, H. K., and Leppänen, P. H. T. (2012). Basic Auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential / field evidence. *J. Learn. Disabil.* 46, 413–427. doi: 10.1177/0022219411436213

Hashimoto, R., Homae, F., Nakajima, K., Miyashita, Y., and Sakai, K. L. (2000). Functional differentiation in the human auditory and language areas revealed by a dichotic listening task. *Neuroimage* 12, 147–158. doi: 10.1006/nimg.2000.0603

Huss, M., Verney, J. P., Fosker, T., Mead, N., and Goswami, U. (2011). Music, rhythm, rise time perception and developmental dyslexia: perception of musical meter predicts reading and phonology. *Cortex* 47, 674–689. doi: 10.1016/j.cortex.2010.07.010

Iliadou, V., Bamiou, D. E., Kaprinis, S., Kandylis, D., and Kaprinis, G. (2009). Auditory processing disorders in children suspected of learning disabilities-a need for screening? *Int. J. Pediatr. Otorhinolaryngol.* 73, 1029–1034. doi: 10.1016/j.ijporl.2009.04.004

Jerger, J., and Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing disorders in school-aged children. *J. Am. Acad. Audiol.* 11, 467–474.

Krishnan, A., Gandour, J. T., and Suresh, C. H. (2014). Cortical pitch response components show differential sensitivity to native and nonnative pitch contours. *Brain Lang.* 138, 51–60. doi: 10.1016/j.bandl.2014.09.005

Landerl, K., and Willburger, E. (2010). Temporal processing, attention, and learning disorders. *Learn. Individ. Differ.* 20, 393–401. doi: 10.1016/j.lindif.2010.03.008

Leppänen, P. H. T., Hämäläinen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., et al. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex* 6, 1362–1376. doi: 10.1016/j.cortex.2010.06.003

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Robertson, I. H., Watson, P., et al. (2001). The differential assessment of children's attention: the test of everyday attention for children (TEA-Ch), normative sample and ADHD performance. *J. Child Psychol. Psychiatry* 42, 1065–1081. doi: 10.1111/1469-7610.00806

Massa, I., and Rivera, V. (2009). Test reviews: Wechsler, D., *& Naglieri, J.A.* (2006). wechsler nonverbal scale of ability. San Antonio, TX: Harcourt Assessment. *J. Psychoeduc. Assess.* 27, 426–432. doi: 10.1177/0734282908329108

McArthur, G. M., and Bishop, D. V. M. (2004). Frequency discrimination deficits in people with specific language impairment: reliability, validity, and linguistic correlates. *J. Speech Lang. Hear. Res.* 47, 527–541. doi: 10.1044/1092-4388(2004/041)

McArthur, G. M., Ellis, D., Atkinson, C. M., and Coltheart, M. (2008). Auditory processing deficits in children with reading and language impairments: can they (and should they) be treated? *Cognition* 107, 946–977. doi: 10.1016/j.cognition.2007.12.005

McArthur, G. M., and Hogben, J. H. (2012). Poor auditory task scores in children with specific reading and language difficulties: some poor scores are more equal than others. *Sci. Stud. Read.* 16, 63–89. doi: 10.1080/10888438.2010.542526

Medwetsky, L., Burkard, R. F., and Hood, L. (2009). *Handbook of Clinical Audiology*, ed. J. Katz, (Philadelphia, PA: Lippincott William & Wilkins).

Metsala, J. L. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *J. Educ. Psychol.* 91, 3-19. doi: 10.1037/0022-0663.91.1.3

Moncrieff, D. W., and Black, J. R. (2007). Dichotic listening deficits in children with dyslexia. *Dyslexia* 14, 54–75. doi: 10.1002/dys

Moore, D., Amitay, S., and Hawkey, D. (2003). Auditory perceptual learning. *Learn. Mem.* 10, 83–85. doi: 10.1101/lm.59703.were

Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., and Ferguson, M. A. (2011). Development of Auditory Processing in 6- to 11-Yr-Old Children. *Ear Hear.* 32, 269–285. doi: 10.1097/AUD.0b013e318201c468

Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., and Riley, A. (2010). Nature of auditory processing disorder in children. *Pediatrics* 126, e382–e390. doi: 10.1542/peds.2009-2826

Moore, D. R., Ferguson, M. A., Halliday, L. F., and Riley, A. (2008). Frequency discrimination in children: perception, learning and attention. *Hear. Res.* 238, 147–154. doi: 10.1016/j.heares.2007.11.013

Musiek, F. E. (2002). The frequency pattern test: a guide. *Hear. J.* 55:58. doi: 10.1097/01.hj.0000293280.99394.dd

Muter, V., Hulme, C., Snowling, M. J., and Stevenson, J. (2004). Phonemes, rimes, vocabulary, and grammatical skills as foundations of early reading development: evidence from a longitudinal study. *Dev. Psychol.* 40, 665–681. doi: 10.1037/0012-1649.40.5.665

Ouellette, G. P. (2006). What's meaning got to do with it: the role of vocabulary in word reading and reading comprehension. *J. Educ. Psychol.* 98, 554–566. doi: 10.1037/0022-0663.98.3.554

Paul, I., Bott, C., Heim, S., Wienbruch, C., and Elbert, T. R. (2006). Phonological but not auditory discrimination is impaired in dyslexia. *Eur. J. Neurosci.* 24, 2945–2953. doi: 10.1111/j.1460-9568.2006.05153.x

Perfetti, C. A., Marron, M. A., and Foltz, P. W. (1996). "Sources of comprehension failure: theoretical perspectives and case studies," in *Reading Comprehension Difficulties: Processes and Intervention*, eds C. Cornoldi, and J. Oakhill, (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 137–165.

Peter, V., Wong, K., Narne, V. K., Sharma, M., Purdy, S. C., and McMahon, C. (2014). Assessing spectral and temporal processing in children and adults using temporal modulation transfer function (TMTF), iterated ripple noise (IRN) perception, and spectral ripple discrimination (SRD). *J. Am. Acad. Audiol.* 25, 210–218. doi: 10.3766/jaaa.25.2.9

Rabiner, D. L., and Coie, J. D. (2000). Early attention problems and children's reading achievement: a longitudinal investigation. *J. Am. Acad. Child Adolesc. Psychiatry* 39, 859–867. doi: 10.1097/00004583-200007000-00014

Ramus, F. (2001). Dyslexia: talk of two theories. *Nature* 412, 393–395. doi: 10.1038/35086683

Reid, M., Johnson, B., McArthur, G., Castles, A., and Hautus, M. (2010). "Auditory Processing in the Dyslexic brain," in *Proceedings of the 9th Conference of the Australasian Society for Cognitive Science, 293–297*. Sydney: Macquarie Centre for Cognitive Science.

Rocheron, I., Lorenzi, C., and Dumont, A. (2002). Temporal envelope perception in dyslexic children. *Children* 13, 3–7.

Scarborough, H. S. (1990). very early language deficits in dyslexic children. *Child Dev.* 61, 1728–1743. doi: 10.1111/j.1467-8624.1990.tb03562.x

Scarborough, H. S., Neuman, S., and Dickinson, D. (2009). "Connecting early language and literacy to later reading (dis) abilities: evidence, theory, and practice," in *Approaching Difficulties in Literacy Development: Assessment, Pedagogy, and Programmes*, eds F. Fletcher-Campbell, J. Soler, and G. Reid, (London: Sage), 23–39.

Semel, E., Wiig, E., and Secord, W. (2006). *Clinical Evaluation of Language Fundamentals Fourth Edition, Australian Standardised Edition (CELF-4 Australian)*. Marrickville, NSW: Harcourt Assessment.

Sharma, M., Cupples, L., and Purdy, S. (2019). Predictors of reading skills in children with listening concerns. *Ear Hear.* 40, 243–252. doi: 10.1097/AUD.0000000000000608

Sharma, M., Dhamani, I., Leung, J., and Carlile, S. (2014). Attention, memory, and auditory processing in 10- to 15-Year-Old children with listening difficulties. *J. Speech Lang. Hear Res.* 57, 2308–2321. doi: 10.1044/2014

Sharma, M., Purdy, S. C., and Kelly, A. S. (2009). Comorbidity of auditory processing, language, and reading disorders. *J. Speech Lang. Hear Res.* 52, 706–722. doi: 10.1044/1092-4388(2008/07-0226)

Sharma, M., Purdy, S. C., Newall, P., Wheldall, K., Beaman, R., and Dillon, H. (2006). Electrophysiological and behavioral evidence of auditory processing deficits in children with reading disorder. *Clin. Neurophysiol.* 117, 1130–1144. doi: 10.1016/j.clinph.2006.02.001

Shinn, J. B., Chermak, G. D., and Musiek, F. E. (2009). GIN (Gaps-In-Noise) performance in the pediatric population. *J. Am. Acad. Audiol.* 20, 229–238. doi: 10.3766/jaaa.20.4.3

Skarzynski, P. H., Wlodarczyk, A. W., Kochanek, K., Pilka, A., Jedrzejczak, W. W., Olszewski, L., et al. (2015). Central auditory processing disorder (CAPD) tests in a school-age hearing screening programme – analysis of 76,429 children. *Ann. Agric. Environ. Med.* 22, 90–95. doi: 10.5604/12321966.1141375

Snowling, M., Gooch, D., McArthur, G., and Hulme, C. (2018). Language skills, but not frequency discrimination, predict reading skills in children at risk of dyslexia. *Psychol. Sci.* 29, 1270–1282. doi: 10.1177/0956797618763090

Stevens, C., and Bavelier, D. (2012). The role of selective attention on academic foundations: a cognitive neuroscience perspective. *Dev. Cogn. Neurosci.* 2S(Suppl. 1), S30–S48. doi: 10.1016/j.dcn.2011.11.001

Swanson, H. L. (1993). Working memory in learning disability subgroups. *J. Exp. Child Psychol.* 56, 87–114. doi: 10.1006/jecp.1993.1027

Swanson, H. L., Cochran, K. F., and Ewers, C. A. (1989). Working memory in skilled and less skilled readers. *J. Abnorm. Child Psychol.* 17, 145–156. doi: 10.1007/BF00913790

Swanson, H. L., Zheng, X., and Jerman, O. (2009). Working memory, short-term memory, and reading disabilities. *J. Learn. Disabil.* 42, 260–287. doi: 10.1177/0022219409331958

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* 9, 182–198. doi: 10.1016/0093-934X(80)90139-X

Tallal, P. (1984). Temporal or phonetic processing deficit in dyslexia? That is the question. *Appl. Psycholinguist.* 5, 167–169. doi: 10.1017/S0142716400004963

Tomlin, D., Dillon, H., Sharma, M., and Rance, G. (2015). The impact of auditory processing and cognitive abilities in children. *Ear Hear.* 36, 527–542. doi: 10.1097/AUD.0000000000000172

Wagner, R., Torgesen, J., and Rashotte, C. (1999). *Comprehensive Test of Phonological Processing*. Austin,TX: Pro-Ed.

Wagner, R. K., and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychol. Bull.* 101, 192–212. doi: 10.1037/0033-2909.101.2.192

Walker, M. M., Givens, G. D., Cranford, J. L., Holbert, D., and Walker, L. (2006). Auditory pattern recognition and brief tone discrimination of children with reading disorders. *J. Commun. Disord.* 39, 442–455. doi: 10.1016/j.jcomdis.2005.12.006

Walley, A. C., Metsala, J. L., and Victoria, M. (2003). Spoken vocabulary growth: Its role in the development of phoneme awareness and early reading ability. *Read. Writ. Q.* 16, 5–20.

Wechsler, D., and Naglieri, J. A. (2006). *Wechsler Nonverbal Scale of Ability*. San Antonio, TX: Harcourt Assessment.

Whitehurst, G. J., and Lonigan, C. J. (1998). Child development and emergent literacy. *Child Dev.* 69, 848–872. doi: 10.1111/j.1467-8624.1998.tb06247.x

Wible, B., Nicol, T., and Kraus, N. (2005). Correlation between brainstem and cortical auditory processes in normal and language-impaired children. *Brain* 128, 417–423. doi: 10.1093/brain/awh367

Willcutt, E. G., and Pennington, B. F. (2000). Comorbidity of reading disability and attention-deficit/hyperactivity disorder: differences by gender and subtype. *J. Learn. Disabil.* 33, 179–191. doi: 10.1177/002221940003300206

Willcutt, E. G., Pennington, B. F., Olson, R. K., Chhabildas, N., and Hulslander, J. (2005). Neuropsychological analyses of comorbidity between reading disability and attention deficit hyperactivity disorder: in search of the common deficit. *Dev. Neuropsychol.* 27, 35–78. doi: 10.1207/s15326942dn2701

Wilson, W. J., and Arnott, W. (2013). Using different criteria to diagnose (central) auditory processing disorder : how big a difference does it make? *J. Speech Lang. Hear. Res.* 56, 63–70. doi: 10.1044/1092-4388(2012/11-0352)b

Wise, J. C., Sevcik, R. A., Morris, R. D., Lovett, M. W., and Wolf, M. (2007). The relationship among receptive and expressive vocabulary, listening comprehension, pre-reading skills, word identification skills, and reading comprehension by children with reading disabilities. *J. Learn. Disabil.* 50, 1093–1109. doi: 10.1177/002221948301600203

Wright, C. M., and Conlon, E. G. (2009). Auditory and visual processing in children with dyslexia. *Dev. Neuropsychol.* 34, 330–355. doi: 10.1080/87565640902801882

Zaidan, E., and Baran, J. A. (2013). Gaps-in-noise (GIN©) test results in children with and without reading disabilities and phonological processing deficits. *Int. J. Audiol.* 52, 113–123. doi: 10.3109/14992027.2012.733421

Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (2009). Speech-perception-in-noise deficits in dyslexia. *Dev. Sci.* 12, 732–745. doi: 10.1111/j.1467-7687.2009.00817.x

Longitudinal Speech Recognition in Noise in Children: Effects of Hearing Status and Vocabulary

Elizabeth A. Walker[1]*, Caitlin Sapp[1], Jacob J. Oleson[2] and Ryan W. McCreery[3]

[1]Pediatric Audiology Laboratory, Department of Communication Sciences and Disorders, University of Iowa, Iowa City, IA, United States, [2]Department of Biostatistics, University of Iowa, Iowa City, IA, United States, [3]Center for Hearing Research, Audibility, Perception, and Cognition Laboratory, Boys Town National Research Hospital, Omaha, NE, United States

Objectives: The aims of the current study were: (1) to compare growth trajectories of speech recognition in noise for children with normal hearing (CNH) and children who are hard of hearing (CHH) and (2) to determine the effects of auditory access, vocabulary size, and working memory on growth trajectories of speech recognition in noise in CHH.

Design: Participants included 290 children enrolled in a longitudinal study. Children received a comprehensive battery of measures annually, including speech recognition in noise, vocabulary, and working memory. We collected measures of unaided and aided hearing and daily hearing aid (HA) use to quantify aided auditory experience (i.e., HA dosage). We used a longitudinal regression framework to examine the trajectories of speech recognition in noise in CNH and CHH. To determine factors that were associated with growth trajectories for CHH, we used a longitudinal regression model in which the dependent variable was speech recognition in noise scores, and the independent variables were grade, maternal education level, age at confirmation of hearing loss, vocabulary scores, working memory scores, and HA dosage.

Results: We found a significant effect of grade and hearing status. Older children and CNH showed stronger speech recognition in noise scores compared to younger children and CHH. The growth trajectories for both groups were parallel over time. For CHH, older age, stronger vocabulary skills, and greater average HA dosage supported speech recognition in noise.

Conclusion: The current study is among the first to compare developmental growth rates in speech recognition for CHH and CNH. CHH demonstrated persistent deficits in speech recognition in noise out to age 11, with no evidence of convergence or divergence between groups. These trends highlight the need to provide support for children with all degrees of hearing loss in the academic setting as they transition into secondary grades. The results also elucidate factors that influence growth trajectories for speech recognition in noise for children; stronger vocabulary skills and higher HA dosage supported speech recognition in degraded situations. This knowledge helps us to develop a more comprehensive model of spoken word recognition in children.

Keywords: children, vocabulary, working memory, hearing loss, speech recognition

INTRODUCTION

Every year, approximately three in 1,000 children are born with a significant hearing loss (Mehra et al., 2009). Children who are hard of hearing (CHH) have sufficient residual hearing to benefit from amplification. With the advent of newborn hearing screening, they are now being identified and fitted with hearing aids (HAs) during infancy (Holte et al., 2012). Early access to technology and services is posited to have a positive, long-term impact on functional outcomes, which results in the vast majority of CHH being educated in regular education settings (Page et al., 2018). As most CHH rely entirely on spoken language to communicate, they face significant challenges as they enter classrooms that are likely to have poor acoustics (Knecht et al., 2002). Most academic and extracurricular settings are characterized by background noise, which negatively affects speech recognition and academic outcomes in children with normal hearing (CNH), and has even greater consequences for CHH. Even though CHH have documented weaknesses with listening in noise (Crandell, 1993; Uhler et al., 2011; Caldwell and Nittrouer, 2013; Leibold et al., 2013; McCreery et al., 2015; Klein et al., 2017; Ching et al., 2018), there is little research on how their ability to recognize speech in noise develops over time during the school-age years. Increased knowledge in this area impacts both clinical decision-making and theoretical understanding of the mechanisms that drive listening in noise. The goals of the current study are twofold: (1) to investigate growth rates in speech recognition in noise for school-age CHH and CNH, and (2) to investigate the impact of auditory access and cognitive-linguistic abilities on CHH's ability to listen in adverse acoustic conditions over time.

Given their reduced access to spectral and temporal cues in the speech signal, as well as reduced binaural processing, it is not surprising that listening in noise is a challenge for CHH. McCreery et al. (2015) examined word and phoneme recognition in noise in 7- to 9-year-old CHH and age-matched CNH. Even with amplification, CHH rarely reached the same level of performance as CNH in noise. Caldwell and Nittrouer et al. (2013) evaluated kindergartners with normal hearing, HAs, or cochlear implants (CIs) on measures of speech recognition in quiet and in noise and found significant group differences in favor of CNH. The question that then arises is whether children with hearing loss can eventually catch up with their peers, if the gap in speech recognition in noise widens over time, or if they show persistent but stable deficits in recognizing speech in noise. Given that adults with hearing loss show difficulties with listening in noise (Dubno, 2015), we would predict that CHH will not show speech recognition scores that are commensurate with CNH. On the other hand, CNH might reach a floor level on speech recognition in noise tasks, allowing CHH to eventually close the gap. It also seems improbable that the gap in speech recognition would widen over time; however, a recent study by Walker et al. (2019) indicated an increasing gap with age between CHH and CNH in identifying words during a gating paradigm. The third option, parallel growth rates between CHH and CHH, would appear to be the most reasonable prediction given what we know

from previous research. This hypothesis has not been tested empirically, however, because much of the research to date is cross-sectional or has too few subjects or data points to conduct longitudinal analyses. Thus, there is minimal knowledge about the developmental aspects of speech recognition for CHH compared to CNH, or the cognitive and peripheral factors that support growth in listening skills over time. The question of developmental trajectories in speech recognition in noise can only be effectively addressed with longitudinal data sets, which are lacking in the research literature on CHH.

In addition to limited longitudinal data, previous large-scale studies of speech recognition in children with hearing loss have focused primarily on children with congenital, severe-profound hearing loss who use CIs (Davidson et al., 2011; Robinson et al., 2012; Ching et al., 2014; Dunn et al., 2014; Easwar et al., 2018). CHH are either excluded from these research studies or combined with children who are deaf, making it difficult to isolate the effects of mild to severe hearing loss on speech recognition. The studies that have been conducted with CHH have some limitations. First, children have been tested with words in quiet, rather than word or sentence recognition in noise (Stiles et al., 2012). Identifying monosyllabic words in quiet is not representative of the everyday listening experiences of children (Magimairaj et al., 2018) and may restrict individual differences for CHH, as many of these children will perform at or near ceiling levels (McCreery et al., 2015). Furthermore, speech recognition testing with background noise more accurately reflects listening experiences in realistic settings than monosyllabic word recognition in quiet (Kirk et al., 2012; Hillock-Dunn et al., 2014). Monosyllabic word recognition in quiet has minimal cognitive and linguistic processing demands, which are required in real-world listening environments (Walker et al., 2019). A second limitation of the prior research is that the focus is often on the influence of age at confirmation of hearing loss or age at amplification on speech recognition in noise (Sininger et al., 2010; Ching et al., 2013). Although it is important to evaluate the effectiveness of early hearing detection and intervention services, it is also important to understand the combined effects of auditory access, cognitive and linguistic abilities on listening development. There has been a great deal of attention directed toward understanding speech recognition skills in children with hearing loss, but we still lack a clear understanding of the mechanisms that drive developmental growth.

In environments with degraded signals (either due to poor acoustics or reduced hearing levels), listeners rely on higher level cognitive and linguistic skills to interpret information about the input (Nittrouer and Boothroyd, 1990). According to the Ease of Language Understanding (ELU) model (Rönnberg et al., 2013), adults with higher cognitive skills compensate in listening situations with distorted or missing information because they can use their memory and linguistic skills to repair the distorted signal (Akeroyd, 2008; Rönnberg et al., 2008; Tun et al., 2010; Zekveld et al., 2011). The findings supporting the predictions of the ELU model in children are mixed. Lalonde and Holt (2014) reported that parent report measures of working memory were positively correlated with speech recognition in quiet with 2-year-old CNH. McCreery et al. (2017) evaluated monosyllabic

word and sentence recognition in noise for 96 5- to 12-year-old CNH. Children with higher working memory skills (measured as a combination of complex visual and verbal working memory span scores) had better speech recognition in noise skills than children with lower working memory. On the other hand, there are several studies that do not support the predictions of the ELU model in children. Eisenberg et al. (2000) did not find an association between working memory capacity (measured with forward digit span) and spectrally degraded speech recognition in CNH after controlling for age. Magimairaj et al. (2018) also did not find that working memory capacity (measured with forward digit span, auditory working memory, and complex working memory span tasks) was predictive of speech recognition in noise for 7- to 11-year-old CNH. The differences in findings may be due to the predictor variables and/or the outcome measures. Eisenberg et al. used a short-term working memory test (i.e., storage only), as opposed to complex working memory span measures (i.e., storage and processing). The proponents of the ELU model have posited that simple span tests like digit span are not good predictors of speech recognition (Rönnberg et al., 2013). McCreery et al. used sentences with no semantic context (which increased the memory load), whereas Magimairaj et al. used the Bamford-Kowal-Bench Speech in Noise sentences (BKB-SIN; Bench et al., 1979) which include semantic cues. It is also important to note that the effects of complex working memory span have not been thoroughly explored in CHH. More studies are needed to disentangle the associations between working memory, language, auditory access, and speech recognition in noise for children with hearing loss, who are most impacted by degraded acoustic input.

In addition to exploring the role of working memory capacity, the ELU model also predicts that language abilities will influence the ability to recognize degraded speech (Zekveld et al., 2011). Performance on sentence repetition tasks (which are used to measure speech recognition in noise) is likely tied to oral language skills (Klem et al., 2015). Stronger language skills allow individuals to make better predictions about an incoming message, even in the presence of limited sensory input (Nittrouer et al., 2013). Vocabulary knowledge accounts for a significant proportion of variance in word and sentence recognition in quiet for children with CIs and/or HAs (Blamey et al., 2001; Caldwell and Nittrouer, 2013), and language skills are significant predictors of speech recognition in noise for school-age CHH (McCreery et al., 2015; Klein et al., 2017; Ching et al., 2018). None of these studies included longitudinal data, so it was not possible to determine how these underlying mechanisms influence developmental trajectories of speech recognition in noise. In contrast to the former studies, Magimairaj et al. (2018) did not find that language skills were related to BKB-SIN scores, which they interpreted as an indication that speech recognition in noise is dissociated from language on that clinical measure. They did not include CHH as participants, however, and their language metric was a combined measure of receptive and expressive vocabulary, language comprehension, sentence recall, and inference-making. Thus, their composite language measure may have lacked sensitivity and masked variability, resulting in their reported finding of a dissociation between language and speech recognition in noise.

A third relevant factor to consider when examining sources of variance in speech recognition in noise is auditory access, particularly because CHH show large individual differences in this variable (McCreery et al., 2013; Walker et al., 2013). Auditory access has been explored as a predictor in several ways. One method is to use degree of hearing loss (i.e., pure tone average; PTA) as a predictor. Blamey et al. (2001) found that lower PTA was associated with better speech recognition in noise in children with moderate to profound hearing loss. In contrast, Sininger et al. (2010) examined auditory outcomes in young children with mild to profound hearing loss and found that PTA did not contribute to speech recognition skills. These mixed results may be related to the fact that PTA does not capture the everyday aided listening experiences of CHH. Because PTA measures only unaided audibility for very soft sounds, it does not reflect a child's access to supra-threshold speech while wearing HAs.

An alternative to relying on PTA is to examine audibility levels, as measured by the Speech Intelligibility Index (SII). SII is a measurement that describes the proportion of speech accessible to the listener, with or without HAs. It accounts for the configuration of hearing loss, differences in ear canal size, and amplification characteristics of HAs. Studies have shown an association between SII and speech recognition in CHH (Stelmachowicz et al., 2000; Davidson and Skinner, 2006; Scollie, 2008; Stiles et al., 2012; McCreery et al., 2017); however, Ching et al. (2018) reported that aided SII did not contribute any additional variance to speech recognition in noise for 5-year-old CHH, after controlling for unaided hearing thresholds, non-verbal intelligence, and language skills. Children with HAs in the Ching et al. study were fitted within 3 dB of HA prescriptive targets, which likely reduced variability in SII.

A third way to examine auditory access is to consider individual differences in the amount of daily HA use. Only a few studies have looked at hours of HA use as a predictor variable of speech recognition in noise. McCreery et al. (2015) found that children with more hours of HA use showed higher scores on parent report measures of auditory skills and word recognition in quiet for toddlers and preschoolers with hearing loss. In contrast, Klein et al. (2017) did not find an effect of HA use on word and nonword recognition in school-age CHH. They acknowledged, however, that there was little variability in this factor among the participants, who were mostly consistent HA users.

To better understand the impact of auditory access on listening in noise, we propose to conceptualize the auditory experience of CHH as a combination of unaided hearing, aided SII, and amount of HA use (Walker et al., accepted). Our past studies showed that CHH demonstrate large individual differences in aided audibility (McCreery et al., 2013) and amount of HA use (Walker et al., 2013) over time. We have found unique effects of unaided SII, aided SII and amount of HA use on listening and language outcomes (McCreery et al., 2015; Tomblin et al., 2015), but we have not empirically tested the combined effects of these three factors on speech recognition. In pursuit of this goal, we have developed a metric we call hearing aid (HA) dosage. The concept of dosage has been applied to pharmacological and child language intervention

research to study the effect of different treatment intensities (Warren et al., 2007), but it has not been utilized in the literature on childhood hearing loss. Combining HA dosage measures with longitudinal data on speech recognition in noise for children with hearing loss can inform us of the long-term effects of specific approaches to intervention and auditory access. For example, it is unclear whether higher HA dosage levels averaged across time is sufficient to support the development of speech recognition in noise, or whether fluctuations in auditory access (either due to inconsistency with wearing HAs or changes in hearing levels or aided audibility) could have a negative impact on listening in noise. The need to demonstrate the effects of aided auditory access is particularly relevant for school age children, some of whom receive less academic support in later grades (Page et al., 2018; Klein et al., 2019) and are at risk for inconsistent HA use in the classroom as they enter adolescence (Gustafson et al., 2015). Greater knowledge of the effects of HA dosage on speech recognition in noise can guide implementation of effective interventions for children with hearing loss and has the potential to motivate parents, teachers, and service providers to encourage increased HA usage.

In summary, no studies have compared developmental trajectories in speech recognition in noise between CNH and CHH. This paper describes results from a longitudinal study in which speech recognition in noise measures were collected on an annual basis in school-age CNH and CHH. The aims of this study were to: (1) compare the growth rates for speech recognition in noise for CNH and CHH, (2) determine whether CHH and CNH show similar growth rates over time, and (3) identify the auditory, cognitive, and linguistic factors that are associated with individual differences in growth rates for speech recognition in noise for CHH. It is expected that this knowledge will provide us with further insight into the everyday functional listening skills of children with and without hearing loss.

METHOD

Participants

Participants included 290 children (CHH, $n = 199$; CNH, $n = 92$) who were enrolled in a multicenter, longitudinal study on outcomes of children with mild to severe hearing loss, Outcomes of School-Age Children who are Hard of Hearing (OSACHH). The primary recruitment sites were the University of Iowa, Boys Town National Research Hospital, and University of North Carolina-Chapel Hill. Some of the children from the Iowa and Boys Town test sites also participated in a second longitudinal project that was conducted during the same time period as OSACHH. This second project was called Complex Listening in School-Age Hard of Hearing Children.

CHH had a permanent bilateral hearing loss with a better-ear four-frequency PTA in the mild to moderately severe range. One hundred seventy-nine children had a sensorineural or mixed hearing loss, 15 had a conductive hearing loss, and two had auditory neuropathy spectrum disorder. Three children did not have the type of hearing loss reported. Both CHH and CNH used spoken English as the primary communication

mode and had no major vision, motor, or cognitive impairments. CNH and CHH were matched by age. There was no significant between-group difference in maternal education level [$t(130) = -1.61$, $p = 0.11$]. Demographic information, including audiologic data for the CHH, is provided in **Table 1**.

Data reported in the current analyses occurred when the children were approximately 7, 8, 9, or 10 years of age (respectively, first, second, third, or fourth grade). Children were seen for Complex Listening during first and third grade and OSACHH during second and fourth grade. All participants had completed the BKB-SIN (see description below) during at least one visit over the course of the studies.

Procedures

This study was carried out in accordance with the recommendations of the University of Iowa Institutional Review Board, with written informed consent from all subjects. All parents of the participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the University of Iowa Institutional Review Board.

For the current analysis, participants contributed data from the BKB-SIN at up to four visits: first grade (CHH, $n = 74$; CNH, $n = 44$); second grade (CHH, $n = 145$; CNH, $n = 79$); third grade (CHH, $n = 93$; CNH, $n = 56$); and fourth grade (CHH, $n = 128$; CNH, $n = 69$). Because participants entered the study at different time points, they varied in terms of their number of visits. Furthermore, some participants missed visits between years. We had 88 CHH and CNH with one visit, 63 with two visits, 82 with three visits, and 57 with four visits.

Audiology Measures

Audiologic measures, HA measures, and speech recognition in noise tests were collected at every visit. For CHH, a trained clinician obtained air-conduction thresholds at 250, 500, 1,000, 2,000, 4,000, 6,000, and 8,000 Hz. Bone-conduction thresholds were obtained at 500, 1,000, 2,000, and 4,000 Hz. The four-frequency (500, 1,000, 2,000, and 4,000 Hz) better-ear pure-tone average (BEPTA) was then calculated. CNH passed a hearing screen in both ears at 20 dB HL at these four frequencies.

TABLE 1 | Demographic characteristics for children who are hard of hearing (CHH) and children with normal hearing (CNH).

Variable	CHH		CNH	
	Mean (SD)	Range	Mean (SD)	Range
BEPTA (dB HL)*	46.16 (15.18)	7.5–90.0	<20	
Aided BESII	0.79 (0.14)	0.36–0.99	N/A	
HA dosage	10.48 (4.19)	2.56–23.91	N/A	
Age at confirmation (months)	16.75 (21.21)	0.00–92	N/A	
Age at HA fitting (months)	19.47 (21.59)	1–95	N/A	
Maternal education level (years)	15.40 (2.48)	8–22	16.05 (3.44)	8–22

*BEPTA, better-ear pure-tone average in dB hearing level (HL); BESII, better-ear speech intelligibility index; HA, hearing aid; N/A, not applicable. *The criterion for study enrollment for children who were hard of hearing was BEPTA no better than 25 dB HL. Exceptions were made to include children with mild high-frequency HL (3-frequency PTA less than 25 dB HL in the better ear, but thresholds greater than 25 dB HL at 3, 4, or 6 kHz).*

Hearing Aid Verification

At each visit, the audiologist verified that participants' HAs were functioning appropriately. The SII (ANSI, 1997) was calculated for both ears to estimate the speech audibility based on ear canal acoustics (measured real-ear-to-coupler difference or age-average real-ear-to-coupler difference) and hearing thresholds. SII represents access of the audible speech spectrum at a conversational speech level (65 dB SPL) from a distance of 1 m. Both better-ear aided and unaided SII were calculated; CHH who did not use HAs only had unaided SII included in the analysis.

Hearing Aid Use

During each test visit, the caregiver completed a questionnaire related to daily HA use (available at: https://ochlstudy.org/assessment-tools). Caregivers reported average number of hours that the child wore HAs during the week and weekends, which was calculated as a weighted HA use measure [weekday use × 0.71 (5/7 days of the week) + weekend use × 0.29 (2/7 days of the week)].

Hearing Aid Dosage

To measure the combined effects of HA use and audibility levels (aided and unaided), we calculated a variable termed "HA dosage." This metric can be conceptualized as how much daily access a child receives from HAs. HA dosage combines the number of hours of daily HA use with aided and unaided hearing into one weighted measure of how much auditory access a child experiences during the day[1] from their HAs. It is calculated as HA Dosage = Daily HA Use hours$^{\text{Aided Better-ear SII}}$ − (24 − Daily HA Use hours)$^{\text{Unaided Better-ear SII}}$. The number of hours of daily HA use is weighted by aided SII (access to speech with HAs). If SII = 1, the child has full access to the speech spectrum for that number of hours throughout the day. The amount of time the child does not wear HAs during the day, weighted by unaided SII (access to speech without HAs), is then subtracted from the hours of use weighted by aided SII. A smaller value indicates lower HA dosage and a greater value indicates higher HA dosage.

Speech Recognition in Noise

We administered the BKB-SIN test (Bench et al., 1979) at each test visit. The BKB-SIN was developed to be used with children and includes short sentences with semantic and syntactic content at a first-grade reading level (Wilson et al., 2007). Recorded sentences were presented with a male talker in multi-talker background noise. The signal was calibrated at 65 dBA prior to administration. Each child received one list consisting of Part A and Part B (10 sentences per part) per visit. Lists 1–8 were administered randomly to participants; however, no participants received the same list 2 years in a row. Each

[1]The 24°h would include periods of time when the child is sleeping and presumably receiving little to no input. We did not ask parents how many hours their children slept on average. In the absence of those data, it seemed appropriate to calculate a full 24°h of possible input, rather than try to estimate individual sleep patterns for children.

sentence was presented at a different SNR, starting at 21 dB SNR and decreasing in 3 dB decrements. The tenth sentence was presented at −6 dB SNR. The test was scored in terms of the SNR needed to accurately identify 50% of the key words (i.e., SNR-50) rather than percent-correct of the total word list. Thus, a lower SNR-50 represents less difficulty understanding speech in background noise, and growth over time is seen as a downward trajectory.

Language Measures

Test protocols were developed to be appropriate for children utilizing spoken English in first through fourth grade. Test protocols varied depending on the year of testing. First and third grade test batteries were the same, and second and fourth grade test batteries were the same.

Vocabulary

At first and third grade, we administered two measures of vocabulary knowledge. The Wechsler Abbreviated Scale of Intelligence-2 (WASI-2; Wechsler and Hsiao-pin, 2011) Vocabulary subtest is a standardized measure of expressive vocabulary. The examiner instructs the participant to define a series of words. Responses are scored as 0, 1, or 2 points based on the accuracy of the definition. Also at first and third grade, examiners administered the Peabody Picture Vocabulary Test-4 (PPVT-4; Dunn and Dunn, 2007). The PPVT-4 assesses receptive vocabulary; the examiner says a target word that corresponds to one of four pictures in a set, and the participant indicates the correct word. The correlation between the WASI-2 Vocabulary raw scores and the PPVT-4 raw scores was 0.81. Because the raw scores for WASI-2 Vocabulary and PPVT-4 are on different scales, we transformed each participant's score to z-scores and averaged the z-scores together to create a single vocabulary composite score. The conversion to z-scores allowed us to standardize performance relative to our own population of participants and better measure individual growth. At second and fourth grade, we administered the Woodcock-Johnson Tests of Achievement-III Picture Vocabulary subtest (WJTA-III; Woodcock et al., 2001), which measures expressive vocabulary via picture naming. Again, we transformed the raw scores to z-scores so they would be on the same scale as the other vocabulary measures.

Working Memory

At first and third grade, we administered two standardized working memory measures from the Automated Working Memory Assessment (AWMA; Alloway, 2007). The Odd One Out subtest is a visual–spatial complex working memory span task. The participant sees three shapes in a three-square matrix on a computer screen. Two of the shapes are the same and one is different. The participant points to the shape that is the "odd one out." The participant is then shown three empty boxes and indicates where the odd shape was located. The task is administered using a span procedure, in which the participant is asked to indicate the location of an increasing number of items. When four out of six spans within a set

are identified correctly, the participant moves to the next level and the span increases by one item. The task is discontinued after three incorrect span responses within a set.

The Listening Recall subtest is a verbal complex working memory span task. The participant hears a sentence (e.g., "You eat soup with a knife") and must determine if it is true or false. After hearing a set of two sentences, the participant repeats back the last word of each sentence in the order that he/she heard them. This task is also administered using a span procedure; if the participant accurately identifies the last words of the sentences in the correct order, the span increases by one sentence. The correlation for Listening Recall and Odd One Out raw scores for participants in first and third grade was 0.61. Raw scores were transformed into z-scores and averaged together to form a composite score.

At second and fourth grade, we administered the Listening Recall and Odd One Out tasks. In addition, we administered Backward Digit span, a working memory span measure in which the participant hears a series of numbers and is instructed to verbally repeat them back in reverse order. The correlation between raw scores for Backward Digit Span and Listening Recall was 0.50, the correlation for Backward Digit Span and Odd One Out was 0.52, and the correlation for Listening Recall and Odd One Out was 0.52 for participants in second and fourth grade. The raw scores of the three variables were transformed into z-scores and averaged together to compute a composite working memory score at second and fourth grade.

Statistical Analyses

Our first two research questions evaluated the growth trajectories of the BKB-SIN SNR-50 scores for CHH and CNH, and whether the two groups showed similarities or differences in their rate of growth. To address these research questions, we constructed a longitudinal regression model. The fixed effects in the regression model were grade (first, second, third, and fourth); hearing status (CHH, CNH); and an interaction between grade and hearing status. To account for the correlation due to repeated measures, we included a correlation structure on the residuals. The Akaike Information Criterion (AIC; Akaike, 1974) was used to select the appropriate correlation structure within the statistical model with lower AIC values meaning better fitting models. A heterogeneous compound symmetric covariance matrix (AIC = 3315.2) was chosen over an unstructured covariance matrix (AIC = 3318.6). Therefore, the correlations between grades were approximately equal, but the variances at each time point were different.

The third research question examined which factors were associated with individual differences in growth rate for speech recognition in noise for only CHH. To construct this analysis, we used a linear regression model with a heterogeneous compound symmetric error structure to account for correlation between grades and unequal variances between grades. The dependent variable was again growth rate on BKB-SIN SNR-50 scores. The fixed effects were grade, maternal education level, age at confirmation of hearing loss, vocabulary composite z-scores, and working memory composite z-scores. Maternal

education level was coded as ordinal levels (1 = High School or less, 2 = Some college, 3 = Bachelor's degree, 4 = Post graduate, with 4 as the reference level). We also included average HA dosage across visits and change in HA dosage as separate fixed effects because HA dosage is a time-varying covariate. Change in HA dosage is calculated as each participant's HA dosage at a given visit subtracted from the average HA dosage across visits. These separate variables allowed us to determine whether the average levels of HA dosage across visits or change in HA dosage were associated with growth rate.

RESULTS

Changes in Speech Recognition in Noise Over Time

We found a significant main effect for grade, $F(3, 389) = 23.78$, $p < 0.0001$. Each older grade had a lower SNR-50 compared to younger grades (see **Table 2**). There was also a significant main effect for hearing status, $t(284) = 8.19$, $p < 0.001$. The interaction between grade and hearing status was not statistically significant, $F(3, 389) = 1.18$, $p = 0.3154$. This lack of an interaction is evident in **Figure 1**. On average, CHH demonstrated a SNR-50 that was 3.14 dB SNR higher than CNH, and the growth rate was consistent between groups.

Factors Associated With Growth Rate in Speech Recognition in Noise for Children Who Are Hard of Hearing

As described in the "Statistical Analyses" section, the fixed factors were grade, maternal education level, age at confirmation of hearing loss, vocabulary composite z-score, working memory composite z-score, average HA dosage, and change in HA dosage. Interactions were not significant, so they were not included in the final model. **Table 3** shows the parameters of the linear regression models. Grade level [$F(3, 199) = 6.04$, $p = 0.0006$], vocabulary composite z-scores [$F(1, 171) = 6.00$, $p = 0.0153$], and average HA dosage [$F(1, 171) = 12.19$, $p = 0.0006$] were significantly associated with rates of growth in BKB-SIN SNR-50 scores. Maternal education level [$F(3, 171) = 0.77$, $p = 0.5098$], age at confirmation of hearing loss [$F(1, 171) = 0.04$, $p = 0.8343$], working memory composite z-scores [$F(1, 171) = 2.42$, $p = 0.1219$], and change in HA dosage [$F(1, 199) = 17$, $p = 0.6802$] were not significant

TABLE 2 | Summary statistics for Bamford-Kowal-Bench SNR-50 scores at each grade level.

Grade	N	Mean	SD	Min	Max
First	118	3.64	3.64	−2.5	18.5
Second	220	2.45	3.56	−7.0	14.5
Third	149	1.87	3.29	−5.5	11.5
Fourth	194	1.36	2.97	−7.0	13.0

SNR, signal-to-noise ratio.
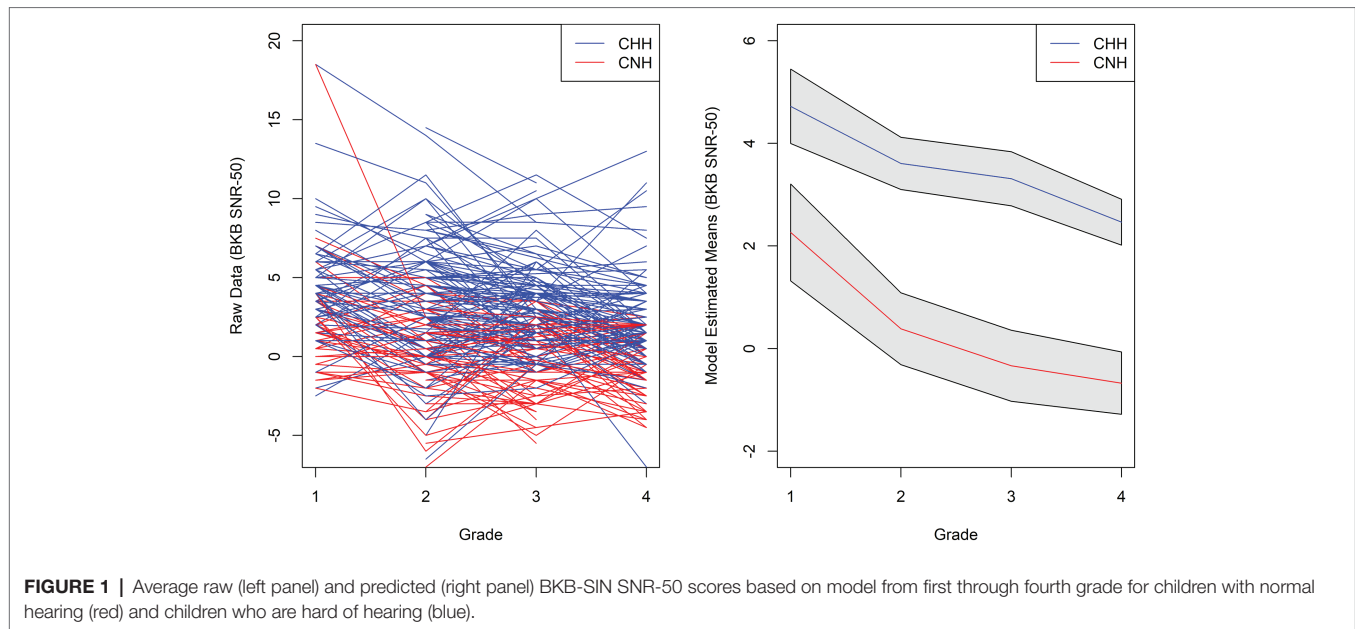
FIGURE 1 | Average raw (left panel) and predicted (right panel) BKB-SIN SNR-50 scores based on model from first through fourth grade for children with normal hearing (red) and children who are hard of hearing (blue).

TABLE 3 | Linear regression model with grade, maternal education level, age at confirmation of hearing loss, average vocabulary, average working memory, average HA dosage, and change in HA dosage as fixed effects and BKB-SIN SNR-50 as the dependent variable.

Parameter	Estimate	Standard error	t (df)	p
Grade (fourth grade = ref. level)				
First	1.6494	0.4192	3.93 (199)	0.0001*
Second	0.8767	0.3356	2.61 (199)	0.0097*
Third	0.3706	0.3384	1.10 (199)	0.2748
Maternal education level (post graduate = ref. level)				
High school or less	0.6879	0.6118	1.12 (171)	0.2625
Some college	−0.2326	0.5260	−0.44 (171)	0.6589
College degree	0.1497	0.4971	0.30 (171)	0.7636
Age at confirmation of hearing loss	−0.0019	0.0093	−0.21 (171)	0.8343
Average vocabulary	−0.6915	0.2822	−2.45 (171)	0.0153*
Average working memory	−0.4795	0.3085	−1.55 (171)	0.1219
HA dosage	−0.1761	0.0504	−3.49 (171)	0.0006*
Change in HA dosage	0.0373	0.0903	0.41 (199)	0.6802

*BKB-SIN, Bamford Kowal Bench Speech in Noise; SNR, signal-to-noise ratio; df, degrees of freedom. *p < 0.05.*

predictors. Stronger vocabulary skills (**Figure 2**), and greater average HA dosage (**Figure 3**) were related to better recognition of speech in noise and these patterns were consistent across age.

DISCUSSION

The primary aim of the current study was to compare speech recognition in noise in a large group of CHH and age-matched hearing peers who have been followed on an annual basis out to fourth grade. To our knowledge, this study is among the first to track the same group of children over time and compare developmental growth rates in speech recognition for CHH compared to CNH. We also evaluated the effects of auditory access, complex working memory span, and vocabulary size on listening in noise in CHH. Identifying the mechanisms that underlie speech recognition in degraded contexts will guide clinical decision-making process for optimizing outcomes (Ching et al., 2018) and inform theories about how auditory access shapes development for CHH (Moeller and Tomblin, 2015).

Group Differences in Growth Trajectories

Prior work on speech recognition in CHH have used cross-sectional designs with a focus on children in the 5- to 12-year-old age range (McCreery et al., 2015; Klein et al., 2017; Ching et al., 2018). CNH appear to improve in their ability to recognize words with age, reaching adult-like levels by adolescence (Eisenberg et al., 2000; Corbin et al., 2016). Based on these previous studies, we expected that CHH would have more difficulty with listening in background noise at the initial test visits and both groups would improve over time, but we were unsure of the between-group developmental patterns of these deficits. There were three possible options: (1) CHH would eventually catch up as a group to the CNH, (2) the gap in speech recognition in noise skills would widen over time, or (3) the gap would remain constant over time. Based on prior literature with adults, it seemed unlikely that CHH would catch up to their hearing peers. Results by Walker et al. (2019) provided some support for the possibility of an increasing gap in speech recognition; however, the results from the linear regression models pointed toward the third option: CHH showed a significant delay in speech recognition in noise skills at the initial visit around first grade, both groups improved in their speech recognition in noise skills over time, and the size of this gap remained approximately the same from first through fourth grade. In effect, both groups appeared to be progressing similarly with time, but the children with hearing loss started off delayed and stayed delayed. These data inform our knowledge

FIGURE 2 | Average predicted BKB-SIN SNR-50 scores as a function of vocabulary scores for children who are hard of hearing. The solid line represents z-scores of −1 (1 SD below average), the dashed line represents z-scores of 0 (average), and the dotted line represents z-scores of +1 (1 SD above the mean).

FIGURE 3 | Average predicted BKB-SIN SNR-50 scores as a function of HA dosage for children who are hard of hearing. The solid line represents the 5th percentile (HA dosage = 5), the dashed line represents the 50th percentile (HA dosage = 10), and the dotted line represents the 95th percentile (HA dosage = 15).

about long-term trajectories in speech recognition in noise for children, as we do not see evidence of convergence or divergence between groups. These data also have important clinical implications because they highlight the need to continue providing support for children with all degrees of hearing loss in the general education setting as they transition from elementary grades into secondary grades. This support may take the form of resource support with a speech-language pathologist or teacher of the deaf/hard of hearing, classroom audio distribution systems, personal remote microphone systems, and/or preferential seating in the classroom.

Individual Differences in Growth Trajectories for Children Who Are Hard of Hearing

Our second aim was to examine the factors that support growth for speech recognition in noise for CHH. Previous studies have examined age at service delivery (Sininger et al., 2010), aided audibility (Davidson and Skinner, 2006; Scollie, 2008; Stiles et al., 2012), and language (Blamey et al., 2001; Nittrouer et al., 2013) as predictive factors, but only a few have looked at the combination of auditory access, cognition, and language (McCreery et al., 2015; Klein et al., 2017; Ching et al., 2018). The findings from these previous studies have been mixed. Ching et al. found that non-verbal IQ and global language skills predicted speech recognition in noise skills for CHH, but auditory access (measured with aided SII, after controlling for unaided hearing levels) did not contribute significant variance. Klein et al. found an effect of vocabulary size, but not working memory (measured with a phonological short-term memory task) or auditory access (measured with aided SII and HA use as separate variables). McCreery et al. (2015) showed significant associations between all three factors (vocabulary

size, aided SII, and phonological working memory) and word recognition in noise.

Taken together, the results of the current study may be viewed as partial support of the predictions of the ELU model. Children with stronger language skills were better able to recognize degraded speech, and children with poorer language skills had more difficulty with speech recognition in noise. Our longitudinal results indicate not just that better vocabulary skills support the ability to perceive a degraded message, but the effect of vocabulary size is stable across time. As discussed in Ching et al. (2018), these findings point toward the critical importance of language development as a focus of intervention for children with hearing loss. For some CHH who demonstrate extreme difficulty with listening in noise, this intervention may need to continue into the school age years, a time period when the intervention needs of CHH are sometimes overlooked (Antia et al., 2009). We also acknowledge that reduced auditory access in early childhood may lead to poorer speech recognition in noise skills, which in turn makes the word learning process more difficult for children with hearing loss (Walker and McGregor, 2013; Blaiser et al., 2015). We are unable to determine the direction of the relationship between vocabulary size and speech recognition noise with our current analysis approach, but future studies could employ cross-lagged analysis models or mediation analysis to infer directionality.

In contrast to the effect of vocabulary, we did not find an impact of working memory on speech recognition. The lack of an association is consistent with Magimairaj et al. (2018), and inconsistent with McCreery et al. (2015, 2017). Magimairaj and colleagues used the same clinical outcome measure, BKB-SIN, as the current study. McCreery et al. (2017) used sentences

that were either syntactically correct but had no semantic meaning or had no syntactic structure or semantic meaning. Thus, the stimuli in McCreery et al. may have required children to rely on memory skills to recall the words, because they could not use linguistic bootstrapping. The BKB-SIN sentences had less of a memory load because children could use linguistic skills to remember the sentence, leading to reduced need to use working memory to repeat target words even in high levels of noise. Another possibility is that the shared variance in the vocabulary and working memory composite measures may have resulted in only vocabulary accounting for unique variance in speech recognition in noise. A larger sample size might have been able to demonstrate unique effects of both variables.

If future studies continue to support a stronger effect of language skills compared to working memory on speech recognition in children, these findings may point toward a need to modify the predictions of the ELU model. The ELU model emphasizes working memory skills as a compensatory mechanism in complex listening situations, with less focus on language skills. Because children show more variability in vocabulary breadth and depth than adults, language ability may take on a more important role in understanding distorted or masked speech, relative to working memory. Additional research is needed to test the applicability of the ELU model to the pediatric population.

In addition to cognitive and linguistic measures, we looked at how auditory access impacts individual differences in speech recognition in noise for CHH. The effects of auditory access have been inconsistent across studies (Blamey et al., 2001; Sininger et al., 2010; McCreery et al., 2015; Klein et al., 2017; Ching et al., 2018). Part of this inconsistency is due to different approaches in quantifying how much access CHH have to speech. Our measure of auditory access represents a novel approach to quantifying the HA experience of CHH. Here we developed a metric, HA dosage, that considers specific effects of amplification by weighting the amount of time children wore amplification throughout the day with aided and unaided hearing levels. The measurement of HA dosage is an improvement on previous attempts to look at auditory access in CHH because it combines sources of variability related to amplification (aided SII and HA use). It also accounts for the differential impact of HA use time based on unaided SII. When we averaged HA dosage across visits for participants, it was a significant predictor of growth rates. Like vocabulary knowledge, as HA dosage increases, CHH show better speech recognition in noise, but the patterns of change do not vary in relation to levels of HA dosage. These results highlight the need for interventions that include well-fitted HAs and consistent HA use, even in cases of mild or moderate hearing loss. While CHH with more residual hearing may perform well in quiet with or without amplification, most listening and learning situations occur in suboptimal or adverse conditions (Shield and Dockrell, 2008; Mattys et al., 2012; Ambrose et al., 2014). Increased HA dosage appears to offer some protection against the difficulties of listening in noise for these children.

We also examined whether change in HA dosage over time influenced growth rates and did not find a significant effect.

CHH show variation in the consistency of auditory access during childhood (McCreery et al., 2013; Walker et al., 2013). By the school-age years, these fluctuations in auditory access do not appear to have an impact on longitudinal growth trajectories in speech recognition in noise. In addition to a lack of a significant effect for change in HA dosage, we did not find an association between speech recognition in noise with maternal education level or age at confirmation of hearing loss. Both variables have been shown to have a positive effect on auditory outcomes in children with hearing loss in previous studies (Sininger et al., 2010; McCreery et al., 2015), but the children in these earlier studies were younger than the children in the current study. Other studies with this same cohort of children indicate that CHH who receive audiologic services later demonstrate initial delays in language outcomes, but show a pattern of catching up to CHH who received services earlier by age 6 years (Tomblin et al., 2015). Thus, timing of service provision may initially affect language and listening outcomes, but the impact of age at confirmation (which is highly correlated with age at HA fitting) gradually weakens over time as other factors (vocabulary skills, aided audibility, HA use) support speech recognition in noise and ameliorate the negative effects of later confirmation of hearing loss and lower maternal education levels.

Limitations
A strength of this study is that it is the first to document longitudinal change in growth trajectories for CNH and CHH on measures of speech recognition in noise. There are also several limitations that should be discussed. Due the study design, children were tested at different time points rather than all children participating at the same time points. This issue of inconsistent time points is a common obstacle in longitudinal research studies, as participants often start late, drop out, or skip test visits (Krueger and Tian, 2004). The use of linear mixed models for the statistical analysis accommodates data where individuals are measured at different time points (Oleson et al., 2019; Walker et al., 2019). The linear mixed model creates individual-specific trends through weighted averages of the individual observed data and the population average data so that all scores can be used in the analysis even if they are at differing time points.

Another limitation is that testing took place over a fairly limited time span (up to four visits). Further, we only tested participants up to 11 years of age, which is still a period of early adolescence. While the current data trends suggest that CNH and CHH show parallel rates of development in speech recognition in noise, it is possible that we may see differences in growth trajectories past 11 years (Corbin et al., 2016), particularly if CNH reach adult-like performance but CHH continue to improve. Future studies would need to include longitudinal data at older ages to determine if CHH eventually catch up to their hearing peers or if deficits persist with age.

We also note that the inclusionary and exclusionary criteria for this study resulted in a homogeneous cohort of children from English-speaking backgrounds with no additional motor or cognitive deficits. Thus, the current results may not generalize

to linguistically diverse populations or children with hearing loss who have additional disabilities. We excluded children with profound hearing loss because we were interested in the impact of hearing loss in the mild to severe range. It is possible that we would have seen a stronger impact of age at confirmation of hearing loss if children who are deaf had been included in the sample. We did not control for the type of hearing loss because our goal was to recruit as many children with permanent hearing loss as possible; however, the majority of children presented with sensorineural hearing loss. We acknowledge that the consequences of sensorineural and conductive hearing loss can impact speech recognition in noise differently, but our limited number of children with conductive hearing loss prevents us from analyzing these children as a separate group.

A final limitation is that we restricted our speech in noise measure to the BKB-SIN test, which uses a four-talker babble as the competing signal. Other studies have shown that informational masking is increased as the number of competing talkers is decreased (Freyman et al., 2004), CNH demonstrate different developmental trajectories for two-talker maskers compared to more energetic masking signals (Corbin et al., 2016), and CHH have more difficulty with two-talker maskers than CNH (Leibold et al., 2013). We did not evaluate the effects of age, hearing status, and masker type in the present study, but this would be an important future direction in order to fully understand children's susceptibility to background noise.

Conclusions

The current study established longitudinal growth trajectories of speech recognition in noise for school-age CHH and CNH. As a group, CHH demonstrated deficits in speech recognition in noise. These deficits do not appear to converge toward or diverge from CNH, as the growth rates were parallel for the CHH and CNH. These findings also helped us identify the underlying mechanisms that drive growth in speech recognition, with stronger vocabulary and higher HA dosage supporting speech recognition in degraded situations.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

EW and RM conceived of the presented idea. EW took the lead in writing the manuscript. CS and RM contributed to the interpretation of the result and the final version of the manuscript. JO performed the analytic and statistical calculations and data visualization. All authors provided critical feedback and helped shape the research, analysis, and manuscript.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19, 716–723. doi: 10.1109/TAC.1974.1100705

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47(Suppl. 2), S53–S71. doi: 10.1080/14992020802301142

Alloway, T. (2007). *The automatic working memory assessment (awma)[computer software]*. London, UK: Harcourt Assessment.

Ambrose, S. E., VanDam, M., and Moeller, M. P. (2014). Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear Hear.* 35, 139–147. doi: 10.1097/AUD.0b013e3182a76768

ANSI (1997). *Methods for calculation of the speech intelligibility index. Technical report S3. 5-1997*. New York, NY: American National Standards Institute.

Antia, S. D., Jones, P. B., Reed, S., and Kreimeyer, K. H. (2009). Academic status and progress of deaf and hard-of-hearing students in general education classrooms. *J. Deaf. Stud. Deaf. Educ.* 14, 293–311. doi: 10.1093/deafed/enp009

Bench, J., Kowal, Å., and Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br. J. Audiol.* 13, 108–112.

Blaiser, K. M., Nelson, P. B., and Kohnert, K. (2015). Effect of repeated exposures on word learning in quiet and noise. *Commun. Disord. Q.* 37, 25–35. doi: 10.1177/1525740114554483

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., et al. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *J. Speech Lang. Hear. Res.* 44, 264–285. doi: 10.1044/1092-4388(2001/022)

Caldwell, A., and Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *J. Speech Lang. Hear. Res.* 56, 13–30. doi: 10.1044/1092-4388(2012/11-0338)

Ching, T. Y., Day, J., Van Buynder, P., Hou, S., Zhang, V., Seeto, M., et al. (2014). Language and speech perception of young children with bimodal fitting or bilateral cochlear implants. *Cochlear Implants Int.* 15(Suppl. 1), S43–S46. doi: 10.1179/1467010014Z.000000000168

Ching, T. Y., Dillon, H., Marnane, V., Hou, S., Day, J., Seeto, M., et al. (2013). Outcomes of early-and late-identified children at 3 years of age: findings from a prospective population-based study. *Ear Hear.* 34, 535–552. doi: 10.1097/AUD.0b013e3182857718

Ching, T. Y., Zhang, V. W., Flynn, C., Burns, L., Button, L., Hou, S., et al. (2018). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57(Suppl. 2), S70–S80. doi: 10.1080/14992027.2017.1346307

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (2016). Development of open-set word recognition in children: speech-shaped noise and two-talker speech maskers. *Ear Hear.* 37, 55–63. doi: 10.1097/AUD.0000000000000201

Crandell, C. C. (1993). Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear Hear.* 14, 210–216. doi: 10.1097/00003446-199306000-00008

Davidson, L. S., Geers, A. E., Blamey, P. J., Tobey, E., and Brenner, C. (2011). Factors contributing to speech perception scores in long-term pediatric CI users. *Ear Hear.* 32, 19S–26S. doi: 10.1097/AUD.0b013e3181ffdb8b

Davidson, L. S., and Skinner, M. W. (2006). Audibility and speech perception of children using wide dynamic range compression hearing aids. *Am. J. Audiol.* 15, 141–153. doi: 10.1044/1059-0889(2006/018)

Dubno, J. R. (2015). Speech recognition across the life span: longitudinal changes from middle-age to older adults. *Am. J. Audiol.* 24, 84–87. doi: 10.1044/2015_AJA-14-0052

Dunn, D., and Dunn, L. (2007). *Peabody picture vocabulary test.* 4th edn. Minneapolis, MN: NCS Pearson, Inc.

Dunn, C. C., Walker, E. A., Oleson, J., Kenworthy, M., Van Voorst, T., Tomblin, J. B., et al. (2014). Longitudinal speech perception and language performance in pediatric cochlear implant users: the effect of age at implantation. *Ear Hear.* 35, 148–160. doi: 10.1097/AUD.0b013e3182a4a8f0

Easwar, V., Sanfilippo, J., Papsin, B., and Gordon, K. J. (2018). Impact of consistency in daily device use on speech perception abilities in children with cochlear implants: datalogging evidence. *J. Am. Acad. Audiol.* 29, 835–846. doi: 10.3766/jaaa.17051

Eisenberg, L. S., Shannon, R. V., Schaefer Martinez, A., Wygonski, J., and Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *J. Acoust. Soc. Am.* 107, 2704–2710. doi: 10.1121/1.428656

Freyman, R. L., Balakrishnan, U., and Helfer, K. S. (2004). Effect of number of masking talkers and auditory priming on informational masking in speech recognition. *J. Acoust. Soc. Am.* 115, 2246–2256. doi: 10.1121/1.1689343

Gustafson, S. J., Davis, H., Hornsby, B. W., and Bess, F. H. (2015). Factors influencing hearing aid use in the classroom: a pilot study. *Am. J. Audiol.* 24, 563–568. doi: 10.1044/2015_AJA-15-0024

Hillock-Dunn, A., Taylor, C., Buss, E., and Leibold, L. J. (2014). Assessing speech perception in children with hearing loss: what conventional clinical tools may miss. *Ear Hear.* 36, 57–60. doi: 10.1097/AUD.0000000000000007

Holte, L., Walker, E., Oleson, J., Spratford, M., Moeller, M. P., Roush, P., et al. (2012). Factors influencing follow-up to newborn hearing screening for infants who are hard of hearing. *Am. J. Audiol.* 21, 163–174. doi: 10.1044/1059-0889(2012/12-0016)

Kirk, K. I., Prusick, L., French, B., Gotch, C., Eisenberg, L. S., and Young, N. (2012). Assessing spoken word recognition in children who are deaf or hard of hearing: a translational approach. *J. Am. Acad. Audiol.* 23, 464–475. doi: 10.3766/jaaa.23.6.8

Klein, K. E., Spratford, M., Redfern, A., and Walker, E. A. (2019). Effects of grade and school services on children's responsibility for hearing aid care. *Am. J. Audiol.* 28, 673–685. doi: 10.1044/2019_AJA-19-0005

Klein, K. E., Walker, E. A., Kirby, B., and McCreery, R. W. (2017). Vocabulary facilitates speech perception in children with hearing aids. *J. Speech Lang. Hear. Res.* 60, 2281–2296. doi: 10.1044/2017_JSLHR-H-16-0086

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., and Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Dev. Sci.* 18, 146–154. doi: 10.1111/desc.12202

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., and Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms. *Am. J. Audiol.* 11, 65–71. doi: 10.1044/1059-0889(2002/009)

Krueger, C., and Tian, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biol. Res. Nurs.* 6, 151–157. doi: 10.1177/1099800404267682

Lalonde, K., and Holt, R. F. (2014). Cognitive and linguistic sources of variance in 2-year-olds' speech-sound discrimination: a preliminary investigation. *J. Speech Lang. Hear. Res.* 57, 308–326. doi: 10.1044/1092-4388(2013/12-0227)

Leibold, L. J., Hillock-Dunn, A., Duncan, N., Roush, P. A., and Buss, E. (2013). Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear Hear.* 34, 575–584. doi: 10.1097/AUD.0b013e3182857742

Magimairaj, B. M., Nagaraj, N. K., and Benafield, N. J. (2018). Children's speech perception in noise: evidence for dissociation from language and working memory. *J. Speech Lang. Hear. Res.* 61, 1294–1305. doi: 10.1044/2018_JSLHR-H-17-0312

Mattys, S. L., Davis, M. H., Bradlow, A. R., and Scott, S. K. (2012). Speech recognition in adverse conditions: a review. *Lang. Cogn. Process.* 27, 953–978. doi: 10.1080/01690965.2012.705006

McCreery, R. W., Bentler, R. A., and Roush, P. A. (2013). Characteristics of hearing aid fittings in infants and young children. *Ear Hear.* 34, 701–710. doi: 10.1097/AUD.0b013e31828f1033

McCreery, R. W., Spratford, M., Kirby, B., and Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *Int. J. Audiol.* 56, 306–315. doi: 10.1080/14992027.2016.1266703

McCreery, R. W., Walker, E. A., Spratford, M., Oleson, J., Bentler, R., Holte, L., et al. (2015). Speech recognition and parent ratings from auditory development questionnaires in children who are hard of hearing. *Ear Hear.* 36, 60S–75S. doi: 10.1097/AUD.0000000000000213

Mehra, S., Eavey, R. D., and Keamy, D. G. Jr. (2009). The epidemiology of hearing impairment in the United States: newborns, children, and adolescents. *Otolaryngol. Head Neck Surg.* 140, 461–472. doi: 10.1016/j.otohns.2008.12.022

Moeller, M. P., and Tomblin, J. B. (2015). An introduction to the outcomes of children with hearing loss study. *Ear Hear.* 36, 4S–13S. doi: 10.1097/AUD.0000000000000210

Nittrouer, S., and Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *J. Acoust. Soc. Am.* 87, 2705–2715. doi: 10.1121/1.399061

Nittrouer, S., Caldwell-Tarr, A., Tarr, E., Lowenstein, J. H., Rice, C., and Moberly, A. C. (2013). Improving speech-in-noise recognition for children with hearing loss: potential effects of language abilities, binaural summation, and head shadow. *Int. J. Audiol.* 52, 513–525. doi: 10.3109/14992027.2013.792957

Oleson, J. J., Brown, G. D., and McCreery, R. (2019). The evolution of statistical methods in speech, language, and hearing sciences. *J. Speech Lang. Hear. Res.* 62, 498–506. doi: 10.1044/2018_JSLHR-H-ASTM-18-0378

Page, T. A., Harrison, M., Moeller, M. P., Oleson, J., Arenas, R. M., and Spratford, M. (2018). Service provision for children who are hard of hearing at preschool and elementary school ages. *Lang. Speech Hear. Serv. Sch.* 49, 965–981. doi: 10.1044/2018_LSHSS-17-0145

Robinson, E. J., Davidson, L. S., Uchanski, R. M., Brenner, C. M., and Geers, A. E. (2012). A longitudinal study of speech perception skills and device characteristics of adolescent cochlear implant users. *J. Am. Acad. Audiol.* 23, 341–349. doi: 10.3766/jaaa.23.5.5

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., et al. (2013). The ease of language understanding (ELU) model: theoretical, empirical, and clinical advances. *Front. Syst. Neurosci.* 7, 7–24. doi: 10.3389/fnsys.2013.00031

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S99–S105. doi: 10.1080/14992020802301167

Scollie, S. D. J. (2008). Children's speech recognition scores: the speech intelligibility index and proficiency factors for age and hearing level. *Ear Hear.* 29, 543–556. doi: 10.1097/AUD.0b013e3181734a02

Shield, B. M., and Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Sininger, Y. S., Grimes, A., and Christensen, E. (2010). Auditory development in early amplified children: factors influencing auditory-based communication outcomes in children with hearing loss. *Ear Hear.* 31, 166–185. doi: 10.1097/AUD.0b013e3181c8e7b6

Stelmachowicz, P. G., Hoover, B. M., Lewis, D. E., Kortekaas, R. W., and Pittman, A. L. (2000). The relation between stimulus context, speech audibility, and perception for normal-hearing and hearing-impaired children. *J. Speech Lang. Hear. Res.* 43, 902–914. doi: 10.1044/jslhr.4304.902

Stiles, D. J., Bentler, R. A., and McGregor, K. K. (2012). The speech intelligibility index and the pure-tone average as predictors of lexical ability in children fit with hearing aids. *J. Speech Lang. Hear. Res.* 55, 764–778. doi: 10.1044/1092-4388(2011/10-0264)

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., and Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss. *Ear Hear.* 36, 76S–91S. doi: 10.1097/AUD.0000000000000219

Tun, P. A., Benichov, J., and Wingfield, A. (2010). Response latencies in auditory sentence comprehension: effects of linguistic versus perceptual challenge. *Psychol. Aging* 25, 730–735. doi: 10.1037/a0019300

Uhler, K., Yoshinaga-Itano, C., Gabbard, S. A., Rothpletz, A. M., and Jenkins, H. (2011). Longitudinal infant speech perception in young cochlear implant users. *J. Am. Acad. Audiol.* 22, 129–142. doi: 10.3766/jaaa.22.3.2

Walker, E. A., Kessler, D., Klein, K., Spratford, M., Oleson, J. J., Welhaven, A., et al. (2019). Time-gated word recognition in children: effects of auditory access, age, and semantic context. *J. Speech Lang. Hear. Res.* 62, 2519–2534. doi: 10.1044/2019_JSLHR-H-18-0407

Walker, E. A., and McGregor, K. K. (2013). Word learning processes in children with cochlear implants. *J. Speech Lang. Hear. Res.* 58, 1611–1625. doi: 10.1044/1092-4388(2012/11-0343)

Walker, E. A., Redfern, A., and Oleson, J. J. (2019). Linear mixed-model analysis to examine longitudinal trajectories in vocabulary depth and breadth in children who are hard of hearing. *J. Speech Lang. Hear. Res.* 62, 525–542. doi: 10.1044/2018_JSLHR-L-ASTM-18-0250

Walker, E. A., Sapp, C., Dallapiazza, M., Spratford, M., McCreery, R., and Oleson, J. J. (accepted). Language and reading outcomes in fourth-grade children with mild hearing loss compared to age-matched hearing peers. *Lang. Speech Hear. Serv. Sch.*

Walker, E. A., Spratford, M., Moeller, M. P., Oleson, J., Ou, H., Roush, P., et al. (2013). Predictors of hearing aid use time in children with mild-to-severe hearing loss. *Lang. Speech Hear. Serv. Sch.* 44, 73–88. doi: 10.1044/0161-1461(2012/12-0005)

Warren, S. F., Fey, M. E., and Yoder, P. J. (2007). Differential treatment intensity research. A missing link to creating optimally effective communication interventions. *Ment. Retard. Dev. Disabil. Res. Rev.* 13, 70–77. doi: 10.1002/mrdd.20139

Wechsler, D., and Hsiao-pin, C. (2011). *Wechsler abbreviated scale of intelligence.* San Antonio, TX: Pearson.

Wilson, R. H., McArdle, R. A., and Smith, S. L. (2007). An evaluation of the BKB-SIN, HINT, QUICKSIN, and WIN materials on listeners with normal hearing and listeners with hearing loss. *J. Speech Lang. Hear. Res.* 50, 844–856. doi: 10.1044/1092-4388(2007/059)

Woodcock, R. W., McGrew, K. S., Mather, N., and Schrank, F. (2001). *Woodcock-Johnson III NU tests of achievement.* Rolling Meadows, IL: Riverside Publishing.

Zekveld, A. A., Rudner, M., Johnsrude, I. S., Festen, J. M., Van Beek, J. H., and Rönnberg, J. (2011). The influence of semantically related and unrelated text cues on the intelligibility of sentences in noise. *Ear Hear.* 32, e16–e25. doi: 10.1097/AUD.0b013e318228036a

Spelling in Deaf, Hard of Hearing and Hearing Children With Sign Language Knowledge

Moa Gärdenfors[1]*, Victoria Johansson[2] and Krister Schönström[1]

[1] Department of Linguistics, Faculty of Humanities, Stockholm University, Stockholm, Sweden, [2] Center for Languages and Literature, The Joint Faculties of Humanities and Theology, Lund University, Lund, Sweden

What do spelling errors look like in children with sign language knowledge but with variation in hearing background, and what strategies do these children rely on when they learn how to spell in written language? Earlier research suggests that the spelling of children with hearing loss is different, because of their lack of hearing, which requires them to rely on other strategies. In this study, we examine whether, and how, different variables such as hearing degree, sign language knowledge and bilingualism may affect the spelling strategies of children with Swedish sign language, *Svenskt teckenspråk*, (STS) knowledge, and whether these variables can be mirrored in these children's spelling. The spelling process of nineteen children with STS knowledge (mean age: 10.9) with different hearing degrees, born into deaf families, is described and compared with a group of fourteen hearing children without STS knowledge (mean age: 10.9). Keystroke logging was used to investigate the participants' writing process. The spelling behavior of the children was further analyzed and categorized into different spelling error categories. The results indicate that many children showed exceptionally few spelling errors compared to earlier studies, that may derive from their early exposure of STS, enabling them to use the fingerspelling strategy. All of the children also demonstrated similar typing skills. The deaf children showed a tendency to rely on a visual strategy during spelling, which may result in incorrect, but visually similar, words, i.e., a type of spelling errors not found in texts by hearing children with STS knowledge. The deaf children also showed direct transfer from STS in their spelling. It was found that hard-of-hearing children together with hearing children of deaf adults (CODAs), both with STS knowledge, used a sounding strategy, rather than a visual strategy. Overall, this study suggests that the ability to hear and to use sign language, together and respectively, play a significant role for the spelling patterns and spelling strategies used by the children with and without hearing loss.

Keywords: spelling, sign language, deaf, hard of hearing, CODA, writing processes, keystroke logging, spelling strategies

INTRODUCTION

This article concerns the writing skills of deaf and hard of hearing (henceforth, DHH) children, and focuses on the processes of writing and spelling. Having various degrees of hearing, or different language backgrounds may lead to different opportunities to develop a spelling ability, but the question remains whether and how those variables together, or separately, will mirror the spelling features of the deaf, and hard of hearing (henceforth, HoH) children.

To give an example, it is likely that when young, normally hearing children start learning how to spell, they will begin by basing the spelling on a sounding strategy, which in turn will cause typical misspelled words, with a close mapping of grapheme and phoneme (e.g., Nauclér, 1989). Bilinguals may exhibit cross-linguistic influence patterns in their language production, that is when structures in any language are influenced by their bilingual competence (see Figueredo, 2006, for a review). But we know next to nothing about the spelling patterns of children with hearing loss, who *at the same* time are bimodal bilinguals. A person is bimodal bilingual when their languages operates in two different modalities, for example using a sign language and a spoken language. Would such a context lead us to expect a pattern of transfer from sign language in the children's spelling (i.e., will the children use a set of visual strategies for their spelling)? Is this case comparable to children who are hearing and bimodal bilinguals (i.e., those who use both sound-based and visual-based cues for their spelling)? Earlier research on the writing of DHH has mostly focused on deviations and errors (see Albertini and Schley, 2010 for an overview), and very few writing studies have included a sign language or a bilingual perspective over different language modalities *and* degrees of hearing.

Wengelin (2002) has reported that deaf adults with sign language knowledge misspelled fewer words compared to adults with reading and writing difficulties, and on the word-level this group barely demonstrated any doubling errors, which is an error type that is common in Swedish. Swedish spelling conventions require that many words include doubled consonants [e.g., "komma" ('come')]. To understand when a consonant should be doubled, and when not, requires both phonological and morphological knowledge, and spelling mistakes in this category are very common for all children in the targeted age group. Doubling errors include on the one hand errors when a consonant is erroneously doubled e.g., "villja" instead of "vilja" ('will'), and on the other hand when the second consonant is erroneously missing [e.g., "tuga" instead of "tugga" ('chew')]. The deaf adults in Wengelin's study, by contrast, showed more reversals, insertions, and morphological errors. The same study also showed that the deaf adults had a higher tendency to choose words which are visually similar to the target word, which resulted in a strategy that can be described as 'spell as it looks' – which was compared to a group of adults with reading and writing difficulties who spelled 'as it sounds.' Another finding was that the deaf adults showed a heterogeneous pattern, without common production problems, while the pattern was more homogenous in the adults with reading and writing difficulties. The strategies of the deaf group are thought to most likely be based on visual cues, where some patterns possibly could be derived from Swedish Sign Language (Svenskt Teckenspråk, henceforth, STS). Wengelin stresses that to find out with more certainty, an investigation of possible STS-influence, including what types of strategies or visual cues deaf people use to spell words, is needed.

The purpose of this study is therefore to perform a descriptive analysis of *whether* and *how* the spelling pattern is linked to children's linguistic knowledge, not only of STS, but also of bilingualism, and hearing loss respectively, by looking at the spelling process *and* the final product through comparing children with and without STS knowledge, and with and without hearing loss. To our knowledge, this is the first study of its kind.

BACKGROUND

It is known that many DHH children face considerable challenges when learning to write. One factor behind these challenges is the absence of, or limited, hearing ability. Another factor is linked to the language acquisition background of the child. The literature often refers to the fact that more than 90% of deaf children are born into a hearing family without any contact with sign language (Mitchell and Karchmer, 2004). This may lead to a delayed start of language acquisition, and, in turn, the acquisition of written language can become a real challenge for the DHH children (Hall, 2017; Glickman and Hall, 2018). Nevertheless, studies show that deaf learners can become skilled readers and writers as well (Hoffmeister and Caldwell-Harris, 2014): a correlation between sign language knowledge and written language proficiency has been consistently reported (Strong and Prinz, 1997; Chamberlain and Mayberry, 2008; Freel et al., 2011; Kuntze et al., 2014). Previous research suggests that DHH children who are born into deaf families or, in exceptional cases, into families who started learning sign language early, may face a considerable advantage in their language development (see e.g., Svartholm, 2010 for an overview). Other studies have also shown that children with cochlear implant (henceforth, CI) with sign language knowledge outperform their DHH-peers born into hearing families without sign language knowledge in almost all intelligence tests (Amraei et al., 2017), in their speech and auditory development (Hassanzadeh, 2012), and showed comparable English scores with their hearing peers with sign language knowledge (Davidson et al., 2014), due to early first language acquisition. By contrast, some researchers have analyzed written outcomes for the deaf using the theoretical framework of Second Language Acquisition, arriving at the conclusion that deaf children exhibit grammatical structures similar to those of hearing second language learners in written Swedish (e.g., Svartholm, 2008; Schönström, 2014).

However, due to the variation in the DHH children's different language (and hearing) backgrounds, it is difficult to arrive at general conclusions, as the relation between (or effect of) language experience (spoken/signed) and use, versus language

proficiency and acquisition background (L1 or L2) remains understudied (cf. Hirshorn et al., 2015).

Sounding Strategies in Spelling

A great deal of the research on literacy concerns phonological awareness. The first stage of developing literacy (in hearing children) is the development of phonological awareness that is, the knowledge of sounds, how sounds can be categorized into phonemes, and how sounds build words. An established phonological awareness has been proposed to constitute an essential foundation for reading, writing and spelling development (Nation and Hulme, 1997). According to the Swedish curriculum for the compulsory school, hearing students in grade three are assessed on their understanding of the relation between graphemes and sounds, but also the spelling rules for regular words, their mastering of the structure of Swedish, and their use of capital letters, question marks, exclamation marks and other punctuation. It can thus be expected that the fundaments of spelling are established for Swedish children when they begin 4th grade, which in Sweden means children of around 10 years of age (Skolverket, 2018).

An overreliance on phonological strategies as the foundation for spelling in hearing children can cause more spelling errors, since other factors (e.g., morphology) influence the orthographic rules (Frith, 1985), for instance Swedish orthography emphasizes doubling errors and letter substitutions (with recurrent examples from Bäck, 2011; Gärdenfors, 2016; Raatikainen, 2018).

From this it follows that normally hearing children have an advantage compared to the DHH, since their ability to hear helps in developing spoken language phonological awareness. When it comes to deaf readers and phonological awareness, Mayberry et al. (2011) conducted a meta-analysis of phonological awareness and reading skills, arriving at the conclusion that phonological awareness as a factor for deaf (and hearing) readers' reading skills is overstated. Instead they found that language ability is a stronger predictor of reading achievement among deaf children. It should also be noted that DHH children can develop phonological awareness in sign language. Research has found positive correlations between signed phonological awareness and literacy skills. Profoundly deaf children can decode phonological information in sign language based on global characteristics from written words, fingerspelling or lipreading. Their solutions to code whole-words may therefore result in different misspelled words such as omissions (e.g., writing "orng" for 'orange'), or letter reversals (e.g., writing "sorpt" for 'sport'). Omissions, insertions and consonant errors have also been found in texts written by deaf children with sign language knowledge. The high number of consonant errors was explained as a consequence of lipreading, since the vowels are more distinct compared to the consonants (Sutcliffe et al., 1999). However, words will be easier to spell if they follow regular spelling patterns and children's ability to decode spellings seems to emerge by second grade. DHH children with residual hearing (i.e., HoH and children with CI) have, however, been shown to be more sensitive to phonological information compared to their deaf peers with sign language knowledge (Marschark, 2009).

Some of those children seem to show certain similar misspelled word patterns, as has previously been reported in hearing children. Studies on children with CI or hearing aids have reported that the children's access to sound enables many of them to use sounding strategies while spelling, causing "plausible" spelling errors (spelling errors based on sounds) (e.g., Geers and Hayes, 2011; Gärdenfors, 2016). Interestingly, Geers and Hayes (2011) further reported that CI-users using sign language in addition to the oral language made more errors that were not plausible. This compared to CI-users who only used spoken language. However, this group still faced difficulties in spelling due to lack of phonological awareness.

In a study on American English, Straley et al. (2016) explored spelling in narrative texts from twenty children using CI. The children were between 8.9 and 12.7 years of age, and all had spoken language as their primary language. The study found that, on average, 14% of all words were misspelled in the narratives. The children demonstrated an ability to represent correct sounds in words, which nevertheless resulted in misspelled words, such as doubling errors and omissions. However, it was shown that their spellings were not always conventional, which Straley et al. (2016) demonstrate with examples such as omissions ("cash" instead of "crash"), insertions ("drivier" instead of "driver") and doublings ("ticet" instead of "ticket"). In the last example, the child is able to represent each sound in the target word, but fails to express the conventional spelling of the/k/sound, using the 'ck' diagraph (i.e., a combination of two letters that represents one sound).

In yet another study, 69 DHH children (with CI and hearing aids) using spoken language, between the age of 10 and 11, were compared with children with dyslexia. They were provided with a test battery consisting of standardized assessments such as non-verbal intelligence, reading and spelling, speech and language skills. The authors found striking similarities in spelling, word reading and non-word reading in both DHH children and children with dyslexia, and in line with earlier studies, the DHH-children showed poor phonological awareness. The children with dyslexia had a larger vocabulary than the DHH-children, and vocabulary was shown to be a strong predictor for good literacy outcomes for the group of deaf children, but not for the group with children with dyslexia (Herman et al., 2019).

Taken together, DHH children often face difficulties in phonological awareness (e.g., Harris and Beech, 1998; Geers and Hayes, 2011; Harris and Terlektsi, 2011; Bowers et al., 2016), likely due to their hearing loss (Sterne and Goswami, 2000; Alamargot et al., 2007). The degree to which sign phonological awareness is transferred to spelling in writing, however, still remains unexplored. Our starting point in this study is that there might be differences depending on children's hearing status as well as their language abilities and backgrounds (cf. Hirshorn et al., 2015).

Spelling in Bilinguals

Swedish spelling research in bilinguals or second language learners on the word-level is limited, to our knowledge, to a handful of student essays, thus referring us to the international spelling literature. Figueredo (2006) reviewed twenty-seven

English as a second language (ESL) spelling studies. The most apparent difference between ESL and monolinguals regarding spelling development is that ESL-learners have an additional language that may be mirrored in their English spellings – i.e., a pattern of cross-linguistic influence through L1 transfer. An assumption within the concept of transfer is that if the language patterns in both languages are similar, the transfer would be facilitating, but if it differs, it would cause an interference, i.e., a so-called negative transfer. An example of interference is that Chinese and Japanese do not have relative clauses, resulting in Chinese and Japanese speakers using fewer relative clauses in English compared with other L2 who have relative clauses in their languages (Benson, 2002). Figueredo (2006) reported that the more the ESL-learners acquired the spelling norms of the second language, the less they relied on their first language, and that they followed the same spelling development as monolinguals.

A large study of reading and spelling development in the first two grades of elementary school included 1,812 children who were native speakers of Dutch and 331 bilinguals from Mediterranean or from Dutch colonies. The results showed that the spelling in the L2 was less efficient, and the children lagged in their phoneme–grapheme knowledge as well as phonemic segmentation compared to the L1. This was explained as difficulties in phoneme distribution rules, and that the children using the L2 had not developed the same automaticity for complex orthographic patterns and phonemic mapping as their L1 peers (Verhoeven, 2000).

Another study compared the ESL-learners with deaf children (mean age 10.7), also taking sign language into consideration in order to discuss possible transfer patterns from sign language to written English. The study found that deaf children made more omissions, insertions, and consonant errors and that the ESL-children showed more vowel errors and substitutions. Many spelling patterns of the deaf were non-phonetic and differed from the errors of the ESL children, who were more phonologically aware than the deaf children (Sutcliffe et al., 1999). The authors also found influence from British Sign Language (BSL) in the spelling of the deaf children. One fifth of those misspelled words of the British deaf children represented the initial letters only, which was explained by a fingerspelling influence from BSL through many incorporations of initialized signs (i.e., signs with a handshape corresponding to the fingerspelling of the word in the written/oral language) (Sutcliffe et al., 1999).

However, this point is debatable since Brown and Cormier (2017) reported that initialized signs in BSL are rare compared to one-handed systems such as American Sign language (ASL), indicating that BSL should be less amenable to initialization. Nevertheless, initialized signs are more common in one-hand systems such as in ASL. Lepic (2015) reported that approximately 15% of conventional lexical signs in ASL are initialized. There is no published study on initialized signs in STS, nevertheless a search in the STS corpus (Svensk Teckenspråkskorpus, 2019) shows that 13% of the STS signs are initialized. Padden and Ramsey (2000) report that skilled deaf signers could take advantage of initialized signs by using them as clues and translate them into English words. But, a "(non)initialized sign" can also cause false clues. Bowers et al. (2016) examined spelling in deaf

children with ASL knowledge, and found that initial handshapes from ASL influenced the English spelling, such as "vorival" instead of "funeral." This influence comes from the fact that the corresponding sign of "funeral" is expressed with two "V's" using both hands.

Another study reported that deaf children showed fewer function words and had a high repetitiveness of the same words. It was suggested that this was a result of the fact that the function words in ASL are limited compared to English, and consequently this was a form of transfer from ASL (Singleton and Newport, 2004). In yet another study, this time of Dutch sign language, deaf students were divided into two groups: low- or high proficiency signing groups. The high-proficiency signing group was found to omit more obligatory articles compared to the low-proficiency signing group. This was explained to be an artifact of Dutch sign language, since sign languages often lack obligatory articles (Van Beijsterveldt and van Hell, 2010).

A very limited number of studies describe the literacy development of CODAs. Some report a similar literacy development pattern for CODAs as for hearing children's first language acquisition (e.g., Brackenbury, 2005); others show that their language is reminiscent of second language learners (Larsson, 2015).

Swedish Research on Spelling

In Sweden, the most comprehensive study of the spelling of hearing children is Nauclér (1980), who provides a deeper insight into different kinds of misspelled words typical for hearing children, especially concerning doubling errors which often are challenging for younger children. Here, doublings errors are defined as when a misspelled word lacks "required doubling and non-required doubling" (Nauclér, 1980, p. 55). In Swedish schools, children are often told to use the strategy to "listen to how it sounds" to find out the spelling of a word. This reflects a common misunderstanding about Swedish spelling that many Swedish teachers share; in fact, the Swedish spelling conventions can, in many cases, be better described as based on long, short, stressed or unstressed vowels. Spellers can use this information to figure out if the following letters will consist of one or two consonants, since the length or stress of an underlying vowel will determine the following number of consonants. However, there are several exceptions violating this rule (Nauclér, 1989). Beyond the doubling errors, there are other spelling error categories, such as insertions, omissions, inversions, letter substitutions, errors in diacritic letters, confusions of similar words and influence from STS.

Wengelin (2002) was the first Swedish researcher who observed the writing process of DHH adults with help of a keystroke logging tool. Today, there is a handful of writing process studies in DHH, using keystroke logging tools (i.e., Asker-Árnason et al., 2010, 2012). Keystroke logging has advantages for research on spelling. If misspelled words are analyzed in the final version of the text (which is the most common way to analyze spelling), we miss the opportunity to study the writing process during which the words were written (Wengelin, 2002). In the final text, we know nothing about which words may have been deleted from the text, or which words

that may have been revised. Neither can we know about spelling attempts or cognitive efforts (measured by, e.g., pauses before, within or after a word with a spelling error).

Fingerspelling

There are slightly different views on the linguistic status of fingerspelling in the sign language linguistics literature. Nevertheless, it is a regularly used component in many sign languages, including STS, as fully lexical signs (Johnston and Schembri, 2010; Hodge and Johnston, 2014). Using a manual alphabet, is used to convey places, personal names, or other words for which there is no sign equivalent. Fingerspelling is expressed in representations of written words and enables connections between a sign language and written words (Bergman, 2012). Fingerspelling is learned naturally and early, and studies have shown that younger deaf children understand fingerspelling as soon as they start learning to communicate, that they perceive it as signs, and they are also able to show attempts to fingerspell themselves. However, their attempts will naturally be limited due to motoric reasons (Padden, 2005; Bergman, 2012). Padden (2005) describes how deaf children learn to fingerspell twice – as young children they will first identify fingerspelling as a sign but as they get older, they will learn that fingerspelling has further linguistic patterns, and that a handshape represents a letter.

Kelly (1995) and Humphries and MacDougall (1999) showed that deaf adults (teachers as well as parents) use fingerspelling and chaining considerably more during communication with their students, compared to hearing teachers. *Chaining* is when an adult pedagogically gives a sign and/or points out a printed word and fingerspells it again to establish a connection between the sign and its written word. The difference between deaf and hearing adults lies in the fact that the deaf adults themselves had the experience of learning to understand the meaning of fingerspelled words before they could recognize printed words. Haptonstall-Nykaza and Schick (2007) showed that deaf children of deaf parents showed better results from fingerspelling training compared to deaf children of hearing parents. The same authors compared two ways to acquire English vocabulary: by a signing condition and by a fingerspelling condition. The results showed that the deaf children did better in the fingerspelling condition, under which they could recognize and produce more English words. The authors suggest that the lexicalized fingerspelling method is an appropriate way to establish a phonological link to printed words. Padden and Ramsey (2000) found a strong relationship between fingerspelling and reading ability in deaf children, and those who were skilled readers demonstrated good ASL-skills. The good signers were also more able to write down English words that were fingerspelled to them.

Lipreading and Mouth Actions

Many children as well as adults with residual hearing need lipreading as a support to understand spoken language. However, trying to teach profoundly deaf children or adults to lipread is described as "difficult" and "frustrating," since vowels are often visually distinct, while consonants are not. Deaf children have been shown to make more consonant errors in their spelling during writing as a result of trying to lipread a "silent"

spoken word with invisible consonants (Sutcliffe et al., 1999; Marschark, 2009).

STS, as well as other sign languages, contains mouth movements i.e., mouth actions too. In sign language research, two main mouth categories have been identified so far: mouthing and mouth gestures (Crasborn et al., 2008). Beyond the lexical mouthing (mouth action patterns based on spoken language), there are also mouth gestures that provide a sign with further adverbial meanings such as regularity and intensity (Bergman, 1982, 2012). For this study, "mouthing" is relevant, as the visual phonetic elements from words in spoken languages are expressed without voice and used simultaneously with a manual sign, for example the Swedish sign for "HUS" ('HOUSE'), uses mouthing based on the Swedish word for the house i.e., "hus." However, unlike the spoken language, mouthing in STS follows a strict pattern, that is reduced in comparison to spoken languages. An example is the Swedish word "medlem" ('member') which is reduced to "mem" while signing it (Bergman and Wallin, 2001; SOU, 2006:29). In our data, spelling errors based on such reduced mouthing have been found in deaf children. Two recurrent errors are "falska" and "börd" instead of the correct "flaska" and "bröd" ('bottle' and 'bread'). The reason is that STS mouth movements are reduced to "fa" and "bö" (Gärdenfors, 2016). Also, since "falska" and "börd" are existing words in Swedish ('false' and 'descent'), consisting of the same, but reversed letters, it may be challenging for deaf children to learn the difference.

The Present Study

In this study, we aim to examine how the four background variables of the DHH children: *STS knowledge*, *hearing loss*, *deafness* (including hard-of-hearing children without use of spoken language) and *bilingual* experience together, and separately, contribute to children's spelling skills. Connected to this aim, we discuss which strategies and patterns DHH children, especially children with STS knowledge, show and use in their spelling. In this investigation, we have carefully selected children with different linguistic and hearing backgrounds, based on the four studied variables above. The participants consist of 33 children with variation in their degrees of hearing, use of spoken language, and in their language backgrounds, as being monolinguals or bilinguals. Each participant is categorized as a monolingual, unimodal bilingual (bilingual in two spoken languages), bimodal bilingual, (bilingual in spoken Swedish and in STS) or a sign-print bilingual, (bilingual in Swedish sign language and in written Swedish). Our research questions are the following:

- What do the writing processes and the spelling patterns look like in children with different linguistic and hearing backgrounds?
- Are any of the following variables: STS knowledge, bilingualism, hearing loss or, deafness, mirrored in these children's writing and spelling patterns? If so, in which group of children?
- Can we identify which strategy the children with STS knowledge use in order to spell?

TABLE 1 | Thirty three children participated in this study.

YES = X	Bilingual	STS	Hearing loss	Deafness
1	X	X	X	X
2	X	X	X	X
3	X	X	X	X
4	X	X	X	X
5	X	X	X	X
6	X	X	X	X
7	X	X	X	X
8	X	X	X	*
9	X	X	X	*
10	X	X	X	*
11	X	X	X	*
12	X	X	X	*
13	X	X	X	*
14	X	X	*	*
15	X	X	*	*
16	X	X	*	*
17	X	X	*	*
18	X	X	*	*
19	X	X	*	*
20	X	*	*	*
21	X	*	*	*
22	X	*	*	*
23	X	*	*	*
24	X	*	*	*
25	*	*	*	*
26	*	*	*	*
27	*	*	*	*
28	*	*	*	*
29	*	*	*	*
30	*	*	*	*
31	*	*	*	*
32	*	*	*	*
33	*	*	*	*

NO = *

Note that some of the children overlap in several variables. For example, a child may have a hearing loss, and be mastering STS, and is therefore a bilingual.

MATERIALS AND METHODS

Participants

For the present study, 33 children (23 girls, 10 boys) between 9.9 and 11.6 years were recruited. Of these, 19 were children with STS knowledge (mean age: 10.9 years) and 14 were children without STS knowledge (mean age: 10.9 years). Their background information is presented in **Table 1**. 19 participants were bimodal bilingual, mastering Swedish and STS consisting of five deaf children, four HoH children, four CI-users and six CODAs. No spoken or written Swedish tests were administered, however, the background questionnaire reported no writing or reading difficulties for the children. Out of the 19 children with STS knowledge, seven DHH-children attended a school class for the deaf and had not developed any spoken language, whereas the other children could communicate by speech and attended

a public school (i.e., mainstreamed with hearing children) or a special school class for hard-of-hearing children in which spoken Swedish was the primary language. All CODAs attended a mainstreamed school.

The remaining 14 children had normal hearing, and were without any knowledge of STS. This group consisted of five unimodal bilinguals and nine monolinguals. Beyond Swedish, the unimodal bilinguals communicated fluently in spoken Spanish, Danish, Thai, Dutch or Kurdish at home with their foreign-born parents. All of them attended a Swedish school and they were reported to master their two languages fairly equally. Unfortunately, we were not able to test their different languages, so our discussion about influence from other languages will be limited to possible influence from STS. The remaining nine participants were normally hearing monolinguals, mastering spoken and written Swedish.[1]

The inclusion criterion for DHH-children was that they should be born with hearing loss. Five children were profoundly deaf (<90 dB) and attended a class for deaf children. Four HoH children had a moderate to severe hearing loss without hearing aids (40–69 dB), and a mild to moderate hearing loss (25–54 dB) with hearing aids. However, two HoH-children have not developed spoken language and were therefore identified as deaf (Deaf HoH). All of the four CI-users were born profoundly deaf, and their first CI was implanted between the age of 9 months and 2 years and 2 months. Three of four CI-users have two implants, and with CI, their hearing was equal to a mild hearing loss (25–39 dB).

The inclusion criteria for the signing group was to be born into deaf families with STS knowledge, or into a family with parents who have started to learn and use STS early in the life of their child. Beyond the CODAs, 11 of 13 DHH-children had two deaf parents, and two children had two hearing parents, however, these parents had taken STS interpreter classes for several years (one of them is a certified STS interpreter) and are very skilled signers. In order to ensure the signing children's STS-knowledge, we provided a STS-test, see the SignRepL2 section.

The scoring of SignRepL2 is based on a five-point scale, i.e., the maximal score for each STS sentence is 4 points (0–4) and the participants with STS as a first language tend to reach total mean score close to 4.0 on this test, while the children without any STS knowledge often are able to copy around the half of the manual signs only, due to the gesture content, but they leave out crucial linguistic parts of the signs, such as grammatical and non-manual functions. The test revealed that the children with STS-knowledge received a mean of 3.78 (SD: 0.19), and the children without STS-knowledge received a mean of 2.11 (SD: 0.20).

Keystroke Logging

In order to capture the children's writing processes, we used keystroke logging, a well-established method for investigating the writing process. In this case we chose to use ScriptLog, which is a program that documents everything the writer does with the keyboard or mouse during the writing session (Wengelin

---

[1]The majority of the students in Swedish schools start learning English in the 3–4th class and may have basic knowledge of English by the age of 10–11.

et al., 2019). This includes documenting revision processes, and pausing behavior. Through replaying or by studying a linear representation of the writing processes the researcher can understand more about the production of a text (Leijten and Van Waes, 2013). To the writer, ScriptLog looks like a simple word processor, with a start and stop button that can be administered by the researcher or the writer. In this version of the program, no spellcheck is included.

For the writing task, the children were provided with a two-page cartoon and were instructed to free-write a story from the cartoon on the computer. They were provided unlimited time, but their average writing time was 29.4 min. The output from ScriptLog consist of, on the one hand, the final text, i.e., the text as it is when the writer has finished writing, and on the other hand, generated information about their writing process, in the form of a linear text. The final text provides a starting point for analyzing linguistic features. Further, ScriptLog's linear representations enable investigations of pauses and revisions, that took place during writing, but are not visible in the final texts.

Writing Task

Children's knowledge of the narrative genre is already established during pre-school years (Berman and Slobin, 1994), and we thus expected all the children in this study to have experience with, and basic knowledge of, the narrative genre. The stimulus for the written task consisted of a two-sided narrative cartoon about the *Pink Panther*. First, the use of picture-elicited narratives is a well-established method that has been used in earlier studies with Swedish DHH children (i.e., Schönström, 2010; Gärdenfors, 2016) and has provided robust outcomes of children's written production, feasible for analysis. Second, as the scope of the cartoon's content is delimited, the children are constrained to this in their writing, which leads to a delimited range of generated vocabulary output.

Further, this design enables us to compare and see how the children spell recurring words, and how they find other solutions such as synonyms and descriptions. The reason why we gave them unlimited time to finish the task was to eliminate the risk of them not showing their best ability if they got interrupted in the middle of the story. This choice was partly made based on the outcome of von Koss, Torkildsen et al. (2016), who provided 10 min writing time for their participants and found that assessment of the participants' narrative competence was not accurate due to the shortness and incompleteness of their texts, caused by the time pressure.

Since the typing speed may be slower in younger children compared to older and more experienced writers, we expect that their low-level-processes, such as typing skills and spelling ability, will not be fully automatized yet, but that this will be evenly divided between the groups (cf. Berninger and Swanson, 1994; Wengelin, 2006). The average writing speed of the participants was 10.5 words per minute.

SignRepL2

In order to measure the participants' STS proficiency, we used a STS repetitive test, called *SignRepL2* (Schönström and
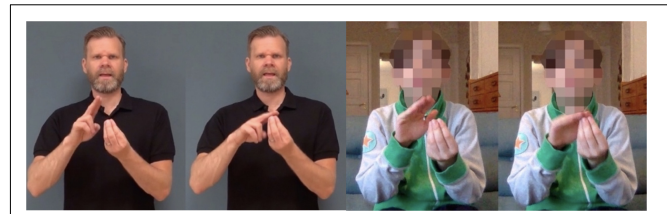
FIGURE 1 | The first print screens from SignRepL2 represent the target sign for "ÄGG" ['EGG']. The second print screens represent a signing attempt of a participant who did not achieve full points because of incorrect hand shape and absence of mouth movement. Permission and written consents for using the print screens were obtained from the individuals and their parents, as well as the copyright holders of SignRepL2 (Schönström and Holmström, 2017).

Holmström, 2017; Holmström, 2018). In the test the participants were shown fifty sentences, presented to them on a computer, and were asked to recall the sentences as presented, as exactly as possible during recording. The sentences increase in difficulty from simple single-sign items to three-sign sentences, see **Figure 1**. The 10-min test was originally developed for measuring L2 learners' STS proficiency, but was used here for assessing the participants' STS knowledge since there is no official standardized STS test for children available. This test has been tried out on 52 Swedish DHH children with STS as L1 or L2 by the developers of this test between 2016 and 2018 (Schönström et al. in preparation), and the measure of their STS results showed a valid difference between L1 signers and L2 signers. Based on this, and since no other tool is available, we expect that SignRepL2 should be suitable for the children of this study too. Testing STS knowledge is motivated by the fact that many writing studies on DHH children do not consider sign language in studying children's reading or written proficiencies. Knowledge of the children's STS proficiency is grounds for discussing possible cross-linguistic influence patterns between STS and writing.

Procedure

The data collection was carried out in three steps. First, the children and parents were recruited through networks, schools or hospitals. After an appointment with a child was booked, the parents filled in a consent and background form about the child's school, language use, and hearing background. The majority of the data collection took place in schools and hospitals. However, for practical reasons, some data was collected in homes in a quiet room. During the test sessions, the children received identical instructions from the first author, and they were informed that they could not ask for any help during the sessions. Every session started with the writing task and ended with the SignRepL2-test.

Analysis of Writing Process and Spelling

An analysis of the writing process was the first step. Due to the automatic output from ScriptLog, a great deal of information from the writing process was retrieved: number of words, writing time, pause length, number of pauses, and pauses before, within and after words. For this study we used an *ad hoc* pause criterion of 1 s, which served

TABLE 2 | Eight spelling error categories with descriptions are presented with examples of Swedish and English corresponding errors.

	Description	Swedish examples	English examples
Doubling errors	When a consonant is doubled or when the second consonant is missing (Nauclér, 1980)	Råta (Råtta) ['rat'] Mussen (musen) ['the mouse'] Äklig (äcklig) ['disgusting']	Faithfull, Ticet
Letter insertion	When an extra letter is inserted (Wengelin, 2002)	Taxsi (Taxi) ['taxi']	Priemary, Dierect
Letter omission	When a letter is missing (Wengelin, 2002)	Ijäl (ihjäl) ['to death']	Belive, Goverment
Inversions	When two letters change place (Wengelin, 2002)	Cylka (cykla) ['bike'/'biking']	Freind
Letter substitution	When an incorrect letter is replaced instead of the intended letter (Wengelin, 2002)	Sjönt (skönt) ['pleasant'] Sengen (sängen) ['the bed']	Repitition, Definitaly
Errors in diacritic letters	Accurate in Swedish, when letters with dots are confused with other letters that look similar (Svartholm, 2006)	A, Å, Ä and O, Ö (Swedish) Äffar (affär) ['store']	
Confusion of similar words	When using a word that looks like another word (Andersson, 1994)	Fantastisk and faktisk ['fantastic' and 'actually']	Expect, except and desert, dessert
Influence from Sign Language	When a child shows any influence from STS, for example when a spelling error is influenced from STS reduced mouth movements (Bergman and Wallin, 2001) or from a handshape (Bowers et al., 2016)	Falska and börd (Flaska and bröd) ['bottle' and 'bread'] Rätt instead of rädd ('right' and 'scared')	Sorpt instead of sport Vorival instead of funeral (the handshape for funeral is formed as a V in ASL)

the purpose of excluding the shortest pauses (which were more likely to be associated with motoric skills and finding a key), while including the longer pauses that could shed light on spelling processes. While we may have missed some relevant pauses, this pause criterion serves the purpose of the focus of the present study. Data further included measures of production rate, i.e., words per minute, and number of pauses per minute.

In addition, we manually identified all occurrences of misspelled words in the final text, and in the linear text, which included misspellings that were removed or corrected during the writing process. All spelling mistakes were sorted into eight different categories, see **Table 2** for an overview. As a result, we could calculate every child's spelling awareness, i.e., how likely it is that the child will detect and correct a spelling mistake. An example is that a child may misspell twenty times in total during the writing process and in the final text, but may only recognize five of them, and remove or correct them. Thus their spelling awareness will be 20% (5/20 = 20%). The higher the percentage spelling awareness is, the better the spelling. To our knowledge, this way of investigating spelling awareness by using keystroke logging is the first of its kind.

The majority of the spelling analysis criteria derive from Wengelin (2002). First, existing words that were ungrammatical such as "Yesterday I have jump," were not counted as misspelled words since the analysis excludes grammatical errors such as morphological errors. Another criterion is that if a word was used incorrectly in terms of meaning, for example "except" instead of the target word "expect," this would be counted as a spelling error. Note, that a misspelled word may be included in two or more spelling error categories (Ejeman and Molloy, 1997). For example, the word "fela" ("fälla," 'trap') belongs to the categories of doubling errors (when a consonant is doubled or when the second consonant is missing) and letter

substitution (when an incorrect letter is replaced instead of the intended letter). Because of this, the concepts of misspelled words versus misspellings will be distinguished from each other, in order to avoid choosing a misspelled word belonging to a particular category by neglecting another. A misspelled word is taken to be the misspelled word itself, and the misspellings on the other hand are the number of misspelling categories counted in a particular misspelled word. The frequency of misspellings will, therefore, be greater than that of the misspelled words.

Writing texts on the computer, with the use of a keyboard, may result in writing errors unrelated to spelling skills. These "typos" typically occur when a writer presses an adjacent key instead of the intended one (e.g., 'anf' instead of 'and' on a QWERTY keyboard), or when a writer presses two keys in the wrong order (e.g., 'adn' instead of 'and'). These so-called "typos" will generally not form any existing word, but may instead often violate the phonology of the language. The research literature that studies writing processes with keystroke logging has often observed that errors with typos are generally not associated with pauses (as an indication of increased cognitive load), and that typos are often immediately corrected. Wengelin (2002) compares typos to "counterparts in writing of articulatory errors in speaking" (p. 79). Research has shown that typos are common errors by children (and adults) who demonstrate no other spelling difficulties (Johansson, 2000). In the current study, we have chosen to exclude writing errors that can be categorized as "typos" from the spelling errors we investigate. The reason for this is that we are interested in describing the children's spelling abilities, and not their general typing abilities, or abilities of error detecting.

Statistical Analysis

In order to compare the results between the overlapping groups, we performed a regression analysis. As the means

and SD were counted (see **Table 3** for a result overview), the statistical analysis was fitted on all results, with help of the statistical program R, including number of words, writing length, words/minute, misspelled words in the final text, misspelled words in total, misspellings, spelling attempts, spelling awareness, pause length, pause/minute, number of pauses per minute, pauses before, within and after words, the STS-test and finally, the spelling error categories that can be found in **Tables 4–6**.

RESULTS

Table 3 provides an overview of the mean and SD, including length measures, writing processes, spelling errors and STS-test result. A regression analysis, based on **Table 3**, can be found in **Tables 4–6**. **Tables 4, 5** show the statistical results on the length measures, writing process, spelling and the STS-test. **Table 6** shows the statistical results on the spelling categories from **Table 2**.

Overall Result of the Groups

In **Table 3**, the column to the left displays the overview result of the length measures, writing process and the spelling error categories divided by the variables: no sign language, no bilingualism, no hearing-loss and deafness displayed on the top of the table. The top column also displays number of participants in each group (N), their mean age, mean and SD of the results.

Regression Analysis

In order to investigate the effects on the spellings, we performed a regression analysis on different writing and spelling measures, which are summarized and divided in **Tables 4, 5**. In these tables, the columns to the left display the results in length measures and writing process divided by the variables: no sign language, no bilingualism, no hearing loss and deafness. The next column displays the output from the regression analysis with the following: estimated difference, degree of freedom (DF), F-value, P-value, t-value, error, adjusted R-square and confidence intervals on a 2.5 and 97.5% level.

In this regression analysis, six effects ($p \leq 0.1$) were found in **Tables 4, 5**, of which three were significant ($p \leq 0.05$). The first effect was found on the number of words and deafness, $F(4.28) = 4.156$, $t = -1.907$, $p = 0.067$. The estimated word difference between the groups was -99.95 words, with a standard error of 52.4. The overall model fit was $F(4.28) = 4.156$, $t = 7.168$, $p = 0.0091$, $R^2 = 0.283$. The second and third effects were found in writing time in minutes on the predictors no bilingualism ($\beta = -12.51$, $p = 0.0802$, with a standard error of 6.9) and deafness ($\beta = -13.54$, $p = 0.059$. with a standard error of 6.8). The overall model fit was $F(4.28) = 2.548$, $t = 6.377$, $p = 0.0614$, $R^2 = 0.162$.

The first significant effect was found on pauses after words on the predictor, no hearing loss ($\beta = -5.7\%$, $p = 0.028^*$, with a standard error of 2.5%), the overall regression model fit was $F(4.28) = 2.837$, $t = 5.052$, $p = 0.043^*$, $R^2 = 0.1867$. Finally, two significant effects were found in SignRepL2, the STS test, (beta coefficient $= -1.5$, $p < 0.000^{***}$ with a standard error of 0.1), and

in deafness: (beta coefficient $= 0.2$, $p = 0.0427^*$ with a standard error of 0.1). The overall model fit was $F(4.28) = 172.7$, $t = 49.837$, $p < 0.000^{***}$, $R^2 = 0.956$.

In **Table 6**, the column to the left displays the investigated results on the *spelling error categories*, divided by the variables: no sign language, no bilingualism, no hearing-loss and deafness. The next column displays the output from the regression analysis with the following: estimated difference, degree of freedom (DF), F-value, P-value, t-value, error, adjusted R-square and confidence intervals on a 2.5 and 97.5% level.

Table 6 represents the second regression analysis that was fit on spelling error categories and four effects ($p \leq 0.1$) were found in which two effects were significant ($p \leq 0.05$). The first effect was found in the doubling error and deafness. $F(4.28) = 0.9148$, $t = -1.857$, $p = 0.074$. The estimated difference between the groups was -3.6%, with a standard error of 1.9%. The overall model fit was $F(4.8) = 0.9148$, $t = 2.924$, $p = 0.469$, $R^2 = -0.0108$. The second effect was found in letter substitutions, with an effect on no bilingualism, $F(4.8) = 1.555$, $t = 1.804$, $p = 0.0820$. The estimated difference between the groups was 1.7%. The overall model fit was $F(4.28) = 1.555$, $t = 2.569$, $p = 0.2136$, $R^2 = 0.06492$. The first significant effect was found in confusion of similar words with an effect on deafness. $F(4.8) = 6.506$, $t = 3.763$, $p = 0.0008^{***}$. The overall model fit was $F(4.28) = 6.506$, $t = 0.462$, $p = 0.0008^{***}$, $R^2 = 0.41$. The last significant effect was found between influence from STS and deafness. $F(4.28) = 7.133$, $t = 3.656$, $p = 0.0011^{**}$. The estimated difference between the groups was 0.08%. The overall model was $F(4.28) = 7.133$, $t = 0.716$, $p = 0.0004^{***}$, $R^2 = 0.434$.

Interpretation of the Results
Writing Length

Effects ($p \leq 0.1$) were found for number of words and writing length. The deaf children wrote on average 100 fewer words than the others. This can be explained by the well-documented fact that bilinguals in general have a smaller vocabulary in each language because of divided inputs from two languages (see Bialystok, 2009 for a review). However, there is yet another factor that explains *why* the deaf children on average wrote fewer words than the other bilinguals. Unlike the other bilinguals, they cannot take advantage of their hearing, so they are physically restricted in acquiring spoken Swedish by using their hearing. They cannot overhear conversations, on either TV or radio (Singleton, 2004). As a result, the constant input of Swedish through hearing is smaller than that of the other bilinguals. Taken together, a combination of shared input from two languages, and limited access to hearing, may be mirrored in a smaller vocabulary and shorter writing length.

Writing Process

Except for the low number of words and shorter writing time, all children, including deaf children, have developed similar typing skills. Between groups there were small differences regarding writing process measures such as words per minute, pause length, number of pauses, pauses before words, and pauses within words. Thus, all children in the study demonstrate similarly good transcription skills.

TABLE 3 | The overview table displays the average and its SD in the overall categories: length measures, writing process, spelling error categories and STS-test in the variable no sign language, no bilingualism, no hearing loss and deafness.

Participants (N)	19		14		24		9		13		20		7		26	
Variable	Sign language skills		No sign language skills		Bilingual		Monolingual		Hearing loss		Hearing		Deafness		Full or residual hearing	
Age	10.9		10.9		11.0		10.6		10.8		11.0		10.7		10.9	
	Mean	*SD*	**Mean**	*SD*	**Mean**	*SD*	**Mean**	*SD*	**Mean**	*SD*	**Mean**	*SD*	**Mean**	*SD*	**Mean**	*SD*
Writing Length																
Number of words	268.5	114.2	296.9	109.0	284.8	121.9	269.4	81.1	221.8	86.5	318.8	110.6	175.7	64.3	308.8	104.6
Writing length in minutes	26.8	11.4	33.0	15.6	29.7	13.3	28.5	14.7	24.9	12.1	32.4	13.8	18.6	6.0	32.3	13.5
Writing Process																
Words per minute	10.7	3.8	10.3	4.5	10.2	3.5	11.3	5.3	10.0	4.5	10.8	3.8	10.4	4.7	10.5	4.0
Misspelled words in final text in%	4.5%	5	3.2%	2.7	4.2%	4.7	3.3%	2.9	4.6%	5.8	3.6%	3.0	3.3%	1.4	4.1%	4.8
Misspelled words in writing process and final text in%	6.5%	5.0	5.3%	4.6	6.1%	4.7	5.9%	5.3	6.7%	5.6	5.6%	4.3	5.6%	1.7	6.2%	5.4
Misspellings in final text in%	5.4%	5.4	3.4%	2.8	5.0%	5.0	3.4%	3.0	5.7%	6.1	3.8%	3.1	5.1%	2.2	4.4%	5.0
Number of spelling attempts	3.5	1.1	3.1	0.7	3.4	1.0	3.2	0.5	3.7	1.2	2.9	0.4	3.9	1.7	3.1	0.5
Spelling awareness in%	30.6%	26.1	40.6%	27.4	31.0%	27.5	43.1%	24.1	31.9%	28.0	36.3%	26.1	40.4%	26.0	34.0%	27.2
Pause length per text in%	61.5%	12.5	65.6%	8.9	62.9%	11.5	64.2%	10.9	62.4%	15.0	63.8%	8.1	61.2%	15.3	63.6%	10.1
Number of pauses per minute > 1 s	10.8	2.1	10.6	2.6	10.6	1.9	11.0	3.2	11.4	2.0	10.2	2.4	12.1	1.4	10.3	2.4
Pauses before words in%	29.1%	11.4	28.9%	10.5	29.5%	10	27.6%	12.4	29.5%	12.6	28.7%	1.0	33.4%	13.6	27.8%	10.0
Pauses within words in%	17.9%	9	19.5%	9.8	18.1%	8.6	19.7%	12.1	20.1%	9.8	17.2%	9.3	21.0%	7.7	17.9%	10.0
Pauses after words in%	7.2%	5.8	4.7%	2.3	6.7%	5.3	4.5%	2.0	9.2%	5.9	4.2%	2.3	9.5%	6.7	5.3%	3.8
Spelling Error Categories																
Doubling errors in%	2.3%	4.1	2.5%	2.3	2.5%	3.6	2.3%	2.6	2.2%	4.9	2.5%	2.2	0.5%	0.7	2.9	3.7
Insertions in%	0.3%	0.4	0.1%	0.2	0.3%	0.4	0.1%	0.2	0.4%	0.4	0.2%	0.4	0.3%	0.5	0.2	0.3
Omissions in%	1.1%	1.1	0.6%	0.8	0.9%	0.9	0.8%	0.9	1.3%	1.3	0.6%	0.7	1.5%	1.7	0.7	0.7
Inversions in%	0.3%	0.6	0.1%	0.2	0.3%	0.5	0.1%	0.2	0.4%	0.7	0.1%	0.2	0.5%	0.9	0.1	0.3
Letter substitutions in%	1.3%	1.6	1.7%	2.0	1.2%	1.4	2.3%	2.2	1.1%	1.4	1.7%	1.9	0.5%	0.6	1.7	1.8
Diacritic letters in%	0.4%	0.5	0.2%	0.5	0.3%	0.5	0.2%	0.4	0.3%	0.5	0.3%	0.5	0.4%	0.6	0.3	0.4
Confusion of similar words in%	0.7%	1.3	0.1%	0.2	0.6%	1.1	0.2%	0.2	1.0%	1.5	0.1%	0.2	1.8%	1.7	0.1	0.2
Influence from STS in%	0.4%	0.6	0.0%	0.0	0.3%	0.6	0.0%	0.0	0.5%	0.7	0.0%	0.0	1.1%	0.8	0.0	0.1
STS Test																
SignRepL2 (max 4.0)	3.8	0.19	2.1	0.20	3.4	0.69	2.1	0.18	3.8	0.19	2.6	0.78	3.9	0.03	2.8	0.83

TABLE 4 | The overview table displays the results from the regression analysis on the investigated results: number of words, writing time, words per minute, pause length, pauses per minute, pauses before, within and after words based on no sign language, no bilingualism, no hearing loss and deafness.

Factor: YES*	Estimated difference	DF	F	P	t	Error	Adjusted R-square	CI 2.5%	CI 97.5%
			LENGTH MEASURES						
Number of words	**275.67**	**4.28**	**4.156**	**0.0091**	**7.168**	**38.5**	**0.283**	**196.9**	**354.4**
No sign language**	−23.27			0.6865	−0.408	57.0		−140.1	93.6
No bilingualism	76.96			0.1542	−1.465	52.5		−184.6	30.7
No hearing-loss	94.00			0.0949	1.728	54.4		−17.4	205.4
Deafness	−99.95			0.067	−1.907	52.4		−207.3	7.4
Writing length in minutes	**32.17**	**4.28**	**2.548**	**0.0614**	**6.377**	**5.0**	**0.162**	**21.8**	**42.5**
No sign language	10.17			0.1851	1.359	7.5		−5.2	25.5
No bilingualism	−12.51			0.0802	−1.816	6.9		−26.6	1.6
No hearing-loss	−1.28			0.8585	−0.180	7.1		−15.9	13.3
Deafness	−13.54			0.059	−1.969	6.8		−27.6	0.5
			WRITING PROCESS						
Words per minute	**9.62**	**4.28**	**0.7163**	**0.5879**	**5.707**	**1.7**	**−0.037**	**6.2**	**13.1**
No sign language	−3.71			0.148	−1.486	2.5		−8.8	1.4
No bilingualism	3.01			0.202	1.307	2.3		−1.7	7.7
No hearing-loss	2.41			0.320	1.013	2.4		−2.5	7.3
Deafness	0.75			0.745	0.329	2.3		−4.0	5.5
Pause length per text in%	**59.5%**	**4.28**	**0.418**	**0.794**	**4.736**	**12.5%**	**-0.078**	**31.7%**	**94.5%**
No sign language	8.6%			0.228	1.232	7.0%		−5.7%	23.0%
No bilingualism	−4.0%			0.537	−0.626	6.5%		−17.2%	9.2%
No hearing-loss	−2.3%			0.727	−0.352	6.6%		−15.8%	11.1%
Deafness	0.6%			0.873	−0.161	3.5%		−6.6%	7.0%
Number of pauses per minute > 1 s	**10.59**	**4.28**	**1.421**	**0.253**	**11.517**	**0.9**	**0.04995**	**8.7**	**12.5**
No sign language	0.43			0.757	0.313	1.4		−2.4	3.2
No bilingualism	1.19			0.351	0.949	1.3		−1.4	3.8
No hearing-loss	−1.20			0.363	−0.924	1.3		−3.9	1.5
Deafness	1.51			0.239	1.203	1.3		−1.1	4.1
Pauses before words in%	**24.9%**	**4.28**	**0.5668**	**0.6888**	**5.454**	**4.6%**	**−0.05725**	**15.5%**	**34.2%**
No sign language	2.8%			0.683	0.413	6.8%		−11.1%	16.6%
No bilingualism	−3.5%			0.581	−0.559	6.2%		−16.2%	9.2%
No hearing-loss	3.5%			0.597	0.535	6.4%		−9.7%	16.6%
Deafness	8.6%			0.178	1.381	6.2%		−4.1%	21.3%
Pauses within words in%	**20.4%**	**4.28**	**0.9471**	**0.4516**	**5.241**	**3.9%**	**-0.0067**	**12.4%**	**28.4%**
No sign language	7.1%			0.226	1.237	5.7%		−4.7%	19.0%
No bilingualism	0.7%			0.892	0.137	5.3%		−10.2%	11.6%
No hearing-loss	−8.5%			0.132	−1.553	5.5%		−19.8%	2.7%
Deafness	5.8%			0.914	0.109	5.3%		−10.3%	11.4%
Pauses after words in%	**8.8%**	**4.28**	**2.837**	**0.043***	**5.052**	**1.7%**	**0.1867**	**5.2%**	**12.4%**
No sign language	1.8%			0.492	0.697	2.6%		−3.5%	7.1%
No bilingualism	−0.4%			0.884	−0.148	2.4%		−5.2%	4.5%
No hearing-loss	−5.7%			0.028*	−2.325	2.5%		−10.7%	−0.7%
Deafness	0.6%			0.791	0.268	2.4%		−4.2%	5.5%

*The regression analysis is based on Factor: YES.
**In this case, it means that the "No sign language: YES" reveals to children who do not master STS, write 23.7 words less than the children with STS knowledge. The gray highlighted values show effects on the results.

Spelling

The range of the misspelled words in percentage was not significant for the studied children, and in order to increase the validity, their results were compared with other Swedish spelling studies on normal-hearing and DHH children. The percentage of misspelled words for the hearing monolinguals of this study was surprisingly low, with an average of 3.3% in their final texts, while previous Swedish studies, including 67 children in similar age, showed an average of 8.5% misspelled words (based on studies by Bäck, 2011; Gärdenfors, 2016; Raatikainen,

TABLE 5 | The overview table displays the spelling results from the regression analysis on the investigated results: misspelled words in final text, misspelled words in total, misspellings, spelling attempts/text, spelling awareness and Sign-RepL2 based on the variables: no sign language, no bilingualism, no hearing loss and deafness.

Factor: YES	Estimated difference	DF	F	P	t	Error	Adjusted R-square	CI 2.5%	CI 97.5%
				SPELLING					
Misspelled words in final text in%	**6.0%**	**4.28**	**0.4921**	**0.742**	**3.349**	**1.8%**	**−0.0678**	**2.3%**	**9.7%**
No sign language	−1.5%			0.585	−0.552	2.7%		−6.9%	4.0%
No bilingualism	0.3%			0.902	0.124	2.5%		−4.7%	5.3%
No hearing-loss	−1.5%			0.549	−0.607	2.5%		−6.8%	3.7%
Deafness	−2.7%			0.283	−1.96	2.4%		−7.7%	2.3%
Misspelled words in writing process and final text in%	**8.0%**	**4.28**	**0.3708**	**0.827**	**3.930**	**2.0%**	**−0.854**	**3.8%**	**12.2%**
No sign language	−1.7%			0.585	−0.552	3.0%		−7.9%	4.5%
No bilingualism	1.4%			0.623	0.498	2.7%		−4.3%	7.1%
No hearing-loss	−1.8%			0.527	−0.641	2.8%		−7.8%	4.1%
Deafness	−2.4%			0.391	−0.870	2.8%		−8.1%	3.3%
Misspellings in final text in%	**6.5%**	**4.28**	**0.4897**	**0.743**	**3.369**	**1.9%**	**−0.068**	**2.5%**	**10.4%**
No sign language	−1.4%			0.626	−0.493	2.8%		−7.2%	4.4%
No bilingualism	0.1%			0.967	0.042	2.6%		−5.3%	5.5%
No hearing-loss	−1.7%			0.526	−0.641	2.7%		−7.3%	3.8%
Deafness	−1.3%			0.611	−0.515	2.6%		−6.7%	4.0%
Number of spelling attempts	**3.5**	**4.28**	**1.844**	**0.1485**	**10.079**	**0.4**	**0.0954**	**2.8**	**4.2**
No sign language	−0.1			0.870	−0.165	0.5		−1.1	1.0
No bilingualism	0.0			0.972	0.035	0.5		−1.0	1.0
No hearing-loss	−0.5			0.296	−1.065	0.5		−1.5	0.5
Deafness	0.4			0.421	0.816	0.5		−0.6	1.4
Spelling awareness in%	**44.5%**	**4.28**	**0.2876**	**0.8835**	**3.918**	**11.4%**	**−0.0978**	**21.2%**	**67.7%**
No sign language	8.3%			0.628	0.490	16.8%		−26.2%	42.7%
No bilingualism	−13.0%			0.410	−0.836	15.5%		−44.7%	18.8%
No hearing-loss	−8.5%			0.603	−0.527	16.1%		−41.3%	24.4%
Deafness	−4.9%			0.754	−0.316	15.5%		−36.6%	26.8%
Sign-RepL2 (max 4.0)	**3.7**	**4.28**	**172.7**	**0.000*****	**49.837**	**0.1**	**0.956**	**3.5**	**3.8**
No sign language	−1.5			0.000***	−13.926	0.1		−1.8	−1.3
No bilingualism	−0.1			0.239	−1.204	0.1		−0.3	0.1
No hearing-loss	0.0			0.825	0.223	0.1		−0.2	0.2
Deafness	0.2			0.0427 *	2.124	0.1		0.0	0.4

The gray highlighted values show effects at p ≤ 0.1. *Significance at p ≤ 0.05 level, **Significance at p ≤ 0.01 level, ***Significance at p ≤ 0.001 level.

2018). The teacher of the monolingual children in this study described them as an extraordinary class, so they have "set the bar high". However, the monolinguals were not the only children with very few spelling errors – the CODAs, CI-users and HoH children showed very low percentages, with about half as many misspelled words compared to previous studies (Bäck, 2011; Gärdenfors, 2016; Raatikainen, 2018), except for the deaf group in which the number of misspelled words was equal to that found in Gärdenfors (2016). We have unfortunately not been able to find any comparable Swedish study on unimodal bilingual children.

One reason why the spelling errors in the children with STS knowledge were fewer compared to older studies, may be due to the children's early STS knowledge. Several previous studies have shown a strong correlation between early sign language and literacy and spoken language (e.g., Svartholm, 2010; Hassanzadeh, 2012; Davidson et al., 2014) and it is

likely that this also is the case for spelling knowledge. This suggestion is reinforced by the equal percentage of spelling errors in the deaf children who were the only group who had full STS knowledge since childhood in this, and in the previous study (Gärdenfors, 2016). An explanation may be that the majority of the children have deaf parents. Kelly (1995) and Humphries and MacDougall (1999) have documented that deaf adults are more prompt to use the *chaining*-method (showing a word or a sign and fingerspelling it to strengthen the link between the fingerspelling and the word) than the hearing adults – thanks to their own personal experience of learning to fingerspell twice (Padden, 2005). Using fingerspelling in Sweden is also reported by Bergman (2012) who observed that the adults use fingerspelling as a natural part of their communication with younger deaf children. When asked, many of the deaf parents of the participants of this study confirmed that they used fingerspelling to their children from

TABLE 6 | The overview table displays the results from the regression analysis on the spelling errors based on the variables: no sign language, no bilingualism, no hearing loss and deafness.

Factor: YES	Estimated difference	DF	F	P	t	Adjusted R-square	Error	CI 2.5%	CI 97.5%
Doubling errors in%	**4.1%**	**4.28**	**0.9148**	**0.469**	**2.924**	**−0.0108**	**1.4%**	**1.2%**	**7.0%**
No sign language	0.4%			0.839	0.205		2.1%	−3.9%	4.7%
No bilingualism	−0.7%			0.712	−0.374		1.9%	−4.7%	3.2%
No hearing-loss	−1.6%			0.437	−0.789		2.0%	−5.7%	2.5%
Deafness	−3.6%			0.074	−1.857		1.9%	−7.5%	0.4%
Insertions in%	**0.4%**	**4.28**	**0.881**	**0.4879**	**2.808**	**−0.0151**	**0.1%**	**0.1%**	**0.7%**
No sign language	−0.1%			0.526	−0.643		0.2%	−0.6%	0.3%
No bilingualism	−0.1%			0.704	−0.384		0.2%	−0.5%	0.3%
No hearing-loss	−0.1%			0.602	−0.527		0.2%	−0.5%	0.3%
Deafness	−0.1%			0.509	−0.669		0.2%	−0.6%	0.3%
Omissions in%	**1.0%**	**4.28**	**1.483**	**0.2339**	**2.657**	**0.05697**	**0.4%**	**0.2%**	**1.8%**
No sign language	−0.5%			0.3609	−0.929		0.6%	−1.7%	0.6%
No bilingualism	0.6%			0.2516	1.171		0.5%	−0.5%	1.7%
No hearing-loss	−0.3%			0.5650	−0.582		0.6%	−1.4%	0.8%
Deafness	0.5%			0.4030	0.849		0.5%	−0.6%	1.5%
Inversions in%	**0.2%**	**4.28**	**1.544**	**0.216**	**1.168**	**0.06371**	**0.2%**	**−0.2%**	**0.6%**
No sign language	0.0%			0.902	0.125		0.3%	−0.5%	0.6%
No bilingualism	−0.0%			0.905	−0.120		0.3%	−0.6%	0.5%
No hearing-loss	−0.1%			0.605	−0.523		0.3%	−0.7%	0.4%
Deafness	0.3%			0.156	1.457		0.3%	−0.1%	0.9%
Letter substitutions in%	**1.8%**	**4.28**	**1.555**	**0.2136**	**2.569**	**0.06492**	**0.7%**	**0.4%**	**3.1%**
No sign language	−1.3%			0.2216	−1.250		1.0%	−3.3%	0.8%
No bilingualism	1.7%			0.0820	1.804		0.9%	−0.2%	3.6%
No hearing-loss	0.1%			0.9208	0.100		1.0%	−1.9%	2.1%
Deafness	−1.2%			0.1933	−1.333		0.9%	−3.1%	0.7%
Errors in diacritic letters in%	**0.1%**	**4.28**	**0.9387**	**0.4561**	**0.594**	**−0.0077**	**0.2%**	**−0.3%**	**0.5%**
No sign language	−0.2%			0.586	−0.551		0.3%	−0.7%	0.4%
No bilingualism	−0.2%			0.416	−0.826		0.3%	−0.8%	0.3%
No hearing-loss	0.4%			0.128	1.570		0.3%	−0.1%	1.0%
Deafness	0.3%			0.268	1.131		0.3%	−0.2%	0.8%
Confusion of similar words in%	**0.1%**	**4.28**	**6.506**	**0.0008*****	**0.462**	**0.4077**	**0.3%**	**−0.5%**	**0.8%**
No sign language	0.0%			0.931	0.087		0.5%	−0.9%	1.0%
No bilingualism	0.1%			0.814	0.237		0.4%	−0.8%	1.0%
No hearing-loss	−0.1%			0.746	−0.327		0.5%	−1.1%	0.8%
Deafness	1.7%			0.0008 ***	3.763		0.4%	0.8%	2.5%
Influence from STS in%	**0.0%**	**4.28**	**7.133**	**0.0004*****	**0.716**	**0.434**	**0.2%**	**−0.2%**	**0.4%**
No sign language	−0.0%			1.0	0.000		0.2%	−0.5%	0.5%
No bilingualism	0.0%			1.0	0.000		0.2%	−0.4%	−0.4%
No hearing-loss	−0.01%			0.62	−0.506		0.2%	−0.6%	0.3%
Deafness	0.08%			0.0011**	3.656		0.2%	0.3%	1.2%

The gray highlighted values show effects on the results.

an early age, saying that "fingerspelling is a crucial part of Swedish sign language." Some of the parents had even read about the chaining-method and applied this on their children, since they believed that it would strengthen the relationship between fingerspelling and Swedish letters. The parents with STS knowledge may thus show how a word is spelled by fingerspelling it to their children, and in that way circumvent the sounding strategy by showing the visual alphabetic characteristics of a Swedish word to their CODAs and DHH-children. The understanding of the relationship between fingerspelling and

how a word is spelled would therefore have been facilitated in children with STS knowledge, compared to the other children who had access to the sounding strategy only. This relationship is also in line with Padden and Ramsey (2000) who found a strong relationship between fingerspelling and reading ability.

In the next section, three spelling categories with patterns that were likely to be caused from sounding and visual strategies will be highlighted and discussed to deepen our understanding of the participants' spelling.

ASPECTS OF SPELLING ERRORS

The heterogeneous nature of the quantitative part of this study is complemented by a qualitative inspection of the spelling errors. The aim of the qualitative approach is to investigate some patterns relating to STS knowledge and hearing ability, as revealed by the quantitative part of the study. Below, both the similarities and differences will be presented. All the spelling results were based on misspelled words occurring during the process *and* in the final text, in order to show the relevant tendencies. In this section, patterns in length measures, the writing process, spelling in general, and the spelling categories of doubling errors, confusion of similar words and influence from STS will be discussed.

Doubling Errors

An effect was found in doubling errors with the variable deafness ($p = 0.07$). The deaf children performed only 0.50% doubling errors compared to the others with 2.89%. We also observed that the many doubling errors in hard-of-hearing children were similar to the errors found in hearing children with typical errors such as "chokad" ('shocked'), "öpen" ('open'), "kunnde" ('could'), instead of "chockad," "open," and "kunde".

For the deaf children (black triangles in **Figure 2**), only two doubling errors were observed for five deaf individuals (representing an average of 0.4%), and the HoH and CI children showed on average 3.4% doubling errors, see **Figure 2** for an illustration. However, two of them could not use their residual hearing in order to communicate by speech, and are therefore defined as "deaf" (plotted in black quadrats as HoH deaf). In those two individuals, only four doubling errors were identified, but those "doubling errors" seemed rather to have occurred by accident. Such an example was written as "dröme" instead of "drömmer" ('dream', 'dreaming'). Since the word lacks an "m", it was counted as a doubling error following our criteria, however, this spelling also indicated a limited morphological knowledge. The Swedish noun is "dröm" ('dream'), and the child was likely trying to use this form to create the verb, but did it incorrectly. Thus, this error was probably not caused by a sounding strategy. Taken together, the observations indicate that there is a relationship between deafness and lack of doubling errors, so one of the most important contributions here is that when a visual and a sounding strategy are available, the hearing, hard of hearing and children with CI seem to use the sounding strategy rather than the visual strategy.

Confusion of Similar Words

A significant effect was found in the spelling category confusion of similar words and the predictor deafness. The deaf hearing children showed some patterns of non-semantic, however, visually similar looking words such as "fjälla" ('scale') instead of "fälla" ('trap'), "brev" ('letter') instead of "bredde" ('smeared'), "läder" ('leather') instead of "lägger" ('lay'). The same phenomenon is also described by Wengelin (2002) and Gärdenfors (2016). One explanation may be that when a deaf child cannot confirm the spelling by sounding the letters, they will likely reach for the most salient letters of a word. Since the mental
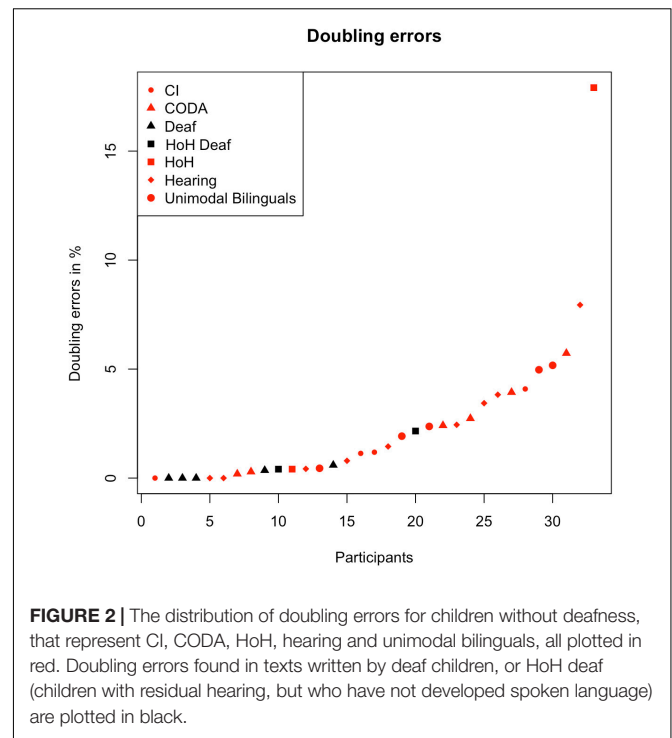
FIGURE 2 | The distribution of doubling errors for children without deafness, that represent CI, CODA, HoH, hearing and unimodal bilinguals, all plotted in red. Doubling errors found in texts written by deaf children, or HoH deaf (children with residual hearing, but who have not developed spoken language) are plotted in black.

representation of the reached-for word may still be diffuse, a visually nearby word will be used instead. Since the children cannot confirm the meaning by sounding this out, the process will as a consequence result in a semantically incorrect word.

Influence From STS

The last significant effect was found in the category of influence from STS and deafness. We identified three different kinds of influence from STS: by mouth; by handshapes; and by signs with different corresponding meanings in Swedish. First, mouth actions are a part of non-manual signals that are essential while signing because they fill important linguistic functions such as negation and adverbs for instance. Bergman (2012, p. 45, our translation from Swedish) writes: "[c]hildren acquire lexically determined mouth actions as natural, visual parts of the signs. Even before the age of two, such oral movements can be observed in children's communication." A similarity can be drawn with hearing children: when they learn new words, they also learn how to stress the vowels correctly. Since STS is the first language for some of the participants, we may expect that DHH participants, particularly those with no use of sound, rely on their acquisition of Swedish by looking at the mouthing (i.e., mouth actions based on borrowed elements from Swedish). However, length of mouthing is reduced to a few prominent segments (Bergman and Wallin, 2001), and we suggest that deaf children develop their phonological awareness on global characteristics, for example by how a word may look on the mouth (cf. Marschark, 2009) meaning that the DHH children rely on the spelling of the most prominent mouth segments, which is also reported by Sutcliffe et al. (1999). As a result, letters will be missing or reversed. This can be compared to how hearing children express words –

FIGURE 3 | The supplied mouth movements while signing "FLASKA" are reduced to the most salient [FA] that result in a spelling error: "faskla" when the deaf relied on the mouth movement while spelling. The images come from https://teckensprakslexikon.su.se (The Swedish Sign Language Lexicon) and are used with permission of the copyright holder.

a hearing child starts learning how to write words by uttering the word (Svensson, 1998), and can then be "misled" by the fact that some sound is missing in the spelling such as "hemst" instead of the correct "hemskt" ('horrible'), in which the "k" is unpronounced.

(a) Deaf: "då gav katten henne (. . .) <u>faskla</u>"
    ('the cat gave her a (. . .) <u>bottle</u>')

Example (a) represents an example of letter reversals that was likely caused by reduced mouthing in STS. The mouthing of the sign "FLASKA" ('bottle') is reduced to "FA" without any distinct movement for "L", see **Figure 3**. Identical patterns of the Swedish word "flaska" have also been reported by Schönström (2010) and Gärdenfors (2016), who found several variants such as "falska," "fasa" and "faka," all started with "fa," and not the supposed "fla" in deaf children's written production.

(b) Deaf: "Mus blir så <u>rätt</u> när se (. . .)"
    ('The mouse becomes so <u>right</u> when it sees (. . .)')

Second, example (b) shows when a profoundly deaf child bases a spelling on the handshape of a sign "as a false clue." The word was supposed to be "rädd," ('scared'), but the word was written as "rätt" ('right'). The STS handshape of the sign "RÄDD" is formed as a "t", so there is a high probability that the child in writing replaced "tt" instead of the supposed "dd," see **Figure 4**. Another interpretation is that this resulted from a confusion of similar words, since "rädd" and "rätt" are visually similar, and prior to not choosing a misspelled word of a particular category by neglecting another, this was also counted as a confusion of similar words.

(c) Deaf 1: "Lilly såg <u>mus din fot</u>"
    ('Lilly saw a mouse <u>your foot</u> (a mouse's footprints)')
(d) Deaf 2: "Och ser rosa katt <u>din</u> säng"
    ('The cat sees a sleeping mouse in <u>your</u> bed (his/her bed)')
(e) Deaf 3: "Katt tog musen och går till <u>ditt</u> säng"
    ('The cat took the mouse and went to <u>your</u> bed (to his/her bed)')
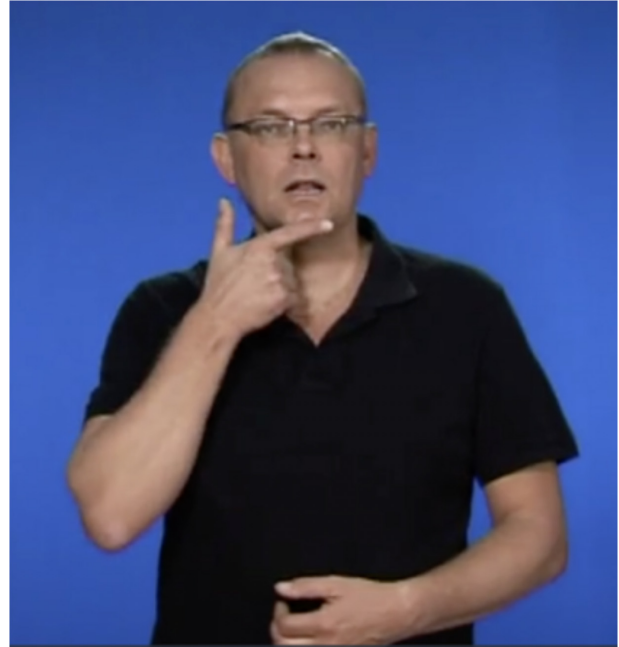
FIGURE 4 | When a spelling error derives from a sign's handshape. The picture shows the sign for "RÄDD" ('scared'), and its handshape is formed as a "t", resulting in the spelling error, "rätt" ('right'). The image comes from https://teckensprakslexikon.su.se (The Swedish Sign Language Lexicon) and is used with permission of the copyright holder.

Lastly, examples (c), (d) and (e) show an overuse of the word "din/ditt" ('your') when the supposed words would be "hans" ('his') or the possessive affix "s", in three profoundly deaf children, resulting in syntactical errors. The STS signs for "s", "DIN," "HENNES," "HANS" and "DERAS" ('your', 'her', 'his' and 'their') are identical, representing a flat hand moving forward from the signer, and as a result, the children choose an incorrect Swedish word, with, however, the same underlying signs in STS. Example (c) indicates that the child did likely not know how to spell "musfotspår" (a mouse's footprint) and tried to sign this word mentally from STS by rephrasing this to "mouse his/her foot". Examples (d) and (e) are similar examples in which when the participants tried to express "his/her" by writing "din/ditt" which has an identical sign to "DIN."

The findings of this study show both similarities and differences between the participants. The similarities could be found in the features of the writing process, particularly in words per minute, pause length in percentage, pauses per minute, pauses before and within words in percentages. Here, we may thus observer patterns typical for this age group. The differences are rather found in the variable deafness that explained the majority of the effects such as number of words, writing time, STS-test, doubling errors, confusion of similar looking words and influence from STS.

The first observation was that the bilinguals who were hard-of-hearing or CI-users showed a larger vocabulary than the bilingual

deaf children. We suggest that it is due to the fact that they could acquire Swedish by means of spoken Swedish.

An essential contribution of this study is also that the hard-of-hearing and CI-users, despite their daily use of STS, seemed to rather rely on the sounding strategy than the visual strategy that was mirrored in recurrent doubling errors and letter substitutions, often caused by the sound. But their access to the visual strategy was not absent since their proportion of spelling errors were considerably lower compared to previous Swedish studies. This was explained as a facilitation from STS, from for example fingerspelling of their deaf parents who can demonstrate how a word is spelled through fingerspelling and circumvent the sounding strategy by showing the visual alphabetic characteristics of a Swedish word. Those visual strategies are reinforced, especially in deaf children, who showed a higher tendency to 'spell as it looks,' and in this have confused similar-looking words since they could not double-check the meaning by sounding it.

A final important finding was that the direct STS transfers (by mouth, by handshapes and by signs with different corresponding meanings in Swedish), could in the first instance be found in deaf children and not in the other children with STS knowledge. Since they did not have access to the sounding strategy, the visual strategy was the only one available. But, due to their limited vocabulary, and when the visual strategy was not available (i.e., due to drawing on their visual memory), they had to use other strategies – characteristics and signs from STS, such as direct translation from STS or spelling a word based on how it looks on the lips.

CONCLUSION

Many of the spelling patterns found in this study confirm earlier findings in the field, that is, that a strategy that uses visual as well as auditory cues can, on the one hand, facilitate spelling, and on the other hand interfere with the spelling. Our present contribution is linked to how those strategies interact both together and separately. Our results indicate that auditory input is a crucial factor; when it is absent, the deaf children resort to visual strategies.

However, with regard to the DHH children, it is difficult to isolate and investigate the impact of auditory and visual input respectively. This needs to be addressed in future studies. Nevertheless, our results indicate that DHH children benefit from using input from both modalities. Further, the results have pedagogical implications and demonstrate the importance of teachers' awareness of the special challenges in learning to spell that the groups of STS and DHH children face. The absence of auditory input calls for an early and continuous input of visual channels, such as exposure to written words through reading, and by explicit training in the relationship between written words and fingerspelling. The latter point has also been shown to be beneficial even for children with residual hearing and also for hearing children as a complement to the auditory strategy.

According to SOU (2016:46) (an official investigation of the Swedish government), the majority of the congenitally deaf Swedish children receive CI before the age of 8–9 months, and some will receive CI as early as 5 months. If the children receive it before the age of 9 months, it is likely that many of them will develop an adequate spoken language. This investigation also reports that 80–90% of those Swedish children with CI attend a mainstream school, and the remainder who do not, attend special schools because of hearing problems or intellectual delays. This study on how sign language relates to spelling makes a significant contribution to the understanding of how basic writing skills are established in this group. Since the children with STS knowledge in the present study showed considerably fewer spelling errors compared to earlier studies, we want to highlight the supporting role that sign language seems to have in developing spelling skills. Having access to a bilingual repertoire with auditory as well as visual input provides these children with a wider range of strategies to make use of for spelling.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Etikprövningsnämnden in Stockholm. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Written informed consent was obtained from the individual(s), and minor(s)' legal guardian/next of kin, for the publication of any potentially identifiable images or data included in this manuscript.

AUTHOR CONTRIBUTIONS

MG (first author) mainly responsible for the data collection and analysis, dissemination of the results, and manuscript writing. VJ (supervisor) and KS (head supervisor) involved in writing parts of the manuscript.

ACKNOWLEDGMENTS

REFERENCES

Alamargot, D., Lambert, E., Thebault, C., and Dansac, C. (2007). Text composition by deaf and hearing middle-school students: the role of working memory. *Read. Writ.* 20, 333–360. doi: 10.1007/s11145-006-9033-y

Albertini, J. A., and Schley, S. (2010). "Writing: characteristics, instruction and assessment," in *The Oxford Handbook of Deaf Studies, Language, and Education*, eds M. Marschark, and P. E. Spencer, (Oxford: Oxford: University Press), 130–143.

Amraei, K., Amirsalari, S., and Ajalloueyan, M. (2017). Comparison of intelligence quotients of first- and second-generation deaf children with cochlear implants. *Int. J. Pediatr. Otorhinolaryngol.* 92, 167–170. doi: 10.1016/j.ijporl.2016.10.005

Andersson, R. (1994). *Sammanblanding av Formlika Svenska Ord Vid Dövas Läsning.* Stockholm: Stockholm University.

Asker-Árnason, L., Akerlund, V., Skoglund, C., Ek-Lagergren, I., Wengelin, A., and Sahlén, B. (2012). Spoken and written narratives in swedish children and adolescents with hearing impairment. *Commun. Disord. Q.* 33, 131–145. doi: 10.1177/1525740111401906

Asker-Árnason, L., Ibertsson, T., Wass, M., Wengelin, A., and Sahlén, B. (2010). Picture-elicited written narratives, process and product, in 18 children with cochlear implants. *Commun. Disord. Q.* 31, 195–212. doi: 10.1177/1525740109337734

Bäck, A. (2011). *Stavfelstyper. Ett antal Elevers Stavningsutveckling Från Årskurs 2 till Årskurs 4*, Master's thesis, Göteborg University, Göteborg.

Benson, C. (2002). Transfer/cross-linguistic influence. *ELT J.* 56, 68–70. doi: 10.1093/elt/56.1.68

Bergman, B. (1982). *Några Satstyper I Det Svenska Teckenspråket. Forskning Om Teckenspråk X.* Stockholm: Stockholm University.

Bergman, B. (2012). *Barns Tidiga Teckenspråksutveckling.* Stockholm: Stockholm University.

Bergman, B., and Wallin, L. (2001). "A preliminary analysis of visual mouth segments in Swedish sign language," in *The Hands are the Head of the Mouth: The Mouth as Articulator in Sign Language*, eds P. Boyes Braem, and R. Sutton-Spence, (Hamburg: Signum), 51–68.

Berman, R. A., and Slobin, D. I. (1994). *Relating Events in Narrative. A crosslinguistic developmental study.* Hillsdale, NJ: Erlbaum Associates, Inc.

Berninger, V. W., and Swanson, H. L. (1994). "modifying hayes and flower's model of skilled writing to explain beginning, developing writing," in *Children's Writing: Toward a Process Theory Of The Development In Skilled Writing*, ed. E. C. Butterfield, (Greenwich, CT: JAI Press Inc.), 57–81.

Bialystok, E. (2009). Bilingualism: the good, the bad, and the indifferent. *Bilingualism* 12, 3–11. doi: 10.1017/S1366728908003477

Bowers, L. M., Dostal, H., McCarthy, J. H., Schwarz, I., and Wolbers, K. (2016). An analysis of deaf students' spelling skills during a year-long instructional writing approach. *Commun. Disord. Q.* 37, 160–170. doi: 10.1177/1525740114567528

Brackenbury, T. (2005). Incidental word learning in a hearing child of deaf adults. *Oxford J.* 11, 76–93. doi: 10.1093/deafed/enj018

Brown, M., and Cormier, K. (2017). Sociolinguistic variation in the nativisation of BSL fingerspelling. *Open Linguist.* 3, 15–144. doi: 10.1515/opli-2017-0007

Chamberlain, C., and Mayberry, R. I. (2008). American sign language syntactic and narrative comprehension in skilled and less skilled readers: bilingual and bimodal evidence for the linguistic basis of reading. *Appl. Psycholinguist.* 29, 367–388. doi: 10.1017/S014271640808017X

Crasborn, O., Els van der Kooij, D. W., Woll, B., and Mesch, J. (2008). Frequency Distribution and spreading behavior of different types of mouth actions in three sign languages. *Sign Lang. Linguist.* 11, 45–67. doi: 10.1075/sl%26l.11.1.04cra

Davidson, K., Lillo-Martin, D., and Pichler, D. C. (2014). Spoken english language development among native signing children with cochlear implants. *J. Deaf Stud. Deaf Educ.* 19, 239–250. doi: 10.1093/deafed/ent045

Ejeman, G., and Molloy, G. (1997). *Svenska i Grundskolan – Metodboken.* Angered: Författarna och Liber AB.

Figueredo, L. (2006). Using the known to chart the unknown: a review of first-language influence on the development of English-as-a-second-language spelling skill. *Read. Writ.* 19, 873–905. doi: 10.1007/s11145-006-9014-1

Freel, B. L., Clark, M. D., Anderson, M. L., Gilbert, G. L., Musyoka, M. M., and Hauser, P. C. (2011). Deaf individuals' bilingual abilities: american sign

language proficiency, reading skills, and family characteristics. *Psychology* 2, 18–23. doi: 10.4236/psych.2011.21003

Frith, U. (1985). "Beneath the surface of surface dyslexia," in *Surface Dyslexia*, eds K. E. Patterson, J. C. Marshall, and M. Coltheart, (London: Erlbaum), 301–330. doi: 10.4324/9781315108346-18

Gärdenfors, M. (2016). *Så vill rätta fantastisk inte har den. Stavningen Hos Döva Och Hörselskadade Elever i Årskurs Fyra Och Fem.* Master's thesis, Uppsala University, Uppsala.

Geers, A. E., and Hayes, H. (2011). Reading, writing, and phonological processing skills of adolescents with 10 or more years of cochlear implant experience. *Ear Hear* 32, 49–59. doi: 10.1097/AUD.0b013e3181fa41fa

Glickman, N. S., and Hall, W. C. (2018). *Language Deprivation and Deaf Mental Health.* New York, NY: Routledge.

Hall, W. C. (2017). What you don't know can hurt you: the risk of language deprivation by impairing sign language development in deaf children. *Matern. Child Health J.* 21, 961–965. doi: 10.1007/s10995-017-2287-y

Haptonstall-Nykaza, T. S., and Schick, B. (2007). The transition from fingerspelling to english print: facilitating english decoding. *J. Deaf Stud. Deaf Educ.* 12, 172–183. doi: 10.1093/deafed/enm003

Harris, M., and Beech, J. R. (1998). Implicit phonological awareness and early reading development in prelingually deaf children. *J. Deaf Stud. Deaf Educ.* 3, 205–216. doi: 10.1093/oxfordjournals.deafed.a014351

Harris, M., and Terlektsi, E. (2011). Reading and spelling abilities of deaf adolescents with cochlear implants and hearing aids. *J. Deaf Stud. Deaf Educ.* 16, 24–34. doi: 10.1093/deafed/enq031

Hassanzadeh, S. (2012). Outcomes of cochlear implantation in deaf children of deaf parents: comparative study. *J. Laryngol. Otol.* 126, 989–994. doi: 10.1017/S0022215112001909

Herman, R. E., Kyle, F., and Roy, P. (2019). Literacy and phonological skills in oral deaf children and hearing children with a history of dyslexia. *Read. Res. Q.* 54, 553–575. doi: 10.1002/rrq.244

Hirshorn, E. A., Dye, M. W. G., Hauser, P., Supalla, T. R., and Bavelier, D. (2015). The contribution of phonological knowledge, memory, and language background to reading comprehension in deaf populations. *Front. Psychol.* 6:1153. doi: 10.3389/fpsyg.2015.01153

Hodge, G., and Johnston, T. (2014). Points, depictions, gestures and enactment: partly lexical and non-lexical signs as core elements of single clause-like units in Auslan (Australian Sign Language). *Austral. J. Linguist.* 34, 262–291. doi: 10.1080/07268602.2014.887408

Hoffmeister, R. J., and Caldwell-Harris, C. L. (2014). Acquiring English as a second language via print: the task for deaf children. *Cognition* 132, 229–242. doi: 10.1016/j.cognition.2014.03.014

Holmström, I. (2018). *Undervisning i Svenskt Teckenspråk Som Andraspråk. En Rapport Från UTL2-PROjektet.* Stockholm: Stockholm University.

Humphries, T., and MacDougall, F. (1999). "Chaining" and other links: making connections between american sign language and english in two types of school settings. *Vis. Anthropol. Rev.* 15, 84–94. doi: 10.1525/var.2000.15.2.84

Johansson, V. (2000). "Developing editing," in *Working Papers in Developing Literacy Across Genres, Modalities, and Languages*, Vol. 3, eds M. Aparici, N. Argerich, J. Perera, E. Rosado, and L. Tolchinsky, (Barcelona: University of Barcelona), 235–249.

Johnston, T., and Schembri, A. (2010). Variation, lexicalization and grammaticalization in signed languages. *Lang. Soc.* 131, 19–35. doi: 10.3917/ls.131.0019

Kelly, A. (1995). ""Fingerspelling interaction: a set of deaf parents and their deaf daughter"," in *Sociolinguistics in Deaf Communities*, ed. C. Lucas, (Washington DC: Gallaudet University Press), 191–210.

Kuntze, M., Golos, D., and Enns, C. (2014). Rethinking literacy: broadening opportunities for visual learners. *Sign Lang. Stud.* 14, 203–224. doi: 10.1353/sls.2014.0002

Larsson, Y. (2015). *Ålderseffekter Av En Kritisk Period för Språkutveckling. Grammatisk Intuition Hos Hörande Infödda.* Göteborg: Göteborg University. master's thesis.

Leijten, M., and Van Waes, L. (2013). Keystroke logging in writing research: using Inputlog to analyze and visualize writing processes. *Writ. Commun.* 30, 358–392. doi: 10.1177/0741088313491692

Lepic, R. (2015). *Motivation in Morphology: Lexical Patterns in ASL and English.* Ph.D. thesis, University of California, San Diego.

Marschark, M. (2009). *Raising and Educating a Deaf Child. A Comprehensive Guide to the Choices, Controversies, and Decisions Faced by Parents and Educators.* Oxford: Oxford University Press.

Mayberry, R. I., Del Giudice, A. A., and Lieberman, A. M. (2011). Reading achievement in relation to phonological coding and awareness in deaf readers: a meta-analysis. *J. Deaf Stud. Deaf Educ.* 16, 164–188. doi: 10.1093/deafed/enq049

Mitchell, R. E., and Karchmer, M. A. (2004). Chasing the mythical ten percent: parental hearing status of deaf and hard of hearing students in the United States. *Sign Lang. Stud.* 4, 138–163. doi: 10.1353/sls.2004.0005

Nation, K., and Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Read. Res. Q.* 32, 154–167. doi: 10.1598/rrq.32.2.2

Nauclér, K. (1980). *Perspectives on misspellings. A Phonetic, Phonological and Psycholinguistic Study.* Lund: Lund University. Ph.D. thesis.

Nauclér, K. (1989). "Hur utvecklas stavningsförmågan under skoltiden?," in *Språkutveckling Under Skoltiden,* eds C. Sandqvist, and U. Teleman, (Lund: Studentlitteratur), 197–237.

Padden, C. (2005). "Learning to fingerspell twice: young signing children's acquisition of fingerspelling," in *Advances in the Sign Language Development of Deaf Children,* eds B. Schick, M. Marschark, and P. E. Spencer, (Oxford: University Press), 189–201. doi: 10.1093/acprof:oso/9780195180947.003.0008

Padden, C., and Ramsey, C. (2000). "American sign language and reading ability in deaf children," in *Language Acquisition by Eye,* eds C. Chamberlain, J. P. Morford, and R. I. Mayberry, (New Jersey: Lawrence Erlbaum), 165–189.

Raatikainen, R. (2018). *Felstavningar. En Studie Av Typiska Stavfel I Svenska Elevers texter i årskurs 3.* Uppsala: Uppsala University. master's thesis.

Schönström, K. (2010). *Tvåspråkighet Hos Döva Skolelever. Processbarhet i Svenska Och Narrativ Struktur I Svenska Och Svenskt Teckenspråk.* Stockholm: Stockholm University.

Schönström, K. (2014). Visual acquisition of Swedish in deaf children: an L2 processability approach. *Linguist. Approac. Bilingual.* 4, 61–88. doi: 10.1075/lab.4.1.03sch

Schönström, K., and Holmström, I. (2017). "Elicited imitation tasks (EITs) as a tool for measuring sign language proficiency in L1 and L2 signers," in *Proceedings of the Paper presentation at ALTE 6th Conference, Learning and Assessment: Making the Connections,* Bologna.

Singleton, J. L. (2004). Vocabulary use by low, moderate, and high ASL-proficient writers compared to hearing ESL and monolingual speakers. *J. Deaf Stud. Deaf Educ.* 9, 86–103. doi: 10.1093/deafed/enh011

Singleton, J. L., and Newport, E. L. (2004). When learners surpass their models: the acquisition of american sign language from inconsistent input. *Cognit. Psychol.* 49, 370–407. doi: 10.1016/j.cogpsych.2004.05.001

Skolverket (2018). *Curriculum for the Compulsory School, Preschool Class And School-Age Educare.* Available at: https://www.skolverket.se/getFile?file=3984 (accessed March 15, 2019).

SOU (2006). *SOU (2006:29) Statens offentliga utredningar. Teckenspråk Och Teckenspråkiga. Kunskaps- Och Forskningsöversikt.* Available at: https://www.regeringen.se/contentassets/f67e90bf3e3d4fa58dd13abe3fb46e66/teckensprak-och-teckensprakiga-kunskaps--och-forskningsoversikt-sou-200629 (accessed April 30, 2019).

SOU (2016). *SOU (2016:46) Statens Offentliga Utredningar. Samordning, Ansvar Och Kommunikation – väGen Till ÖKad Kvalitet i Utbildningen FöR Elever Med Vissa Funktionsnedsättningar.* Available at: https://www.regeringen.se/contentassets/512c9cff5ffc24a8fb9371d3ab538edaa/samordning-ansvar-och-kommunikation--vagen-till-okad-kvalitet-i-utbildningen-for-elever-med-vissa-funktionsnedsattningar-hela-dokumentet-sou-201646.pdf (accessed April 30, 2019).

Sterne, A., and Goswami, U. (2000). Phonological awareness of syllables, rhymes, and phonemes in deaf children. *J. Child Psychol. Psychiatr. Allied Discipl.* 41, 609–625. doi: 10.1017/S0021963000005680

Straley, S. G., Werfel, K. L., and Hendricks, A. E. (2016). Spelling in written stories by school-age children with cochlear implants. *Deaf. Educ. Int.* 18, 67–70. doi: 10.1080/14643154.2016.1143168

Strong, M., and Prinz, P. M. (1997). A study of the relationship between american sign language and english literacy. *J. Deaf Stud. Deaf Educ.* 2, 37–46. doi: 10.1093/oxfordjournals.deafed.a014308

Sutcliffe, A., Dowker, A., and Campbell, R. (1999). Deaf children's spelling: does it show sensitivity to phonology? *J. Deaf Stud. Deaf Educ.* 4, 111–123. doi: 10.1093/deafed/4.2.111

Svartholm, K. (2006). *Svenska Som Andraspråk För Döva – En Ämnesöversikt.* Stockholm: Stockholm University.

Svartholm, K. (2008). "The written swedish of deaf children: a foundation for EFL," in *English in International Deaf Communication. Linguistic insights,* eds C. J. K. Bidoli, and E. Ochse, (Bern: Peter Lang), 211–250.

Svartholm, K. (2010). Bilingual education for deaf children in Sweden. *Int. J. Bilin. Educ. Bilingual.* 13, 159–174. doi: 10.1080/13670050903474077

Svensk Teckenspråkskorpus (2019). Available at: https://www.ling.su.se/teckenspr%C3%A5ksresurser/teckenspr%C3%A5kskorpusar/svensk-teckenspr%C3%A5kskorpus (accessed April 23, 2019).

Svensson, A.-K. (1998). *Barnet, Språket Och Miljön.* Lund: Studentlitteratur.

Van Beijsterveldt, L. M., and van Hell, J. (2010). Lexical noun phrases in texts written by deaf children and adults with different proficiency levels in sign language. *Int. J. Biling. Educ. Biling.* 13, 439–468. doi: 10.1080/13670050903477039

Verhoeven, L. (2000). Components in early second language reading and spelling. *Sci. Stud. Read.* 4, 313–330. doi: 10.1207/S1532799XSSR0404-4

von Koss, Torkildsen, J., Morken, F., Helland, W. A., and Helland, T. (2016). The dynamics of narrative writing in primary grade children: writing process factors predict story quality. *Read. Writ.* 29, 529–554. doi: 10.1007/s11145-015-9618-4

Wengelin, Å. (2006). "Examining pauses in writing: theory, methods and empirical data," in *Computer Keystroke Logging and Writing,* eds K. P. H. Sullivan, and E. Lindgren, (Amsterdam: Elsevier), 107–130.

Wengelin, Å (2002). *Text Production in Adults with Reading and Writing Difficulties.* Gothenburg: Gothenburg University.

Wengelin, Å, Frid, J., Johansson, R., and Johansson, V. (2019). "Combining keystroke logging with other methods. Towards an experimental environment for writing process research," in *Observing Writing: Insights From Keystroke Logging And Handwriting,* eds E. Lindgren, and K. P. H. Sullivan, (Netherlands: Brill Publishing), 30–49.

Cluster Analyses Reveals Subgroups of Children With Suspected Auditory Processing Disorders

Mridula Sharma[1,2]*, Suzanne C. Purdy[3,4] and Peter Humburg[5]

[1] Department of Linguistics, Australian Hearing Hub, Macquarie University, Macquarie Park, NSW, Australia, [2] The HEARing CRC, Audiology, Hearing and Speech Sciences, The University of Melbourne, Parkville, VIC, Australia, [3] Speech Science, School of Psychology, The University of Auckland, Auckland, New Zealand, [4] Eisdell Moore Centre for Hearing and Balance Research, The University of Auckland, Auckland, New Zealand, [5] Faculty of Human Sciences, Macquarie University, Macquarie Park, NSW, Australia

Background: Some children appear to not hear well in class despite normal hearing sensitivity. These children may be referred for auditory processing disorder (APD) assessment but can also have attention, language, and/or reading disorders. Despite presenting with similar concerns regarding hearing difficulties in difficult listening conditions, the overall profile of deficits can vary in children with suspected or confirmed APD. The current study used cluster analysis to determine whether subprofiles of difficulties could be identified within a cohort of children presenting for auditory processing assessment.

Methods: Ninety school-aged children (7–13 years old) with suspected APDs were included in a cluster analysis. All children had their reading, language, cognition and auditory processing assessed. Parents also completed the Children's Auditory Performance Scale (CHAPS). Cluster analysis was based on tasks where age-norms were available, including word reading (Castles and Coltheart irregular and non-words test), phonological awareness (Queensland University Inventory of Literacy), language [Comprehensive Language of Assessment-4, Comprehensive Assessment of Spoken Language (CASL)], sustained attention (Continuous Performance Test), working memory (digits forward and backward), and auditory processing [Frequency Pattern Test (FPT), Dichotic Digits Test (DDT)]. Hierarchical cluster analysis was undertaken to determine the optimal number of clusters for the data, followed by a k-means cluster analysis.

Results: Hierarchical cluster analysis suggested a four-group solution. The four subgroups can be summarized as follows: children with (1) global deficits, $n = 35$; (2) poor auditory processing with good word reading and phonological awareness skills, $n = 22$; (3) poor auditory processing with poor attention and memory but good language skills, $n = 15$; and (4) poor auditory processing and attention with good memory skills, $n = 18$.

Conclusion: The cluster analysis identified distinct subgroups of children. These subgroups display the variation in areas of difficulty observed across different studies in the literature (e.g., not every child with APD has an attention deficit), highlighting

the heterogeneous nature of APD and the need to assess a range of skills in children with suspected APD. It would be valuable for future studies to independently verify these subgroups and to determine whether interventions can be optimized based on these subgroups.

INTRODUCTION

Some school-aged children appear to not hear well in difficult listening situations such as the classroom, in the absence of a hearing loss based on pure tone audiometry (Purdy et al., 2018). These children are often described as having problems hearing in noise, needing to have instructions repeated, being unable to follow verbal instructions and having generally poor listening skills (Chermak et al., 2002). Some of these children also show co-existing reading difficulties and/or attention deficits (Richardson et al., 2004; Sharma et al., 2009; Tomlin et al., 2015). These children are initially tested for hearing loss and in the absence of any audiometric hearing loss they should be referred for auditory processing assessment (Jerger and Musiek, 2000). Clinical practice varies widely, however, despite considerable efforts internationally to develop auditory processing assessment and treatment guidelines (Iliadou et al., 2018).

In children diagnosed with APD, there is impaired processing of auditory information that is not consistent with their hearing thresholds (Moore et al., 2013). Auditory processing includes the ability of the auditory system to localize, discriminate, recognize auditory patterns, and discriminate temporal aspects of sounds (including but not limited to temporal resolution, masking, integration and sequencing) (American Speech and Hearing Association [ASHA], 1996; Jerger and Musiek, 2000). A significant deficit in any of these auditory skills is indicative of APD (American Speech and Hearing Association [ASHA], 1996). Thus to diagnose any child with APD, many established guidelines (American Speech and Hearing Association [ASHA], 1996; Jerger and Musiek, 2000; Wilson, 2018) recommend a test battery that evaluates multiple auditory processing skills.

Clinicians working with children with suspected APD face three important challenges. One is that auditory processing is not a unitary skill and therefore cannot be assessed with one test (Jerger and Musiek, 2000; Wilson, 2018), hence clinicians need to access a range of tests that have age-dependent norms and demonstrated reliability, test efficiency and validity (Musiek et al., 2010; Emanuel et al., 2011; Wilson, 2018; Keith et al., 2019). For example, commonly used tests such as the FPT (Musiek, 1994) and the DDT (Musiek, 1983) have age-related norms while the Random Gap Detection Test (RGDT) has a screening pass level that is applied to all school aged children (Sharma et al., 2006; Kelly, 2007). A second challenge is that children with APD can have co-existing language, attention, and/or reading disorders (Sharma et al., 2009; Wilson, 2018) that may affect test results and/or management choices. A third challenge stems

from the need for efficient, clinically feasible diagnostic protocols that can capture APD in children who are heterogeneous and that assist the children and their families in receiving appropriate management that includes appropriate evidence-based treatments (Wilson, 2018). The current research attempts to address the second and third challenges by determining whether there are identifiable subprofiles of children who are suspected to have APD with other potential co-existing disorders, since such subprofiles may help guide management. The research aim is to determine whether cluster analysis identifies distinct subgroups of children, which would help researchers and clinicians to better understand the range of challenges that children with APD present with, and could guide recommendations to parents and clinicians regarding appropriate clinical referral pathways. There have been attempts to define subgroups of children with APD in the past (Bellis and Ferre, 1999), recognizing the potential values of this approach for clarifying referral pathways and planning treatment, but to our knowledge the current study is the first that uses cluster analysis to define subgroups.

There is ample evidence that children with auditory processing deficits can display reading and language deficits, but typically this is not the case for all children (Jerger and Musiek, 2000; Ramus, 2003; Bishop, 2007; Sharma et al., 2009; Leppänen et al., 2010; Hämäläinen et al., 2013; Halliday et al., 2017; Mealings and Cameron, 2019). A causal relationship between auditory deficits and poor reading and/or language skills has been proposed, or at least it has been suggested that these share some common underlying neurodevelopmental etiology (Leppänen et al., 2010; Moore et al., 2013; Halliday et al., 2017). This is difficult to prove, however, and there is no empirical evidence that confirms this. A theoretical framework has been proposed (Ramus, 2003; Goswami, 2011; Halliday et al., 2017) that attempts to explain why auditory processing and reading disorders are associated (Sharma et al., 2006; Leppänen et al., 2010; Hämäläinen et al., 2013) but there is no agreement on the "nature or magnitude of the link" between auditory processing and reading disorders (Ramus, 2003). There are also reports that children with auditory processing deficits have cognitive (attention and/or working memory) difficulties that account, at least in part, for their poor performance on auditory processing tests (Moore et al., 2013). This is also not straightforward as some children with APD do not have attention and memory deficits (Sharma et al., 2009, 2014a; Tomlin et al., 2015).

Links between auditory, cognitive, reading, and language abilities of children with APD are still not fully understood. It is recommended that children with suspected APD are assessed using a wide range of measures that encompass all these domains (American Speech Language Hearing Association [ASHA], 2005).

Abbreviations: APD, auditory processing disorder; TONI, test of nonverbal intelligence; PA, phonological awareness; FPT, Frequency Pattern Test; DDT, Dichotic Digits Test.

In the current study auditory, cognitive, reading, and language abilities of children with suspected APD were assessed and cluster analysis was used to determine whether the results revealed distinct subgroups of children. The subgrouping was then tested by comparing the groups across a range of related measures not included in the cluster analysis to determine where there were significant differences in performance.

METHODOLOGY

Participants

The University of Auckland Human Research Participants' Ethics Committee approved this study. Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin. Ninety children aged 7–12.8 years old (*Mean* = 9.8 years ± 1.5) with listening concerns participated: 58 males with an average age of 9.8 years ± 1.6 and 32 females with an average age of 9.7 years ± 1.5. Children were referred to the study by speech language pathologists, teachers, educational psychologists, and audiologists. Most came to the research with suspected APD and reports of other reading and language concerns, making this a potentially heterogeneous group of participants. A subset of these children was reported on previously (Sharma et al., 2009, 2019; Gilley et al., 2016).

Methods

Children were tested individually in a sound-treated laboratory booth over two sessions of about 3 h each with multiple breaks. Pure-tone audiometry and behavioral auditory processing tests were administered using a GSI clinical audiometer and TDH-39 earphones. Test materials were presented at 60 dB HL using a CD player (Bass XPander, P882). All children were administered hearing, auditory processing (behavioral and electrophysiological), language, cognitive, and reading assessments.

Inclusion criteria included normal peripheral hearing and a standard score of 80 or more on the TONI (Brown et al., 1990). Parents were invited to report on their children's perceived listening difficulties by completing the Children's Auditory Performance Scale (CHAPS) questionnaire (Smoski et al., 1998), which rates the children's difficulties compared to classroom peers (a score a "0" indicates equivalent performance to peers) (Sharma et al., 2009). Smoski et al. (1998) proposed a normative cut-off of −11 for the overall CHAPS score, with scores lower than this indicating significant listening difficulties. In total 83 parents (92% of participants) returned the CHAPS questionnaire.

All participants had normal hearing sensitivity. Pure tone thresholds were 15 dB HL or better at octave frequencies from 250 to 8000 Hz. All children had Type A tympanograms, measured using a 226-Hz probe tone (Jerger, 1970) and ipsilateral 1000-Hz acoustic reflex thresholds less than 100 dB HL (Silman and Gelfand, 1981) consistent with normal middle ear function. For all children otoacoustic emissions (OAE) strength was within the normal range based on the pass/refer criteria in the TEOAE protocols of the Scout Sport System (Bio-Logic Systems Corp®) (Hall, 2000).

Children were evaluated on multiple measures after completing the peripheral hearing assessments. The tasks included in the cluster analysis are the ones where published age-specific norms were available. Details of the stimuli, procedure and scoring are provided in **Table 1**.

The auditory processing measures were the FPT (Musiek, 1994) and DDT (Musiek, 1983). Cognitive measures were memory (Comprehensive Evaluation of Language Fundamentals, CELF-4, digit span forward and backward) (Semel et al., 2003) and sustained attention [Integrated Visual and Auditory (IVA) Continuous Performance Test] (Sandford and Turner, 2000). Language measures were Receptive and Core Language standard scores from the CELF-4, and Auditory Comprehension and Supralinguistics standard scores from the Comprehensive Assessment of Spoken Language (CASL), as these all rely on auditory perception (Carrow-Woolfolk, 1999). The reading task included was word reading measured using the Castle and Coltheart's word lists (Castles and Coltheart, 1993). Phonological processing was measured using the Queensland Inventory of Literacy (QUIL) (Dodd et al., 1996). Only those QUIL tasks specifically linked to auditory perception were included (syllable identification, segmentation, rhyming, spoonerisms, phoneme detection, phoneme manipulation). Non-word spelling and visual rhyme subtests were not included in the cluster analysis. All analyses were undertaken using Statistica 10.0.

Data Reduction: Correlation and Factor Analysis

The entire dataset included 23 variables. Pearson correlation analyses were undertaken to remove highly correlated variables. This is important step as strongly correlated variables represent potentially the same measure and may receive higher weighting during cluster analysis. Both correlation and factor analysis were undertaken to avoid this (Chiarello et al., 2012). Variables with strong correlations ($r \geq 0.70$) were not placed in the cluster analysis (Taylor, 1990). Following the correlational analysis, exploratory factor analysis was used to further reduce the number of variables.

Cluster Analysis

Before undertaking the cluster analysis, the selected variables were standardized to control for unequal scaling of the data (Clatworthy et al., 2005). The standardization transforms all values (regardless of their distributions and original units of measurement) to compatible units from a distribution with a mean of 0 and a standard deviation of 1. This transformation makes the distributions of values easy to compare across variables and independent of the units of measurements. A hierarchical cluster analysis using Ward's method was performed on the data to determine how many clusters are appropriate for the final selected variables (Clatworthy et al., 2005; Chiarello et al., 2012). Following this, a k-means cluster analysis was performed on the data to determine the membership of the individual cases into the clusters. Once group membership was determined, a discriminant function analysis was undertaken to confirm predicted membership.

TABLE 1 | The details of the all tests that included auditory processing, reading, language, attention, memory in the current study.

	Tests	Procedure
Auditory processing	Frequency Pattern Test (FPT)	*Stimuli*: triplet tones presented in a sequence of high and low frequencies. The high frequency had a frequency of 1100 Hz, and the low tone had a frequency of 880 Hz (Musiek, 1994). The tones were presented through an Audiometer that received the input from a CD player. The tones were 150 ms long with an interstimulus interval of 80 ms. There were six possible sequences, HLL, HLH, HHL, LHH, LHL, LLH.
		Procedure: the participants had to verbally identify the sequence.
		Response and scoring: the verbal response of the participants were marked, with each correct response yielding 1 point. The score out of 15 was calculated to a percentage score and converted to *z* scores.
	Dichotic Digit Test (DDT)	*Stimuli*: in this task two pairs of double digits (four digits) are presented to the two ears simultaneously and the listener has to verbally report back the four digits.
		Procedure: the participants repeat the numbers they hear irrespective of the order (free recall). Twenty digits were presented in each ear.
		Response and scoring: each correctly identified digit is allocated a score of 1 and percentage correct for each ear is calculated out of maximum presented digits (20 for each ear) and converted to *z* scores.
	RGDT	*Stimuli*: in this task two clicks are presented to both ears with varying gap and the listener has to verbally report back whether they heard 1 click (no gap) or 2 clicks (gaps of varying duration).
		Procedure: the participants state they hear 1 or 2 clicks. The gaps include 0, 2, 5, 10, 15, and 20 ms.
		Response and scoring: the shortest gap identified is noted as the threshold.
	MLD	*Stimuli*: 500-Hz tone masking level difference task is a measure of the improvement in detectability of a tone when presented in noise that is phase-shifted between the two ears.
		Procedure: the listener has to identify the presence of tone within noise. Detection of the tone improves for the phase-shifted binaural listening condition compared to the in-phase condition.
		Response and scoring: the MLD is the 500-Hz threshold difference when the $S_\Pi N_0$ (tone presented to the two ears in opposite phase, noise presented in the same phase) and S_0N_0 (tone and noise presented to the two ears in the same phase) conditions are compared.
	AB words	*Stimuli*: in this task, monosyllabic words (Consonant-vowel-consonant) are presented to each ear and the listener has to verbally report back each word.
		Procedure: the participants repeat the words they hear. Ten words were presented in each ear in quiet and modified by 65% compression and 0.3s reverberation.
		Response and scoring: each correctly identified word is allocated 10% and each consonant within each word is allocated a score of 3%. Scores were separately determined for quiet and modified for right and left ear.
	CAEPs	*Stimuli*: /da/ natural token 158 ms long stimulus spoken by Australian female, are presented to both ears via insert earphones in quiet at 70 dB SPL and +3 dB SNR.
		Procedure: the listener does not need to respond and was asked to watch a muted movie of their choice. CAEPs were collected from the vertex (Cz) and frontal (Fz) non-inverting electrodes, with right and left mastoid as inverting electrode.
		Response and scoring: P1-N250 peaks were identified. P1 is a biggest positive peak at about 100 ms after the onset and N250 is the biggest negative peak that occurs at about 250 ms after the onset of the sound. Amplitude and latency for P1-N250 were identified.
Reading	Castles and Coltheart	*Stimuli*: the Castles' Word/Non-word Test (Castles and Coltheart, 1993) was used to measure word reading for regular words (pronounced in accordance with letter-sound rules, e.g., plant), irregular words (pronunciation violates letter-sound rules, e.g., yacht), and non-words (e.g., phot).
		Procedure: the words from the three lists are presented in a random order.
		Response and scoring: the participant receives a score of 1 for each word read correctly. There were 30 words in each list.
Phonological processing	Queensland Inventory of Literacy (QUIL)	*Stimuli*: seven QUIL subtasks were utilized for testing phonological processing; non-word spelling, syllable identification and segmentation, rhyming, spoonerisms, phoneme detection; phoneme manipulation.
		Procedure: non-word spelling; syllable identification (to determine the number of syllables in a given word); syllable segmentation; rhyming (judgment of auditory word pairs, such as "shell and bell," or "bout and bait"); spoonerisms (swapping the first sounds from a pair of spoken words to make new words such as "felt and make" swapped to make "melt and fake"); phoneme detection (identifying the odd word out, which may be different because of first/second or third sound); and phoneme segmentation and manipulation (identifying and removing a sound from a given word and then saying the new word; e.g., "frog" without *r* is "fog").
		Response and scoring: the participant receives a score of 1 for each item within each task completed correctly and converted to scaled scores.
Non-verbal intelligence	TONI	*Stimulus*: the TONI-3 measures a child's reasoning ability with minimal language influence. It is a norm-referenced measure of cognitive ability that assesses intelligence aptitude, abstract thinking, and problem solving for 6–89 year olds.
		Procedure: for each item, the child selected one of six options to complete a matrix pattern that incorporated one or more features, such as shape, direction, position, size, and movement.
		Response and scoring: the correct identification of pattern led to total scores that are then converted to standard scores.

(Continued)

	Tests	Procedure
Memory	Digit Span Test	*Stimulus*: auditory memory was assessed using the CELF-4 forward and backward digits tasks.
		Procedure: the participant recalls a series of numbers in either forward or reverse order immediately after hearing them. The numbers are spoken slowly in a monotone voice and the child is asked to repeat them. The length of the series keeps increasing until the child can no longer repeat the series correctly in the appropriate sequence. Each sequence is presented in pairs.
		Response and scoring: score of 1 is given for every correct sequence till the participant made an error for a pair of sequence. The score is determined independently for forward sequences and reverse order sequences. These were then converted to scaled scores.
Sustained attention	Visual and Auditory Continuous Performance Test	*Stimulus*: children's attention was tested using the IVA (Integrated Visual and Auditory Continuous Performance Test) presented on a laptop computer, which assesses performance in a combined visual and auditory task (Sandford and Turner, 1995).
		Procedure: numbers 1 and 2 are seen and heard in pseudorandom order. The task includes 500 trials and takes approximately 15 min to complete. Feedback is provided only during practice trials. The child was instructed to click the mouse whenever the number "1" was seen or heard, and to ignore the number "2".
		Response and scoring: the task is undertaken for about 15 min and based on the number of correctly identified responses in both auditory and visual modality, the sustained attention standard score is calculated independently for auditory and visual modality.
Language	CELF-4	*Stimulus*: the Clinical Evaluation of Language Fundamentals 4th edition (CELF-4; Semel et al., 2003) was used to assess children's receptive and core language scores.
		Procedure: the Receptive Language score is a cumulative score from the subtests Concepts and Following Directions, Receptive Word Classes (semantic relationships), and Sentence Structure. Expressive Language is a cumulative score from subtests such as recalling sentences, formulated sentences and word classes-expressive.
		Response and scoring: the correct responses in all subtests are added and converted to scaled scores.
	CASL	*Stimuli*: comprehensive Assessment of Spoken Language (CASL) is an orally presented language assessment battery for ages 3–21 years (Carrow-Woolfolk, 1999).
		Procedure: two tasks were included Auditory Comprehension and Supralinguistics. Auditory Comprehension included two subtests paragraph comprehension and synonym while Supralinguistics included subtests non-literal (explaining sentences) and inference drawing (determining meaning).
		Response and scoring: the correct responses in all subtests are added and converted to scaled scores.

Inferential Statistics

The stability of the clustering was determined by comparing groups on the variables that were not used in the cluster analysis to evaluate generalizability of the clusters (Chiarello et al., 2012). The group comparisons included gender distribution, paragraph reading [Wheldall Accuracy of Reading Passages (WARP)] (Madelaine and Wheldall, 1998), non-word spelling (QUIL subtest), CELF-4 Expressive language (Semel et al., 2003), RGDT (Keith, 2000), Masking Level Differences (MLD) (Sweetow and Reddell, 1978; Jerger et al., 1984), word recognition scores (AB words in quiet and in noise with 65% compression and 0.3s reverberation of words) (Sharma et al., 2009), and speech-evoked cortical auditory evoked potential (CAEP) latencies and amplitudes. The procedure for recording CAEPs to /da/ in quiet and in noise (at 3 dB signal-to-noise ratio, SNR) is described elsewhere (Sharma et al., 2014b). All comparisons were performed with age as a covariate and results were adjusted for multiple testing using Bonferroni correction.

RESULTS

Data Reduction: Correlation and Factor Analysis

FPT and DDT scores for right and left ears were significantly correlated ($r = 0.86$, $p < 0.001$ and $r = 0.070$, $p < 0.001$, respectively) and therefore, only FPT and DDT right ear scores were included in the cluster analysis. Castle and Coltheart regular word and irregular word scores were also highly correlated ($r = 0.78$, $p < 0.001$) so only irregular words were included. Non-word scores on the Castle and Coltheart test and QUIL were correlated ($r = 0.72$, $p < 0.001$); the Castle and Coltheart non-word task was included in the cluster analysis and the QUIL subtests were examined separately.

Tables 2A,B provide Pearson's correlational results for QUIL and language tasks respectively. Scores for the QUIL subtests were weakly or modestly correlated with each other (r values in the range 0.28–0.55). All QUIL measures were therefore included in the next stage of data reduction using factor analysis (Taylor, 1990) (**Table 2A**).

Performance on the memory tasks (digit span forward and backward) were significantly correlated but the correlation was weak ($p = 0.001$, $r = 0.35$); both measures were included in the cluster analysis. Auditory and visual sustained attention were strongly correlated ($p < 0.001$, $r = 0.74$) and therefore, only auditory attention scores were included. Language scores were not highly correlated (r values 0.52–0.69) (**Table 2B**), therefore, all languages scores (Receptive, Core, Auditory Comprehension, Supralinguistics) were included in the next stage of data reduction using factor analysis.

Of the now 18 tasks included based on the correlation analysis, there were six measures of phonological processing and four language measures. From the six measures of phonological

(A)

		1	2	3	4	5	6	7
QUIL	Non-word reading							
Syllable identification	0.26^{ns}							
Syllable Segmentation	0.36^{**}	0.19^{ns}						
Spoken rhyme	0.55^{**}	0.28^{*}	0.16^{ns}					
Spoonerism	0.46^{**}	0.43^{**}	0.43^{*}	0.37^{**}				
Phoneme detection	0.44^{**}	0.33^{*}	0.44^{**}	0.34^{*}	0.48^{**}			
Phoneme segmentation	0.24^{ns}	0.21^{ns}	0.15^{ns}	0.25^{ns}	0.22^{ns}	0.21^{ns}		
Phoneme manipulation	0.50^{**}	0.34^{*}	0.34^{*}	0.49^{**}	0.52^{**}	0.50^{**}	0.26^{ns}	

(B)

	Core	Auditory Comprehension	Supralinguistics
CELF: Receptive			
CELF: Core	0.69^{**}		
CASL: Auditory Comprehension	0.52^{**}	0.52^{**}	
CASL: Supralinguistics	0.54^{**}	0.63^{**}	0.52^{**}

(A) Word reading and phonological awareness tasks as measured on QUIL across all children (N = 90). (B) Language subtasks as measured on CELF-4 and CASL across all children (N = 90). $**p < 0.0001$, $*p < 0.001$, $ns = p > 0.01$.

processing, (unrotated) principal component analysis identified only one factor with an eigenvalue greater than 1, explaining 48.6% of the variance. This included all six items with loadings greater than 0.50 (**Table 3**). Individual principal component scores were therefore included in the cluster analysis. Similarly, when the unrotated principal component analysis was undertaken for the four language measures, only one principal component was extracted that explained 68.0% of the variance. All four items within the component had loadings greater than 0.77 and hence the principal component scores were used for the cluster analysis.

Classification Analyses: Predictors of Cluster Membership

For the next stage of analysis, the now remaining 10 measures were standardized. The tasks included were auditory processing (FPT, DDT), reading (irregular, non-word), language (one principal component derived from Receptive, Core, Auditory Comprehension, Supralinguistics), TONI, phonological processing (one principal component derived from syllable identification and segmentation, spoken rhyme, spoonerism, phoneme detection, and manipulation), sustained auditory attention, and both memory measures (digit span forward and backward). All the values for the factors were within two standard deviations of the mean and therefore, no outliers were identified.

The hierarchical cluster analysis, as seen in **Figure 1**, suggested a four-cluster solution appropriate for the final 10 selected variables. Using the plot of linkage distances, **Figure 2** shows a plateau, thus a large number of clusters were at the same linkage distance. A four-cluster solution was determined at a point where the plateau ended between linkage distances of 10–15. The final grouping of cases into four clusters was determined after three iterations of the k-means algorithm, using equally spaced centers.

Figure 3 shows the means of the 10 standardized variables in each of the four clusters.

The discriminant analysis and the cluster analysis showed very similar membership of the cases. The accuracy was 97% for group one, 95% for group 2, 93% for group 3, and 89% for group 4. Box's M test ($p = 0.134$) was not significant, indicating that the assumption of homoscedasticity is justified. A significant Wilks lambda ($\Lambda = 0.072$, $p < 0.001$) shows a good difference in the mean scores between the four clusters. **Table 4** shows the demographics and performance on auditory processing, reading, language and cognitive skills of children across the four clusters.

Four variables that provided a three-function solution had much higher discriminant function coefficients and hence were more relevant in determining cluster membership. These

TABLE 3 | Factor loadings and communalities based on a (unrotated) principal components analysis for six items from the **(A)** QUIL subtests and **(B)** language measures from CELF-IV and CASL (N = 90).

Items	Factor 1*
(A)	
Syllable identification	0.60
Syllable segmentation	0.60
Spoken rhyme	0.62
Spoonerism	0.79
Phoneme detection	0.75
Phoneme Manipulation	0.78
(B)	
CELF: Receptive	0.84
CELF: Core	0.87
CASL: Auditory Comprehension	0.77
CASL: Supralinguistics	0.82

The factor scores were multiplied by -1 to facilitate interpretation of results.

variables were phonological processing, digit span backward (function 1), TONI (function 2), and non-word reading (with backward digit, function 3) ($\Lambda = 0.1$, $p \leq 0.004$ for all) (**Supplementary Table 1**). **Figures 4A,B** shows the scatterplot of cases for the four clusters across the three functions.

Inferential Statistics

Table 5 shows how the four clusters differ on the additional auditory processing and language tasks that were not included in the cluster analysis. The four clusters of children did not differ significantly based on any of the auditory processing measures including CAEPs (**Table 5**). There were significant performance differences between the four clusters, however, for several reading, phonological, and language measures (WARP paragraph reading, non-word spelling, visual rhyme, expressive language) (**Table 5**).

Interpretation of Clusters

Four clusters emerged based on ten tasks included in the cluster analysis. To determine differences between clusters, each skill was scaled against the mean to determine the proportion of children with relatively poor results for the different areas included in the cluster analysis.

Cluster 1 included 35 children who showed overall poor scores on reading, language, and cognitive measures. Dichotic scores were also impacted relatively more in this group compared to other clusters. This cluster of children appear to have global deficits across all domains. All children had scores more than 1 SD below the overall mean ($N = 90$) for more than one measure (**Tables 4A,B**). One quarter of the children in this cluster had TONI standard scores of 80–85 (a standard score of 80 was the lower limit for study inclusion), and 63% had scores more than 1 SD below the mean for sustained auditory attention. About

FIGURE 1 | Dendrogram based on Wald's minimum-variance hierarchical clustering method. The 90 participants were clustered into a single final group. At each generation of clusters, samples were merged into larger clusters to minimize the within cluster sum of squares.

half of the Cluster 1 parents (51%) reported that their children had significant listening difficulties based on the CHAPS scoring criterion proposed by Smoski et al. (1998) (overall score < −11). Children within this cluster also showed performance 2 SD below the mean on FPT ($n = 12$, 34%), DDT ($n = 3$, 9%) or both ($n = 19$, 54%). Pearson's partial correlations within the cluster exploring associations between cognitive, reading, and language skills, and DDT and FPT auditory processing measures (with age as covariate) showed no significant associations (with Bonferroni adjustments).

Cluster 2 included 22 children with good reading and good phonological processing skills. Only one child had a TONI score more than 1 SD below the mean and this child had high reading and phonological skills. Another child had a score more than 2 SD below the mean for the digit span backward test but had average language, reading, and phonological processing scores. Auditory processing skills measured using the FPT and DDT showed that 27% of this group only had FPT scores that were more than 2 SD below the mean, 18% only had DDT scores more than 2 SD below the mean, while 27% had poor performance on both the FPT and DDT. This cluster includes children with auditory processing difficulties in the presence of relatively good reading and phonological processing skills and, like Cluster 1, the CHAPS showed that half of the children in this cluster had parent-reported listening difficulties (overall score < −11) (Smoski et al., 1998). This cluster showed a moderate and significant partial correlation (age as covariate) between non-word reading and paragraph reading ($r = 0.58$, $p = 0.048$).

Cluster 3 included 15 children with relatively high non-verbal IQ scores, and good phonological processing and word reading. Thirty percent of children in this cluster had FPT scores that were more than 2 SD below the mean, while 13% only had DDT scores that were more than 2 SD below the mean, and 40% of children had scores 2 SD below the mean for both FPT and DDT. Four children with scores more than 2 SD below the mean for FPT and DDT also showed scores 1 SD below the mean on sustained auditory attention and working memory (backward digit span) tasks. In general, this cluster had good TONI and language skills with poor auditory processing and poor attention and memory and about 27% of parents (3/11 who completed the questionnaire) reported listening difficulties based on the CHAPS criterion. For this cluster, the DDT showed a significant partial correlation (age as covariate) with digit span forward scores ($r = 0.69$, $p = 0.048$).

Cluster 4 included 18 children mostly with at least average scores on all tasks other than FPT. Three children who showed DDT deficits with scores more than 2 SD below the norm also showed difficulties with the FPT. Forty four percent of children had difficulties only on FPT and seven of these also had poor sustained attention deficits. This cluster represents children with good memory, word reading, and language skills, combined with poor FPT scores and sustained attention. For cluster 4, 35% of the parents (6/17) reported listening difficulties based on responses to the CHAPS questionnaire. Although not significant, a trend was observed for an association between non-word reading and phonological processing ($r = 0.62$, $p = 0.064$) and between Core Language and paragraph reading (WARP, $r = 0.62$, $p = 0.064$).

FIGURE 2 | Linkage distance with a clear plateau means that many clusters were formed at essentially the same linkage distance. The higher the linkage distance (y-axis), the more dissimilar the groups are.

FIGURE 3 | Means for the four clusters. The y-axis shows the means (of standardized scores such that +1 is one standard deviation better than the average sample score) and the x-axis shows the 10 variables used to determine the clusters.
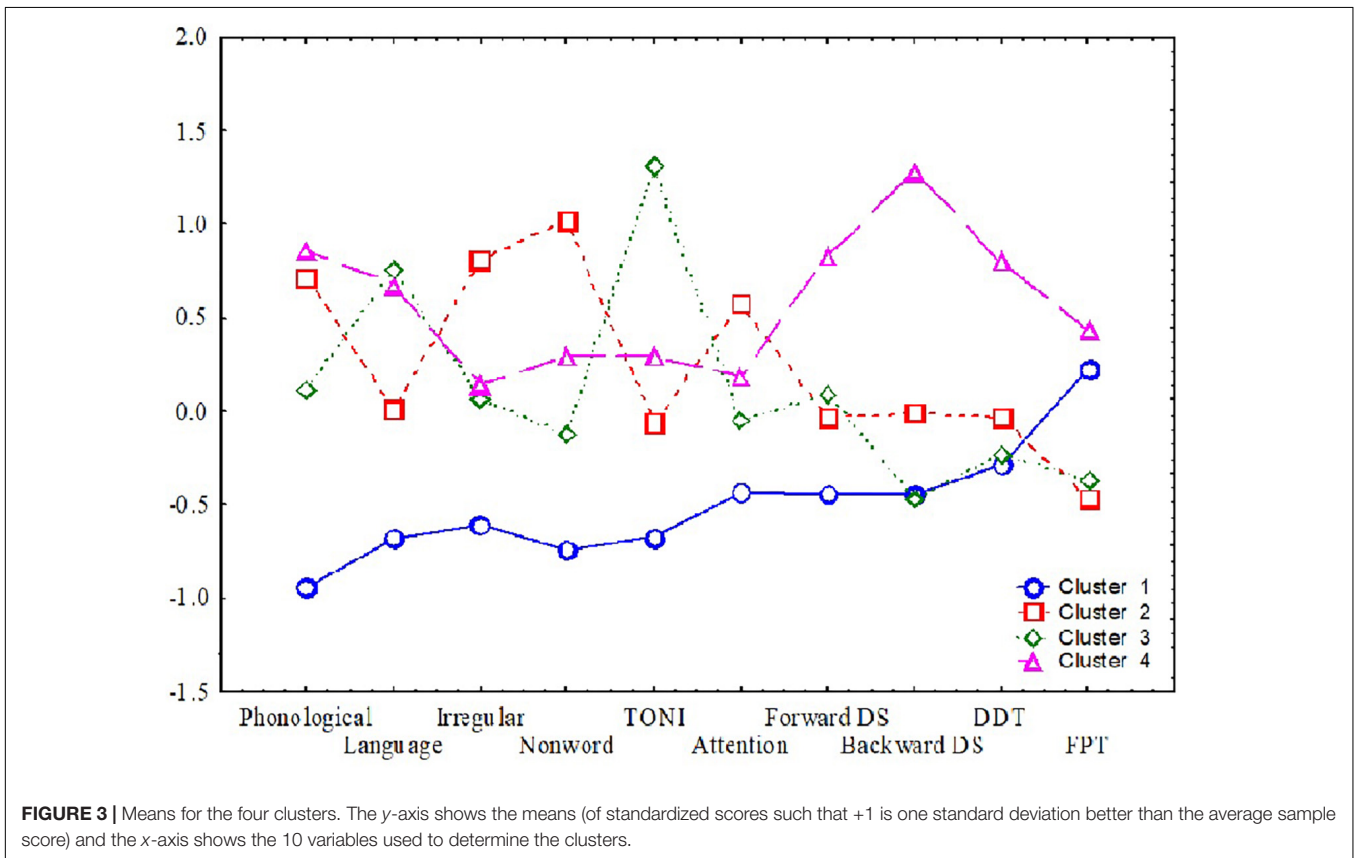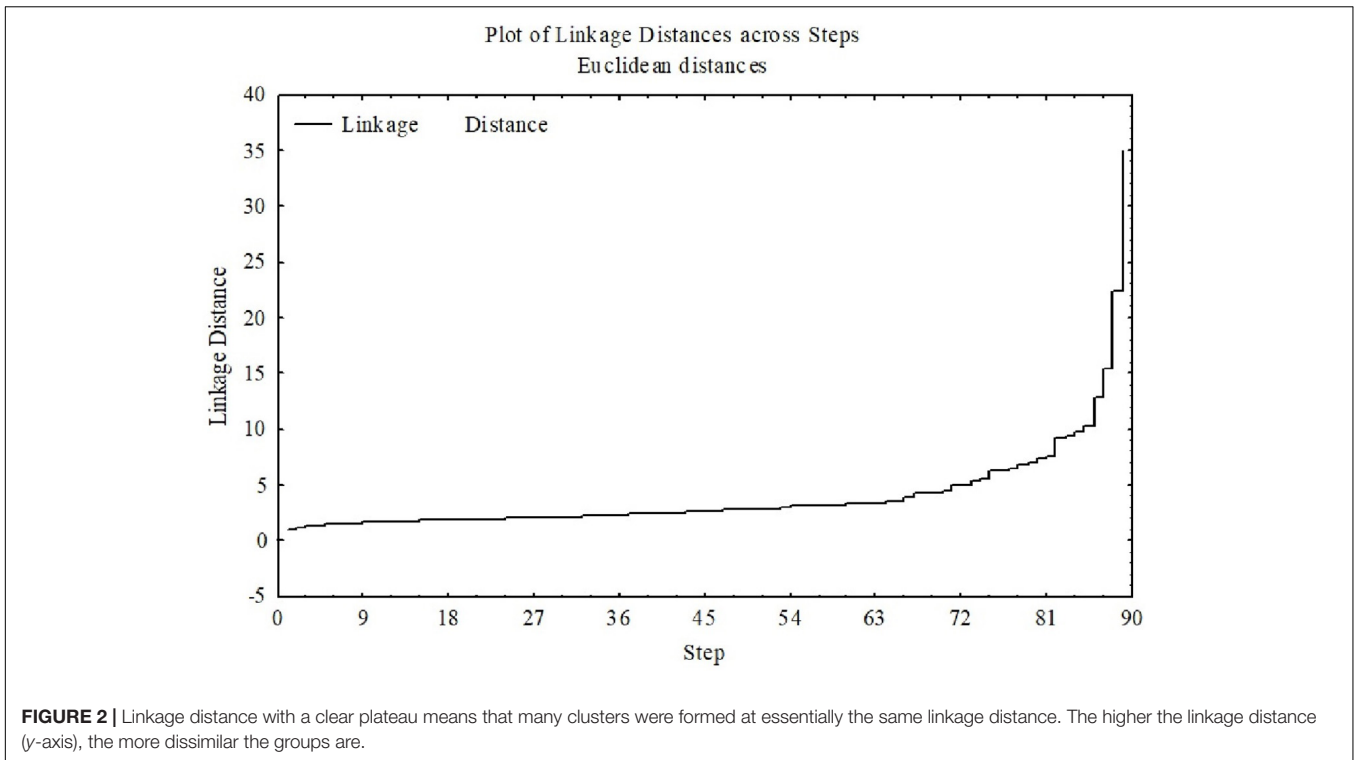
TABLE 4 | (A) Demographics and means (and standard deviation) of participants within each cluster on all skills; **(B)** means (and standard deviation) of participants within each cluster on phonological processing and language.

Variable	Total N = 90	Cluster 1, N = 35 poor reading and language	Cluster 2, N = 22 good word reading skills and attention, poor FPT	Cluster 3, N = 15 good non-verbal IQ (TONI) and language skills, poor AP	Cluster 4, N = 18 good memory (digit span), language skills and AP
(A)					
Age	9.7 (1.5)	9.5 (1.4)	10.4 (1.4)	10.2 (1.6)	9.3 (1.7)
Gender (F, M)	38, 52	13, 22	7, 15	4, 11	8, 10
Irregular word reading	17.4 (6.7)	13.5 (6.3)	**22.6 (3.5)**	18.3 (6.9)	17.6 (6.1)
Non-word reading	16.5 (8.2)	10.1 (6.4)	**24.3 (3.6)**	17.9 (5.2)	18.3 (7.9)
Forward digit span	6.5 (2.4)	5.4 (2.0)	6.4 (2.3)	6.7 (2.3)	**8.5 (2.4)**
Backward digit span	8.5 (2.8)	7.3 (2.0)	8.5 (2.1)	7.2 (2.2)	**12.1 (2.3)**
Auditory attention	77.0 (33.7)	62.4 (37.0)	**96.2 (20.0)**	75.3 (29.0)	83.2 (32.5)
TONI	99.8 (13.1)	91.0 (7.3)	99.1 (8.2)	**117.0 (13.6)**	103.7 (11.2)
DDT raw scores[#]	78.1 (13.9)	72.4 (15.4)	80.2 (9.8)	77.3 (14.1)	87.3 (9.2)
DDT	0.0 (1.0)	−0.3 (1.1)	−0.04 (0.7)	−0.2 (1.1)	**0.8 (0.4)**
FPT raw scores[#]	43.8 (27.9)	28.9 (9.3)	62.2 (24.7)	49.3 (27.8)	45.9 (31.7)
FPT	0.0 (1.0)	0.2 (1.1)	−0.5 (0.6)	−0.4 (0.8)	**0.4 (1.1)**
(B)					
Syllable identification	7.5 (3.6)	5.6 (3.2)	8.7 (3.8)	7.8 (3.0)	**9.7 (2.6)**
Spoken rhyme	6.1 (3.3)	3.7 (1.4)	8.1 (3.0)	6.3 (3.2)	**8.3 (3.2)**
Spoonerism	7.2 (3.8)	4.3 (2.4)	9.2 (3.1)	7.6 (3.8)	**10.0 (3.4)**
Phoneme detection	8.5 (3.7)	6.4 (3.3)	**10.6 (3.0)**	8.3 (3.9)	10.3 (3.1)
Phoneme manipulation	7.4 (3.3)	4.8 (2.8)	9.6 (2.1)	7.7 (2.7)	**9.7 (2.3)**
Receptive	87.6 (14.3)	79.5 (11.9)	89.5 (11.3)	**94.9 (13.9)**	**94.9 (15.3)**
Core	83.7 (13.2)	74.4 (9.7)	84.1 (10.5)	**93.7 (10.2)**	93.0 (12.2)
Auditory Comprehension	98.7 (12.6)	92.7 (11.9)	98.6 (9.6)	**106.7 (12.7)**	103.9 (11.7)
Supralinguistics	93.8 (11.6)	88.31 (12.3)	92.4 (10.3)	100.7 (9.9)	**100.7 (6.4)**

(A) Bold numbers indicate the cluster with the highest average performance on a given skill. All scores are standardized scores unless otherwise stated. [#]DDT and FPT raw scores are percent correct irrespective of age. **(B)** Bold numbers indicate the cluster with the highest average performance on a given skill. All scores are standardized scores unless otherwise stated.

DISCUSSION

Children with suspected APDs have been reported to differ from control group children without auditory difficulties on measures of attention, memory, reading and/or language skills. Comorbidity of APDs with other neurodevelopmental conditions is a norm rather than an exception (Sharma et al., 2009; Musiek et al., 2010; Tomlin et al., 2015); the proportion of children with co-occurring conditions varies across studies but is typically about 40–50% (Sharma et al., 2009; Ferguson et al., 2011). Variations are likely to reflect sampling and test protocol differences across studies. These studies have largely been cross-sectional and have used simple group comparisons, analysis of variance, and regression and correlation analyses to demonstrate links between different domains of neurodevelopmental difficulties.

A cluster analysis is unsupervised, in other words, it does not employ any *a priori* restrictions. Consequently, cluster analysis offers an advantage over other approaches in determining distinct groups based on dominant features or common skills (Clatworthy et al., 2005; Chiarello et al., 2012). The current analyses provide evidence for the validity of four clusters of children amongst the 90 children referred to the study with

suspected APDs. These clusters differentiate largely based on backward digit span, phonological processing, and non-verbal intelligence with smaller contributions from irregular word reading, forward digit span, DDT, and FPT.

Clinical Implications

The cluster analysis does not provide any information on causal relationships, instead the purpose of the clusters is to determine common links and associations within groups of participants presenting with similar difficulties (i.e., listening complaints in the current study). Cluster 1 is the only group showing global difficulties across all domains. The remaining groups all have areas of strength as well as difficulties. The question arises – what makes Cluster 1 different. It is possible that the executive function is the missing link that may explain the poor performance overall of Cluster 1. In a recent paper Snowling et al. (2018) suggested that difficulties with executive control might explain the widely reported associations between language, reading and auditory processing difficulties. Partial correlations were not significant in this cluster and hence do not support a link such as this between these skills, however, this may be due to the relatively small sample in Cluster 1.
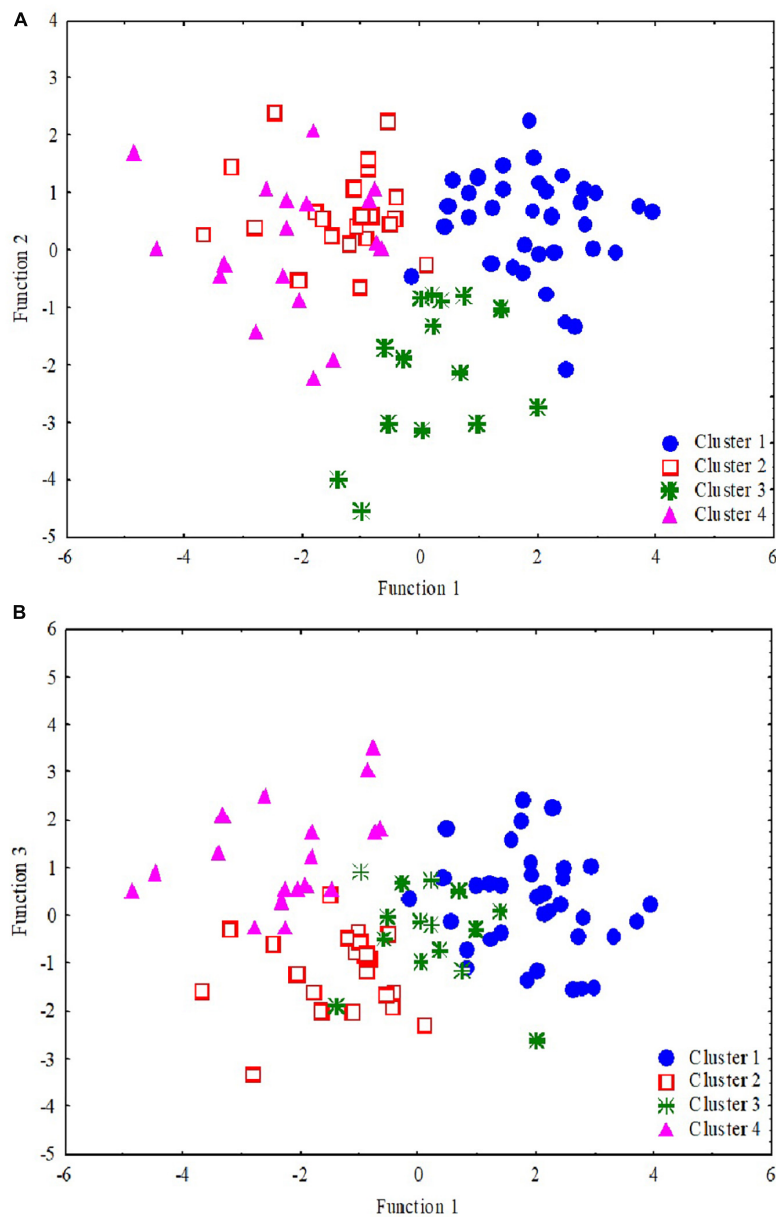
FIGURE 4 | Scatterplot of the discriminant functions. Each data point represents a single participant where cluster 1 (blue dots, n = 35), cluster 2 (red empty square, n = 22), cluster 3 (green star, n = 15), and cluster 4 (pink triangle, n = 18). The two plots depict clustering and separating of four clusters using three factors (of the 10 variables). While **(A)** shows that the clusters 1 and 3 are separated clearly by functions 1 and 2, **(B)** shows that clusters 2 and 4 separate very clearly by functions 1 and 3.

An alternative view is that all children (N = 90) within this cohort are similar and the differences in their profiles are due to strengths the children have developed, which could be compensatory or as a result of previous training or therapy. The children in the current study participated in the research when they were at least 7 years old. There were no reports of any injury or medical misadventure to account for the auditory processing concerns and, therefore, one can assume that all these children have a "developmental" APD (Moore et al., 2013). Could the current clusters be

the consequence of individual compensatory mechanisms? At present, there are no empirical data to answer this question; however, future longitudinal research could consider this question regarding the effects of variations in intervention, neuroplasticity, and maturation on the profile of skills in children with auditory processing difficulties. According to the questionnaire completed by the parents, all children showed mild to extreme deficits on CHAPS, irrespective of their groups. A longitudinal study with intervention for younger children presenting with auditory processing

TABLE 5 | Clusters in predicting reading, auditory processing performance, only significant results on ANOVA are shown.

Variables	Cluster 1 mean (SD) N = 35	Cluster 2 mean (SD) N = 22	Cluster 3 mean (SD) N = 15	Cluster 4 mean (SD) N = 18	Test of significance	Tukey post hoc
CHAPS[#][@]	−11.2 (4.6)	−11.6 (5.3) N = 20	−7.0 (4.4) N = 11	−6.9 (8.8) N = 17	$F_{(3,78)} = 3.28, p = 0.20$[$]	ns*
WARP	60.6 (40.4)	116.6 (27.9)	81.6 (49.7)	84.9 (48.7)	$F_{(3,85)} = 7.41, p = 0.002$[$]	Cluster 1 vs. 2: $p = 0.002$
Non-word spelling (QUIL)	3.5 (1.2)	6.0 (2.3)	4.0 (1.7)	7.4 (2.9)	$F_{(3,85)} = 21.42, p < 0.001$	Cluster 1 vs. 2: $p = 0.0001$; Cluster 1 vs. 4: $p = 0.0008$ Cluster 2 vs. 3: $p = 0.072$; Cluster 3 vs. 4: $p = 0.002$
Visual rhyme (QUIL)	5.3 (2.8)	7.0 (3.8)	5.0 (2.8)	4.0 (1.7)	$F_{(3,85)} = 12.75, p < 0.001$	Cluster 1 vs. 4: $p = 0.0008$; Cluster 2 vs. 4: $p = 0.032$; Cluster 3 vs. 4: $p = 0.002$
Expressive language (CELF-IV)	75.9 (11.0)	85.4 (11.5)	92.0 (12.4)	93.3 (12.7)	$F_{(3,85)} = 11.62, p < 0.001$	Cluster 1 vs. 2: $p = 0.16$; Cluster 1 vs. 3: $p = 0.002$; Cluster 1 vs. 4: $p = 0.002$;
RGDT	16.8 (14.5)	11.3 (10.3)	10.7 (10.3)	14.6 (14.0)	$F_{(3,85)} = 0.45, p = 0.715$[$]	ns*
MLD	10.5 (4.3)	11.6 (2.5)	10.2 (2.1)	11.1 (3.4)	$F_{(3,85)} = 0.61, p = 0.607$[$]	ns*
AB words 65%, 0.3s (R)	76.0 (10.2)	75.0 (9.3)	74.5 (8.3)	75.9 (6.3)	$F_{(3,85)} = 0.03, p = 0.994$	ns*
Hint sentences 65%, 0.3s (R)	57.5 (17.6)	68.7 (21.1)	66.3 (15.6)	73.4 (20.8)	$F_{(3,85)} = 4.48, p = 0.048$	ns*
CAEPS in Q and N 250 amp					$F_{(3,85)} = 1.99, p = 0.121$	ns*
CAEPS in Q and N 250 latency					$F_{(3,85)} = 1.99, p = 0.121$	ns*
CAEPS in Q and P1 amp					$F_{(3,85)} = 1.30, p = 0.279$	ns*
CAEPS in Q and P1 amp					$F_{(3,85)} = 1.30, p = 0.279$	ns*
CAEPS in Q and P1 latency					$F_{(3,85)} = 0.75, p = 0.526$	ns*

For some variables, univariate analysis was conducted (shown with [$]). The variables in bold are distinct across the clusters. Significance is adjusted for multiple testing.
$p > 0.02$ is suppressed. #CHAPS = Children's Auditory Performance Scale only completed by 83 caregivers. @ Not all parents completed the CHAPS (n = 83) and more negative implies more difficulty. $ Univariate analysis.

difficulties (e.g., 5–6 year olds) might be the best way to determine validity of these clusters and to better understand the casual relation (if it exists) between cognitive skills including attention and memory, auditory processing, and language. Leppänen et al. (2010) found electrophysiological evidence for atypical processing of sound frequency in newborns who were later identified as having phonological, reading, and language difficulties.

Another noteworthy finding is the presence of poor FPT and DDT performance in the presence of good reading and phonological processing in Cluster 2. This is an important finding as it challenges the framework suggested by Tallal or Goswami that the auditory processing link to word reading is mediated by phonological processing (Ramus, 2003). It also challenges the proposal that executive control links language, auditory processing, and reading (Snowling et al., 2018). Overall, children in Cluster 2 showed good attention and memory skills. Cluster 2, therefore, appears to be a subgroup of children who have poor auditory processing skills not linked to reading, language, memory, or attention.

Children in Cluster 3 had poor FPT, attention, and memory scores in the presence of relatively good TONI, language, and reading skills. Based on structural equation modeling,

Snowling et al. (2018) observed that executive function was predictive of frequency discrimination; therefore, it is possible that, as was the case for Cluster 1, this group could have poor executive function. The FPT test encompasses a range of skills, however, in addition to frequency discrimination, including pattern perception and verbal reporting skills, so this finding may be unrelated to the frequency discrimination aspect of the task.

Cluster 4 is somewhat similar to Cluster 3 as both groups exhibit poor performance on FPT and poor attention skills. However, Clusters 3 and 4 differ in their backward digit span scores, with Cluster 4 showing higher performance compared to Cluster 3. Poor attention is not an obvious explanation for poor auditory processing, as children in Cluster 2 had poor auditory processing despite the presence of good attention skills. While attention has been linked with performance on the AP tasks in general (Moore et al., 2010), sustained attention has not been found to contribute to the performance on the FPT (Gyldenkærne et al., 2014; Tomlin et al., 2015).

The participants in this study are likely to be representative of the children referred for clinical evaluation of auditory processing (since referrals to the research came from a

range of professionals and parents). Consequently, the four clusters may be representative of children with suspected APD, but the identified clusters are unlikely to be the only ones that exist in the population of children with neurodevelopmental difficulties affecting learning and behavior. Despite this limitation, there are some potential benefits of identifying these subgroups of children presenting for auditory processing assessment. The distinct clusters identified in the current study highlight the heterogeneity of children with suspected APD, and this result encourages clinicians to ensure assessments span all the domains examined here, especially those that contributed most to the separation of the clusters, namely working memory (backward digit span), non-verbal IQ, non-word reading, and phonological processing. Assessment of these areas in children with a diagnosis of APD could assist clinicians to choose appropriate referral pathways and treatments.

Although all groups included children with APD and parent-rated listening difficulties based on the CHAPS questionnaire, some children might make better functional gains if their specific phonological processing, reading, and/or other difficulties are targeted. For instance, Cluster 1 might benefit from referral to a psychologist for cognitive assessment that includes measures of executive control and is likely to need a broad range of supports. Audiologists would best manage children in Cluster 2. Clusters 3 and 4 could benefit from auditory training delivered by an audiologist or speech pathologist and a psychologist would be able to conduct more comprehensive attention and memory assessment and management suggestions. Although this suggests a different pathway for each cluster, all children presented with listening difficulties, and hence all are likely to benefit from treatments such as personal remote microphone technology to improve the signal to noise ratio in difficult listening situations (Sharma et al., 2012).

Leppänen et al. (2010) identified auditory insensitivity 3–5 days after birth in about half of the infants they tested with familial risk for dyslexia, using a mismatch negativity paradigm (event-related potential response to infrequently presented 1100 Hz deviant sinusoidal stimulus among frequently presented 1000 Hz sinusoidal stimulus). They found that about half of the participants in this longitudinal study had impaired differentiation of basic pitch changes at birth and these children were later diagnosed with dyslexia; the other half of the children with normal mismatch negativity responses in infancy did not have problems in reading acquisition when tested 8 years later. This paper highlights the possibility of earlier identification of auditory difficulties using electrophysiological approaches. This would allow the possibility of early interventions targeted at enhancing auditory processing that might prevent later literacy difficulties. This could change the profiles of children with APD in the future.

Limitations and Future Directions

Cluster analysis is an unbiased way to determine subgroups; there are some limitations, however. For instance, the participating

90 children created the current four clusters, and validation using a different, larger sample would be useful to confirm the characteristics of the clusters. With a larger sample, the details of the clustering might change (as in, some children could be assigned to a different cluster if the data looked a bit different, or different tests were included), but the overall differences between clusters identified in the current study are sufficiently pronounced that the interpretation of the subgroups identified here may not change. With a larger sample, the stability of the clusters could be determined by comparing the clustering of the original data set with the clustering obtained on subsamples or with a completely new data set (Levine and Domany, 2001).

The clusters did not differ in the balance of boys to girls (Chi-square = 1.31, $p = 0.73$), although there were more males than females overall. There was a trend for two clusters (Clusters 1 and 4) to be slightly younger [$F(1, 3) = 3.02$, $p = 0.034$] than the other two, however. Higher numbers with equal gender proportions to account for slight age and gender variations may assist in generalization of the clusters.

In the current cluster analysis only two auditory processing measures with established age norms were included (FPT and DDT). Inclusion of other auditory skills, such as spatial listening (LISN-S) (Cameron and Dillon, 2007) or temporal or frequency discrimination (Moore et al., 2010) might yield different results if these auditory skills are more strongly linked than FPT and DDT to cognition and other skill areas. In future research, it would be useful to include a wider range of norm-referenced auditory processing measures that capture the range of auditory skills typically included in the clinical auditory processing test battery. Due to the complexity of reading disorders (Horbach et al., 2019), a more detailed assessment of reading abilities and potential underlying deficits such as temporal or phonological processing might also affect cluster membership.

It is possible that children with neurodevelopmental disorders will show evidence of different difficulties at different ages, even if deficits were solely in the auditory domain at an early age. More longitudinal research is needed to establish the stability of clusters over time as it is possible that training of specific cognitive and/or auditory skills would give rise to different results over time.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this manuscript are not publicly available. Requests to access the datasets should be directed to mridula.sharma@mq.edu.au.

ETHICS STATEMENT

The studies involving the human participants were reviewed and approved by the University of Auckland Human Research Participants' Ethics Committee. Written informed consent to

participate in this study was provided by the participants' legal guardian/next of kin and verbal assent from all participating children was also gathered.

AUTHOR CONTRIBUTIONS

MS: the concept of the project, data analysis, and writing of the manuscript. SP: the concept of the project and editorial of the manuscript. PH: advising on the data analysis and editorial of the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02481/full#supplementary-material

REFERENCES

American Speech and Hearing Association [ASHA] (1996). Central auditory processing: current status of research and implications for clinical practice. *Am. J. Audiol.* 5, 41–52. doi: 10.1044/1059-0889.0502.41

American Speech Language Hearing Association (2005). *(Central) Auditory Processing Disorders – The Role of the Audiologist*. Available at: www.asha.org/policy

Bellis, T. J., and Ferre, J. M. (1999). Multidimensional Approach to. *J. Am. Acad. Audiol.* 10, 319–328.

Bishop, D. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: where are we, and where should we be going? *Psychol. Bull.* 133, 651–672. doi: 10.1037/0033-2909.133.4.651

Brown, L., Sherbenou, R. J., and Johnsen, S. K. (1990). *Test of Nonverbal Intelligence: A Language-Free Measure of Cognitive Ability*. Austin, TX: PRO-ED.

Cameron, S., and Dillon, H. (2007). Development of the listening in spatialized noise-sentences test (LISN-S). *Ear Hear.* 28, 196–211. doi: 10.1097/aud.0b013e318031267f

Carrow-Woolfolk, E. (1999). *CASL: Comprehensive Assessment of Spoken Language*. Circle Pines, MN: American Guidance Services.

Castles, A., and Coltheart, M. (1993). Varieties of developmental dyslexia. *Cognition* 47, 149–180. doi: 10.1016/0010-0277(93)90003-e

Chermak, G. D., Tucker, E., and Seikel, J. A. (2002). Behavioral characteristics of auditory processing disorder and attention-deficit hyperactivity disorder: predominantly inattentive type. *J. Am. Acad. Audiol.* 13, 332–338.

Chiarello, C., Welcome, S. E., and Leonard, C. M. (2012). Individual differences in reading skill and language lateralisation: a cluster analysis. *Laterality* 17, 225–251. doi: 10.1080/1357650X.2011.561860

Clatworthy, J., Buick, D., Hankins, M., Weinman, J., and Horne, R. (2005). The use and reporting of cluster analysis in health psychology: a review. *Br. J. Health Psychol.* 10, 329–358. doi: 10.1348/135910705x25697

Dodd, B., Holm, A., Orelemans, M., and McCormick, M. (1996). *QUIL: Queensland University Inventory of Literacy*. Brisbane: The University of Queensland.

Emanuel, D. C., Ficca, K. N., and Korczak, P. (2011). Survey of the diagnosis and management of auditory processing disorder. *Am. J. Audiol.* 20, 48–60. doi: 10.1044/1059-0889(2011/10-0019)

Ferguson, M. A., Hall, R. L., Riley, A., and Moore, D. R. (2011). Communication, listening, cognitive and speech perception skills in children with auditory processing disorder (APD) or specific language impairment (SLI). *J. Speech Lang. Hear. Res.* 54, 211–227. doi: 10.1044/1092-4388(2010/09-0167)

Gilley, P. M., Sharma, M., and Purdy, S. C. (2016). Oscillatory decoupling differentiates auditory encoding deficits in children with listening problems. *Clin. Neurophysiol.* 127, 1618–1628. doi: 10.1016/j.clinph.2015.11.003

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci.* 15, 3–10. doi: 10.1016/j.tics.2010.10.001

Gyldenkærne, P., Dillon, H., Sharma, M., and Purdy, S. C. (2014). Attend to this: the relationship between auditory processing disorders and attention deficits. *J. Am. Acad. Audiol.* 25, 676–687. doi: 10.3766/jaaa.25.7.6

Hall, J. W. (2000). *Handbook of Otoacoustic Emissions*. San Diego: Singular Publishing Company.

Halliday, L. F., Tuomainen, O., and Rosen, S. (2017). Auditory processing deficits are sometimes necessary and sometimes sufficient for language difficulties in children: evidence from mild to moderate sensorineural hearing loss. *Cognition* 166, 139–151. doi: 10.1016/j.cognition.2017.04.014

Hämäläinen, J. A., Salminen, H. K., and Leppänen, P. H. (2013). Basic auditory processing deficits in dyslexia: systematic review of the behavioral and event-related potential/field evidence. *J. Learn. Disabil.* 46, 413–427. doi: 10.1177/0022219411436213

Horbach, J., Mayer, A., Scharke, W., Heim, S., and Günther, T. (2019). Development of behavior problems in children with and without specific learning disorders in reading and spelling from kindergarten to fifth grade. *Sci. Stud. Read.* 1–15. doi: 10.1080/10888438.2019.1641504

Iliadou, V. V., Ptok, M., Grech, H., Pedersen, E. R., Brechmann, A., Deggouj, N., et al. (2018). European 17 countries consensus endorses more approaches to APD than reported in Wilson, 2018. *Int. J. Audiol.* 57, 395–396. doi: 10.1080/14992027.2018.1442937

Jerger, J. (1970). Clinical experience with impedance audiometry. *Arch. Otolaryngol.* 92, 311–324. doi: 10.1001/archotol.1970.04310040005002

Jerger, J., Brown, D., and Smith, S. (1984). Effect of peripheral hearing loss on the masking level difference. *Arch. Otolaryngol.* 110, 290–296. doi: 10.1001/archotol.1984.00800310014003

Jerger, J., and Musiek, F. (2000). Report of the consensus conference on the diagnosis of auditory processing. *J. Am. Acad. Audiol.* 11, 467–474.

Keith, R. W. (2000). *Random Gap Detection Test*. St. Louis, MO: Auditec.

Keith, W., Purdy, S. C., Baily, M., and Kay, F. (2019). *New Zealand Guidelines on Auditory Processing Disorder*. Available at: https://www.audiology.org.nz/: New Zealand Audiological Society (accessed May 31, 2019).

Kelly, A. (2007). Normative data for behavioural tests of auditory processing for New Zealand school children aged 7 to 12 years. *Aust. N. Z. J. Audiol.* 29, 60–64. doi: 10.1375/audi.29.1.60

Leppänen, P. H., Hämäläinen, J. A., Salminen, H. K., Eklund, K. M., Guttorm, T. K., Lohvansuu, K., et al. (2010). Newborn brain event-related potentials revealing atypical processing of sound frequency and the subsequent association with later literacy skills in children with familial dyslexia. *Cortex* 46, 1362–1376. doi: 10.1016/j.cortex.2010.06.003

Levine, E., and Domany, E. (2001). Resampling method for unsupervised estimation of cluster validity. *Neural Comput.* 13, 2573–2593. doi: 10.1162/089976601753196030

Madelaine, A., and Wheldall, K. (1998). Towards a curriculum-based passage reading test for monitoring the performance of low-progress readers using standardised passages: a validity study. *Educ. Psychol.* 18, 471–478. doi: 10.1080/0144341980180408

Mealings, K., and Cameron, S. (2019). Investigating auditory spectral and temporal resolution deficits in children with reading difficulties. *J. Am. Acad. Audiol.* 30, 533–543. doi: 10.3766/jaaa.17142

Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., and Riley, A. (2010). Nature of auditory processing disorder in children. *Pediatrics* 126, e382–e390. doi: 10.1542/peds.2009-2826

Moore, D. R., Rosen, S., Bamiou, D.-E., Campbell, N. G., and Sirimanna, T. (2013). Evolving concepts of developmental auditory processing disorder (APD): a

British society of audiology APD special interest group "white paper.". *Int. J. Audiol.* 52, 3–13. doi: 10.3109/14992027.2012.723143

Musiek, F. E. (1983). Assessment of central auditory dysfunction: the dichotic digit test revisited. *Ear Hear.* 4, 79–83. doi: 10.1097/00003446-198303000-00002

Musiek, F. E. (1994). Frequency (pitch) and duration pattern tests. *J. Am. Acad. Audiol.* 5, 265–265.

Musiek, F. E., Baran, J. A., Bellis, T. J., Chermak, G. D., Hall, J. W., Keith, R. W., et al. (2010). Guidelines for the diagnosis, treatment and management of children and adults with central auditory processing disorder. *Am. Acad. Audiol.* 1–50.

Purdy, S. C., Sharma, M., and Morgan, A. (2018). Measuring perceptions of classroom listening in typically developing children and children with auditory difficulties using the LIFE-UK questionnaire. *J. Am. Acad. Audiol.* 29, 656–667. doi: 10.3766/jaaa.17053

Ramus, F. (2003). Developmental dyslexia: specific phonological deficit or general sensorimotor dysfunction? *Curr. Opin. Neurobiol.* 13, 212–218. doi: 10.1016/s0959-4388(03)00035-7

Richardson, U., Thomson, J. M., Scott, S. K., and Goswami, U. (2004). Auditory processing skills and phonological representation in dyslexic children. *Dyslexia* 10, 215–233. doi: 10.1002/dys.276

Sandford, J., and Turner, A. (2000). *Integrated Visual and Auditory Continuous Performance Test Manual.* Richmond, VA: Braintrain Inc.

Sandford, J. A., and Turner, A. (1995). *Manual for the Integrated Visual and Auditory Continuous Performance Test.* Richmond: Braintrain.

Semel, E., Wiig, E., and Secord, W. (2003). *Clinical Evaluation of Language Fundamentals (CELF-4).* San Antonio, TX: The psychological corporation.

Sharma, M., Cupples, L., and Purdy, S. C. (2019). Predictors of reading skills in children with listening concerns. *Ear. Hear.* 40, 243–252. doi: 10.1097/AUD.0000000000000608

Sharma, M., Dhamani, I., Leung, J., and Carlile, S. (2014a). Attention, memory, and auditory processing in 10-to 15-year-old children with listening difficulties. *J. Speech Lang. Hear Res.* 57, 2308–2321. doi: 10.1044/2014_JSLHR-H-13-0226

Sharma, M., Purdy, S., and Kelly, A. (2014b). The contribution of speech-evoked cortical auditory evoked potentials to the diagnosis and measurement of intervention outcomes in children with auditory processing disorder. *Semin. Hear.* 35, 51–64.

Sharma, M., Purdy, S., and Kelly, A. (2012). A randomized control trial of interventions in school-aged children with auditory processing disorders. *Int. J. Audiol.* 51, 506–518. doi: 10.3109/14992027.2012.670272

Sharma, M., Purdy, S., Newall, P., Wheldall, K., Beaman, R., and Dillon, H. (2006). Electrophysiological and behavioral evidence of auditory processing deficits in children with reading disorder. *Clin. Neurophysiol.* 117, 1130–1144. doi: 10.1016/j.clinph.2006.02.001

Sharma, M., Purdy, S. C., and Kelly, A. S. (2009). Comorbidity of auditory processing, language, and reading disorders. *J. Speech Lang. Hear. Res* 52, 706–722. doi: 10.1044/1092-4388(2008/07-0226)

Silman, S., and Gelfand, S. A. (1981). The relationship between magnitude of hearing loss and acoustic reflex threshold levels. *J. Speech Hear. Dis.* 46, 312–316. doi: 10.1044/jshd.4603.312

Smoski, W., Brunt, M. A., and Tannahill, J. C. (1998). *Children's Auditory Performance Scale (CHAPS).* Tampa, FL: Educational Audiology Association.

Snowling, M. J., Gooch, D., McArthur, G., and Hulme, C. (2018). Language skills, but not frequency discrimination, predict reading skills in children at risk of dyslexia. *Psychol. Sci.* 29, 1270–1282. doi: 10.1177/0956797617876 3090

Sweetow, R., and Reddell, R. (1978). The use of masking level differences in the identification of children with perceptual problems. *J. Am. Aud. Soc.* 4, 52–56.

Taylor, R. (1990). Interpretation of the correlation coefficient: a basic review. *J. Diagn. Med. Sonogr.* 6, 35–39. doi: 10.1177/8756479390006 00106

Tomlin, D., Dillon, H., Sharma, M., and Rance, G. (2015). The impact of auditory processing and cognitive abilities in children. *Ear Hear* 36, 527–542. doi: 10.1097/AUD.0000000000000172

Wilson, W. J. (2018). Evolving the concept of APD. *Int. J. Audiol.* 57, 240–248. doi: 10.1080/14992027.2017.1409438

frontiers
in Psychology

Speech-in-Noise Perception in Children With Cochlear Implants, Hearing Aids, Developmental Language Disorder and Typical Development: The Effects of Linguistic and Cognitive Abilities

Janne von Koss Torkildsen[1]*, Abigail Hitchins[1,2], Marte Myhrum[3,4] and Ona Bø Wie[1,3]

[1] Department of Special Needs Education, Faculty of Educational Sciences, University of Oslo, Oslo, Norway, [2] Auditory Verbal UK, Oxon, United Kingdom, [3] Division of Head, Neck and Reconstructive Surgery, Department of Otorhinolaryngology and Head and Neck Surgery, Oslo University Hospital, Oslo, Norway, [4] Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

Children with hearing loss, and those with language disorders, can have excellent speech recognition in quiet, but still experience unique challenges when listening to speech in noisy environments. However, little is known about how speech-in-noise (SiN) perception relates to individual differences in cognitive and linguistic abilities in these children. The present study used the Norwegian version of the Hearing in Noise Test (HINT) to investigate SiN perception in 175 children aged 5.5–12.9 years, including children with cochlear implants (CI, $n = 64$), hearing aids (HA, $n = 37$), developmental language disorder (DLD, $n = 16$) and typical development (TD, $n = 58$). Further, the study examined whether general language ability, verbal memory span, non-verbal IQ and speech perception of monosyllables and sentences in quiet were predictors of performance on the HINT. To allow comparisons across ages, scores derived from age-based norms were used for the HINT and the tests of language and cognition. There were significant differences in SiN perception between all the groups except between the HA and DLD groups, with the CI group requiring the highest signal-to-noise ratios (i.e., poorest performance) and the TD group requiring the lowest signal-to-noise ratios. For the full sample, language ability explained significant variance in HINT performance beyond speech perception in quiet. Follow-up analyses for the separate groups revealed that language ability was a significant predictor of HINT performance for children with CI, HA, and DLD, but not for children with TD. Memory span and IQ did not predict variance in SiN perception when language ability and speech perception in quiet were taken into account. The finding of a robust relation between SiN perception and general language skills in all three clinical groups call for further investigation into the mechanisms that underlie this association.

Keywords: hearing in noise, speech in noise perception, children, hearing loss, cochlear implant, hearing aid, language ability, developmental language disorder

INTRODUCTION

Perceiving language in busy and often noisy classrooms and playgrounds is a challenge for all children in mainstream schools. The signal-to-noise ratio (SNR) is defined as the ratio between the speech dB level and the noise dB level. A negative SNR means that the noise is higher than the speech. A SNR of +15 or +20 dB is recommended for classrooms by the American Speech-Language-Hearing Association [ASHA] (1995) and the British Association of Teachers of the Deaf [BATOD] (2001), respectively. However, results from numerous studies in actual classroom situations indicate that the SNR is much lower, and in some cases even negative (Crandell and Smaldino, 1994, 2000). Not only must children listen in noisy settings, attempting to learn new words and facts, but the ambition is for better grades and better peer-to-peer relationships and social functioning. Shield and Dockrell (2008) demonstrated that external and internal noise in classrooms had a negative impact on the academic test results of typically hearing children aged 7 and 11 years. This negative relationship between performance and noise levels was maintained when the data were corrected for socio-economic factors relating to social deprivation, language, and special educational needs.

For children with hearing loss, noise makes mainstream schooling an even bigger challenge. Hearing loss affects speech-in-noise (SiN) perception via at least three main mechanisms. The first is loss of audibility, especially at high frequencies where speech sounds are lower in intensity (Soli and Wong, 2008). The second is distortion, or loss of spectral and temporal processing sensitivity and selectivity, which reduces speech perception in noise even when speech is entirely audible (Plomp, 1978, 1986; Bronkhorst, 2000). The third is less efficient binaural processing compared to typically hearing children, an essential aspect of listening in background noise, especially when speech and noise are not collocated (for a review see Bronkhorst, 2000). The latter two mechanisms will cause deficits in supra-threshold auditory processing tasks in which the audio signals of interest are audible to the listener.

In addition, there may be indirect effects of hearing loss on SiN perception via cognitive skills such as language and phonological working memory. A number of studies have shown that children with hearing loss perform less well than their typically developing peers on measures of language skills such as vocabulary and grammar (e.g., Tomblin et al., 2015; Ching et al., 2019), and on measures of phonological working memory as measured by non-word repetition or digit span tasks (e.g., Pisoni and Cleary, 2003; Lyxell et al., 2008; Davidson et al., 2019). However, the consequences of such linguistic and cognitive deficits on SiN perception are not well understood. Below we review studies that have examined the link between SiN perception and cognition generally and in children with hearing loss and language disorders specifically.

The Ease of Language Understanding (ELU) model provides a theoretical framework for how perceptual input characteristics interact with cognition in noisy listening conditions (Rönnberg et al., 2010). The model posits that as long as listening conditions are optimal, implicit processing mechanisms rapidly map the auditory input to phonological representations in long-term memory. Under noisy conditions, however, the implicit processing mechanisms may fail and lead to mismatches between input and stored phonological representations. In this situation, explicit processing mechanisms are invoked. Specifically, the listener is required to use his or her explicit working memory and linguistic knowledge to prospectively and retrospectively reconstruct the input and infer meaning. The ability to resolve mismatches between input and phonological representations will consequently depend both upon the individual's working memory capacity and linguistic abilities. This theoretical position corresponds well with findings from a literature review by Akeroyd (2008) which examined 20 studies of the relationship between SiN perception and some aspect of cognition in mostly elderly hearing impaired adults. Results showed that measures of working memory were typically significant predictors, whereas measures of general ability, such as IQ, did not significantly predict SiN perception. However, the assumption of an association between working memory and SiN perception may hold only for some populations, such as older hearing-impaired listeners. A recent meta-analysis of studies examining young adults with typical hearing failed to find evidence of a relation between working memory (measured by the reading span task) and SiN perception (Füllgrabe and Rosen, 2016).

To our knowledge, no meta-analysis has investigated the relation between SiN perception and cognition in children, possibly because there are relatively few studies on this topic. However, the literature examining children with typical hearing suggests that the relation between SiN perception and cognitive abilities may depend on the task that is used to measure SiN perception, specifically whether it involves identification of single words or larger linguistic units. A study of school age children and adolescents by Talarico et al. (2007) found no significant correlations between SiN perception of single words and verbal or non-verbal IQ scores. In line with this, a large-scale study of school age children and young adults found no significant association between performance on the Words-in-Noise Test and receptive vocabulary (Wilson et al., 2010). On the other hand, Sullivan et al. (2015) who measured comprehension of orally presented passages, found that the relationship between auditory working memory and comprehension was stronger in noise than in quiet, indicating an increased contribution of working memory in noisy conditions. Additionally, three recent studies of SiN perception in school age children and adolescents found a significant relationship between SiN perception and (backward or forward) memory span (MacCutcheon et al., 2018, 2019, submitted). In these three studies the SiN task involved identification of two changing target words in an otherwise fixed carrier sentence. In sum, the literature to date suggests that phonological working memory and higher-order linguistic skills such as vocabulary and grammar may be more closely associated with SiN perception of sentences and passages than single words.

Studies which have focused on predictors of SiN perception for children with hearing loss specifically indicate a role for cognitive factors such as language and working memory. Ching et al. (2018) studied 252 5-year-old children with hearing aids (HA) and cochlear implants (CI) who were enrolled in the Longitudinal

Outcomes of Children with Hearing Impairment study. Speech in babble perception was measured with either a word identification or sentence repetition task, depending on the language abilities of the child as judged by speech pathologists. The authors found that non-verbal IQ and language abilities were significant predictors of speech perception in babble in children using HA, with the effect size of language ability almost double that of non-verbal IQ. For children using CI, age at implantation and language abilities were significant predictors. After taking into account the effect of language ability, non-verbal IQ was not a significant predictor of speech perception in babble for children with CI. Another study by Caldwell and Nittrouer (2013) examined 27 children who used unilateral or bilateral CI, 8 children who wore bilateral HAs and 19 children with typical hearing. Children completed tasks measuring speech perception of single words in quiet and in three different SNRs (+3, 0, and −3 dB). In addition, they measured phonological awareness, general language, and cognitive skills. Children with typical hearing had better speech recognition in quiet than children with either HA or CI. However, only a small group × SNR interaction effect was observed. Interestingly, when speech perception in quiet was accounted for, there was not a significant interaction effect of group × SNR. This finding suggests that the processing limitations imposed by HA and CI had the biggest effect on recognition in quiet (when measured as phoneme score on consonant-vowel-consonant monosyllabic words), whereas the noise effects on speech perception were comparable for all children. For the participant group as a whole, general language abilities and phonological awareness explained significant variance in both phoneme and word recognition in noise. Short-term memory also explained variance in word recognition (but not phoneme recognition). However, none of these effects reached significance when the groups were considered separately.

If language, and the cognitive abilities that underlie language, do play a substantial role in SiN perception, this may also leave children with typical hearing, but a deficit in language, such as Developmental Language Disorder (DLD), vulnerable to noisy environments. Indeed, some studies have found that children with DLD have speech perception deficits in both silence and noise when tested with non-word monosyllables designed to measure discrimination of phonological contrasts (Ziegler et al., 2005, 2011). In line with this, Ziegler et al. (2011) found a significant association between SiN perception and language ability within the DLD group, but not in the typically developing control group. Another study which measured SiN perception by real word monosyllables, reported that children with DLD and co-occurring literacy impairment had a deficit in SiN perception compared to typically developing peers and children with DLD but no literacy problems (Vandewalle et al., 2012). In contrast to these results, a study by Ferguson et al. (2011) found no differences between unselected children and children with DLD in speech perception, either in quiet or in noise, when measured with both sentence lists and non-word monosyllables. However, group differences in SiN were numerically larger when measured with monosyllables than with sentences. Taken together, it appears that SiN perception may be deviant in children with DLD when tested with syllables designed to tap discrimination of

phonological contrasts. However, it is unclear whether children with DLD also have a deficit when SiN perception is measured by sentences. It is well established that children with DLD exhibit a robust deficit in sentence repetition tasks when sentences are linguistically complex or long (e.g., Conti-Ramsden et al., 2001), but the sentences used in SiN tasks are typically simple and short, as the tests are constructed to measure speech perception and not general language abilities.

In addition to possibly being dependent upon cognitive abilities such as language ability and phonological working memory, SiN perception also appears to develop with age. A main finding from studies with young school-age children (5–12 years) with typical hearing is that speech perception in noise gets better during this period in development (e.g., Elliott, 1979; Fallon et al., 2000; Picard and Bradley, 2001; Jamieson et al., 2004; Stuart, 2005; Neuman et al., 2010; Wilson et al., 2010; Myhrum et al., 2016). This may be due to the protracted development of children's auditory system, for example the gradual maturation of binaural processing (Hall and Grose, 1990; Moore et al., 2001), but also to the language development that happens during this period (Myhrum et al., 2016). This developmental trend could also be partly explained by the maturation of other cognitive abilities that may be involved in SiN perception, such as attention and processing speed (Gomes et al., 2000; Luna et al., 2004).

Over the last 25 years, research has devised a number of adaptive test paradigms, using words or sentences and different maskers, to measure SiN perception. Examples of such tests include the Hearing in Noise Test (HINT; Nilsson et al., 1994), the Speech Recognition in Noise Test (SPRINT, Cord et al., 1992), and the Words-in-Noise Test (WIN; Wilson, 2003). The HINT is one of the most widely used adaptive tests of SiN perception (Harianawala et al., 2019). It is commonly used in clinical practice, and the paper by Soli and Wong (2008) lists normative data for 13 different languages. The developers of the HINT have attempted to address many of the factors which are known to affect SiN, such as the speech material, masking noise and test room acoustics (Soli and Wong, 2008). For example, the speech materials consist of "Short, simple sentences from children's books at a first grade reading level" (Soli and Wong, 2008, p. 356). The sentences are evaluated for naturalness by native speakers, and those with low scores are rejected or modified. When children's versions of the HINT are developed, sentences are also evaluated specifically on appropriateness for the youngest school age children (Myhrum et al., 2016). Still, the previous literature reviewed above suggest that it is possible that the sentences which are acceptable for 5-year-old typically developing children may be challenging in terms of linguistic complexity, cognitive or working memory demands for children with hearing loss or developmental disabilities, especially when presented in noise. This is a critical question, as these are the target populations for the HINT in clinical practice. It will therefore be useful to know whether language skills, working memory or non-verbal abilities may explain variance in performance on the HINT for children with hearing loss or language disorders.

The main aims of the present study were firstly, to investigate the differences in SiN perception, as measured by the HINT, between four groups of school-age children—children using CI,

children using HA, children with DLD and typically developing (TD) children. Secondly, the study aimed to explore whether language ability, working memory or non-verbal IQ could explain variance in HINT performance.

MATERIALS AND METHODS

Participants

The 175 participants in the present study were recruited from a wider project to specifically investigate performance on the HINT in children aged 5.5–12.9 years old. The wider project was approved by the Regional Committees for Medical and Health Research Ethics, South-East Norway. Written informed consent was obtained from the parents of all participants. In addition, oral consent was obtained from the participating children after receiving information about the study and the tasks involved.

The inclusion criteria in the present study were set to investigate performance on the HINT and control for other factors that could affect performance on the HINT. All children, in the CI, HA, DLD and TD groups, met the following inclusion criteria: (1) They had completed the HINT, (2) they all had a standard score of 75 or above on the non-verbal IQ test Raven's Progressive matrices (Raven et al., 2004; Raven and Raven, 2008), (3) the child, and at least one of the child's parents, had spoken Norwegian as their native language, and (4) no diagnosis of other developmental disorders such as autism or ADHD had been made. An additional inclusion criterion for the TD and DLD groups was (5) parent report of normal hearing and the presence of otoacoustic emissions (OAEs) in both ears, indicating no damage in outer-hair-cell function. Presence of OAEs is in most cases associated with normal hearing sensitivity and hearing thresholds (Lucertini et al., 2002; Engdahl et al., 2005; Stach, 2010). An additional inclusion criterion for the CI and HA groups was (6) the child used bilateral HAs or CIs. There were two additional criteria, (7) and (8), used in the recruitment of the DLD group. These are detailed under 'Characteristics of the children in the DLD group' below.

In the present study, no official diagnosis of additional needs were made. Criterion 2 was set to prevent including children who had intellectual disabilities, defined in DSM V as IQ scores below 70, including a margin of measurement error (American Association on Intellectual and Developmental Disabilities, 2010; American Psychiatric Association, 2013). This criterion adhered to the definition of DLD that language difficulties should not be associated with known biomedical conditions such as intellectual disabilities (Bishop et al., 2017).

Participants in the present study included 64 children using bilateral cochlear implants (CI) (38 boys, 59%), 37 children using bilateral hearing aids (HA) (16 boys, 43%), 16 children with DLD (11 boys, 69%), and 58 children with typical development (TD) (22 boys, 38%). Children in all groups were recruited in the same age range (5.5–12.9 years), and the groups had a similar age distribution (see **Table 1** and **Figure 1**). However, the groups were not matched for age. A one-way ANOVA showed a significant effect of group on age [$F(3,171)$ = 5.3, p = 0.002]. *Post hoc* comparisons using the Bonferroni correction, showed that there

TABLE 1 | Descriptive statistics by group for age, speech perception, language ability, non-verbal IQ, and memory span.

	CI (*n* = 64)	HA (*n* = 37)	DLD (*n* = 16)	TD (*n* = 58)
	Mean (*SD*) (range)	Mean (*SD*) (range)	Mean (*SD*) (range)	Mean (*SD*) (range)
Age [years]	10.1 (1.8) (6.9–12.9)	8.7 (2.2) (5.5–12.7)	9.7 (1.8) (6.5–12.5)	9.0 (2.0) (5.7–12.8)
Speech perception (monosyllables)	87.2 (7.0) (68–100)	88.4 (11.0) (48–100)	95.6 (8.4) (70–100)	99.0 (1.6) (92–100)
Speech perception (sentences)	96.3 (5.6) (77–100)	95.8 (6.0) (70–100)	93.6 (16.1) (42–100)	99.7 (0.8) (96–100)
Speech-in-noise perception (SiN)	2.6 (2.5) (−3.2–10.5)	0.6 (3.1) (−3.7–13.6)	−0.8 (2.3) (−4.1–4.4)	−2.9 (1.2) (−5.4−−0.3)
Language ability	76.1 (18.4) (42–114)	85.2 (15.5) (47–117)	66.9 (15.7) (44–102)	102.7 (14.9) (57–135)
Non-verbal IQ	97.7 (11.2) (75–120)	98.5 (16.1) (75–135)	90.6 (11.8) (75–115)	103.6 (13.6) (80–145)
Memory span	6.3 (2.0) (2–11)	7.8 (2.4) (2–13)	5.3 (2.4) (2–10)	8.5 (2.2) (4–15)

SD, standard deviation; Speech perception (monosyllables) (% correct words); speech perception (sentences) (% correct words); SiN perception measured by age-adjusted HINT SRT in noise (dB SNR); language ability measured by The Core Language Index of the Clinical Evaluation of Language Fundamentals (CELF-4) (standard score); non-verbal IQ measured by Raven (standard score); memory span measured by CELF Digit span subtest (scaled score).

were significant differences in age between the TD and CI groups (p = 0.013) and the HA and CI groups (p = 0.004). The other pairwise comparisons had *p*-values equal to 1.0 except for the group comparison between HA and DLD (p = 0.48). Results on the HINT and cognitive tests were adjusted for age, where appropriate, to enable adequate comparison between groups (see section Test Materials and Procedure for details).

Characteristics of the Children in the CI Group

Participants in the wider project were recruited from a clinical population with a wide range of hearing abilities. In this current smaller study, which focused on SiN perception measured by the HINT (sentence repetition), participants were included only if they had hearing abilities which were good enough to repeat sentences presented in quiet. Consequently, the children with CI included in the present study were those with relatively good speech perception, and they are thus not representative of the whole pediatric CI population.

The onset of hearing loss was reported, in medical journals, as before 12 months of age for the majority (81%) of children with CI (see **Table 2**). All children received their first CI (in either sequential or simultaneous bilateral implantation) between November 2002 and December 2014. Amongst the children whose onset of hearing loss was prior to 12 months, 24% were implanted by 12 months of age, 37% were implanted between 1 and 2 years of age, and 39% were implanted after the age of 2 years. There were 26 children who were implanted after 3 years of age, and six of these had an onset of hearing loss prior to 12 months of age. The remaining children either had normal hearing or some residual hearing after birth. In the present study the youngest children to receive CIs were
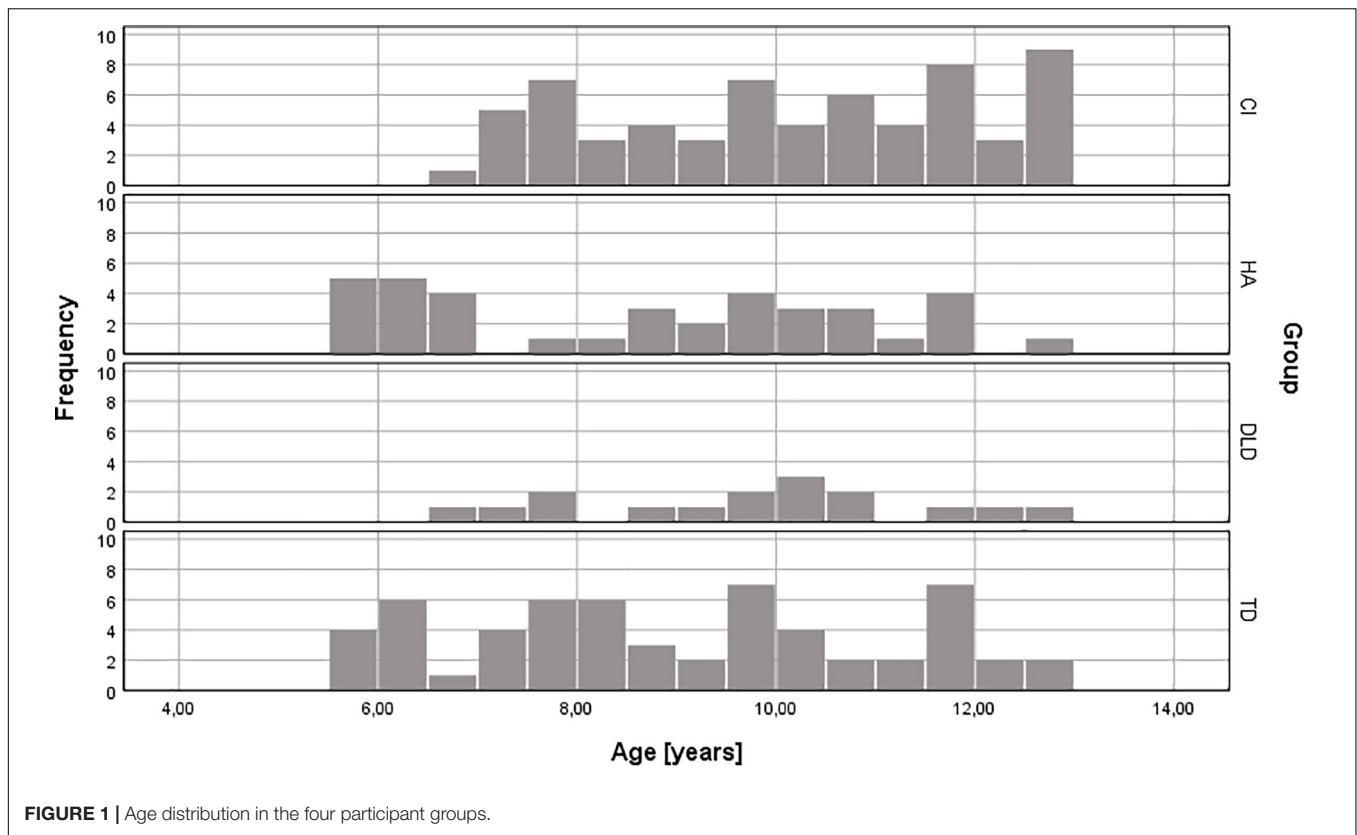
FIGURE 1 | Age distribution in the four participant groups.

5 months old. To investigate the effect of implantation age, the participants in the CI group were classified as either (1) having acquired oral language before implantation ($n = 18$) or (2) having acquired no or very little oral language before implantation ($n = 46$). This classification was made based on medical journals and parental report. It was considered relevant to investigate the effects of implantation age only for children in the second subgroup.

In Norway, the cost of CI implantation is covered by the government and bilateral implantation is offered as the standard

TABLE 2 | Characteristics of the CI and HA groups.

	CI group ($n = 64$)	HA group ($n = 37$)
Onset of hearing loss		
<12 months	81%	76%
>12 months or unknown	19%	24%
Communication approach		
Spoken language	80%	78%
Spoken language with some sign support	11%	11%
Spoken language with lots of sign support	1%	–
Missing data	8%	11%
Educational setting		
Mainstream	92%	95%
Hearing unit in mainstream school	3%	3%
School for children with hearing loss	5%	2%

procedure for children under 18 years. All children in this study wore bilateral CIs. Fifty-five percent were implanted simultaneously and 45% were implanted sequentially.

The make of implants used by participants were Cochlear (55%) and MED-El (45%). Among the 35 children using Cochlear devices, 89% (31 children) were fitted bilaterally with the same model: CI24R/RE (27 users) and CI512 (four users). The remaining four children with Cochlear devices had different models implanted in the right versus the left ear with a combination of CI24RE and CI422 (one user), CI24R/RE and CI512 (two users), CI513 and CI512 (one user). Among the 29 children using MED-EL devices, 69% (20 children) were fitted bilaterally with the same model: C40 + (one user), CONCERTO FLEX24 (one user), PULSARci100 (11 users) and with SONATti100 (seven users). The remaining nine children with MED-EL devices had different models implanted in the right versus the left ear with a combination of PULSARci100 and C40+ (five users), PULSARci100 and CONCERTO (one user), and PULSARci100 and SONATti100 (three users). The use of different models in the right versus the left ear can be explained by the date of implantation, and that the newest available CI model was implanted if the child received sequential CIs or in cases of reimplantation.

The cause of hearing loss was identified for 67% of the children with CI, with a genetic abnormality for the Connexin 26 protein being the most common etiology (29% of the children). Other common causes of hearing loss were Pendred or LVAS (Large Vestibular Aqueduct Syndrome) (16%) and meningitis

infection (19%). For communication approach and educational settings of participants with CI, see **Table 2**.

Characteristics of the Children in the HA Group

The onset of hearing loss was reported, by parents, as before 12 months of age for the majority (76%) of children (see **Table 2**). According to the descriptors recommended by the World Health Organization (WHO), two children (5%) had a mild hearing loss, 27 (73%) had a moderate hearing loss, 7 children (19%) had a severe hearing loss, and 1 child (3%) had profound hearing loss in the better hearing ear. The majority of children (87%) had symmetric hearing loss with a difference of less than 10 dB hearing threshold levels between the ears according to Pure Tone Average (PTA) (measured at frequency 500, 1000, 2000, and 4000 Hz). Five children had PTA differences of 10, 11, 17, 22, and 60 dBHL respectively. All children used bilateral HAs. The children were tested in conjunction with their regular hearing device checkups at the Ear-, Nose and Throat clinic at their local hospital. The assessment session was carried out after the device checkup and adjustment, thus securing that the child had well-functioning hearing devices at the time of assessment. For communication approach and educational settings of participants with HA, see **Table 2**.

Characteristics of the Children in the DLD Group

The participants with DLD comprised a clinical sample which was recruited from the educational and psychological counseling service in municipalities across Norway. This service has the responsibility for the assessment and counseling for children with developmental difficulties in Norway. In addition to the general inclusion criteria 1–5 reported above, two additional inclusion criteria for the DLD group were, (6) referral to the educational and psychological counseling service for language difficulties, and to independently confirm the status as developmentally language disordered, (7) scores 1 SD or more below the normative mean on at least two out of the following five standardized tests: The British Picture Vocabulary Scale II (BPVS-II; Dunn et al., 1997; Lyster et al., 2010), the children's test of non-word repetition (Gathercole et al., 1994; Norwegian version by Furnes and Samuelsson, 2009), and three subtests from the Clinical Evaluation of Language Fundamentals (CELF-4; Semel et al., 2003): Concepts and Following Directions, Formulated Sentences, and Recalling Sentences. One child who was recruited in the DLD group completed all standardized tests, but did not satisfy criteria 7 and so was excluded, leaving 16 children with DLD in the sample. All the children were fully integrated into mainstream school education.

Characteristics of the Children in the TD Group

All of the children with TD were recruited from mainstream educational settings across Norway. Children in this group were defined as 'typically developing' by their teachers.

Test Materials and Procedure

Standard scores which were derived from age-based norms were used for language and cognitive tests, and age norms were used to adjust HINT scores based on normative data (see description

in section Speech-in-Noise Perception). The three tests of speech perception were always administered in the same order: (1) perception of monosyllables in quiet, (2) perception of sentences in quiet, and (3) perception (of sentences) in noise. Except for this, the order of the tests was randomized. All tests were administered in a one-to-one setting in a quiet room. All the participants with hearing loss wore their hearing devices bilaterally during the entire test session, including both the speech perception tasks and the cognitive/linguistic tasks.

Speech Perception in Quiet: Monosyllables

The Norwegian Phonetically Balanced word lists consisting of 50 monosyllabic words each (Øygarden, 2009) were presented from a speaker 2 m in front of the listener at 65 dBA. The main objective of this test was to assess the children's ability to discriminate speech sounds. The monosyllables are Norwegian words with a high usage frequency, but as they are monosyllabic, they are difficult to guess if not all of the speech sounds are identified correctly. Some of the words differ from other Norwegian words by a single phoneme. The child's score represented the percentage of words that were repeated correctly.

Speech Perception in Quiet: Sentences

The HINT sentences (for description, see Speech-in-Noise Perception below) in quiet were presented at 65 dBA. The number of words the child was able to repeat accurately were counted to calculate a percentage score of words in sentences. The sentence repetition test in quiet served two purposes: (1) to measure speech perception in quiet, and (2) as a pretest to the HINT (in noise) to familiarize the child with a test which requires repetition of sentences.

Speech-in-Noise Perception

SiN perception was assessed with the Norwegian HINT for children (NHINT-C). For the sake of brevity, we refer to the NHINT-C as HINT in this paper. The adaptive procedure described in Soli and Wong (2008) was used to estimate the speech reception threshold (SRT) in speech-shaped noise fixed at 60 dBA. The SRT was defined as the mean SNR at which the listener could repeat 50% of the sentences correctly (ignoring inflectional errors and additional words). The SiN performance was evaluated under the condition where speech and noise came from a speaker (Sony SS-MB150H) one meter in front of the participant. Participants were presented with two lists each composed of ten sentences that they were asked to repeat. The speech levels were adjusted for each sentence depending on whether the previous sentence was repeated correctly or not (thereby the name adaptive procedure). The SRTs of the two lists were averaged. HINT SRTs were adjusted for age and room effects as described below, to yield the final measure of SiN perception.

The principle behind the HINT is that the sentences used in the test should be short, the syntax simple, and the vocabulary familiar even to preschool children (Soli and Wong, 2008). Thus, the linguistic and memory demands of the sentences are assumed to be minimal. As reported by Myhrum et al. (2016) the sentences used in the Norwegian child version were a subset

(120 sentences) of the 240 sentences of the Norwegian HINT for adults. The sentences included in the child version were selected in a two-step process: First five adult raters, including three pediatric speech and language pathologists selected 158 of the 240 sentences, which were judged to be comprehensible and repeatable for 5–6-year-old children. Second, the sentences were tested on 11 TD children aged 4.8–5.6 years. The 120 sentences with the highest accuracy scores were identified and divided into 12 phonemically matched 10-sentence lists. Trial-and-error was used to adjust the composition of the lists to obtain the closest match of their phoneme distributions to the overall distribution. The average length of the 120 included sentences was 5.2 words (SD = 1.0, range 3–8).

In the previous study by Myhrum et al. (2016) with typically hearing children from 5;6 years, all the children who were tested with the sentences in quiet performed at ceiling. As a rule of thumb in clinical practice, the word score of sentences in quiet must be above 75% for the child to be tested with the sentences in noise.

For children with typical hearing, HINT results depend on age. In order to know how a child performed compared to a population of normal hearing children of the same age, the normative mean value of his or her age group was subtracted from each HINT SRT. Norwegian HINT normative data across ages from 5;6 to 13;0 years of age were reported in a study described by Myhrum et al. (2016), and the linear regression equation for the age-specific correction for age 5;6 to 10;5 years was used to calculate age-specific correction factors in the current study. HINT results of children older than 10;4 years were not adjusted for age. By adjusting for age, the age effect observed in the normative HINT SRT material is taken into account, and the age-adjusted SRTs can be used as in analyses together with standard scores from the other tests used in the study.

Participants were tested in mainly one room, the anechoic chamber used in the normative study (Myhrum et al., 2016). However, due to large geographical distances from the clinic, 21 participants were tested in two other audiometric test rooms. Since HINT results obtained in a sound field will be influenced by room acoustics, Nilsson et al. (1996) proposed age-specific correction factors relative to adult performance to allow comparison across different sound fields. This was described as a five step procedure in Vaillancourt et al. (2008): (1) obtain adult norm in sound field A, (2) obtain age-specific normative data in sound field A, (3) calculate age-specific correction factors from step one and two, (4) obtain adult norm in sound field B, and (5) calculate age-specific norms for sound field B which are the sums of the age-specific correction factors (3) and the adult norm (4).

The anechoic chamber used in the normative study was defined as sound field A with HINT adult norm −3.9 dB SNR, and the two other audiometric rooms were defined as sound field B1 and sound field B2. The adult norm in sound field B1 was calculated as the average SRT (−2.6 dB SNR) obtained from ten normal hearing adults. Five children from the TD group and eight from the HA group were tested in sound field B1, and their HINT SRTs were corrected for the room effect by −1.3 dB. The adult norm in sound field B2 was not collected, and the HINT SRTs were not corrected for the room effect. Eight children from the HA group were tested in sound field B2.

Language Ability

General language ability was measured by the Norwegian adaptation of the Clinical Evaluation of Language Fundamentals 4 (CELF-4; Semel et al., 2003). CELF-4 is a comprehensive test of language skills, consisting of 13 subtests measuring different aspects of expressive and receptive language as well as verbal memory. There are two slightly different versions of the CELF: one for children aged 5;5–8;9 years and one for children aged 9–12;11 years. The Core Language Index (CLI) is the main index of the test and is intended to be a measure of general language ability that can be used to make decisions about whether a child has a language disorder or not. The CLI for children aged 5;5–8;9 years comprises the following four subtests: Concepts and Following Directions, Word structure, Recalling Sentences and Formulated sentences. The CLI for children aged 9;0–12;11 years comprises the same subtests except that Word Structure has been replaced with Word Classes 2 Total. The Concepts and Following Directions subtest measures the child's ability to follow oral directions of increasing length and complexity by pointing to one or more pictured objects in the correct order. The Word Structure subtest examines the child's morphological knowledge (e.g., plurals and past tense conjugations) by asking the child to complete orally presented sentences in reference to a picture. In the Recalling Sentences subtest the child is asked to repeat orally presented sentences. In the Formulated sentences task, the child is asked to generate sentences in response to orally presented words and a pictures. In the Word Classes 2 task, the child is given three to four orally presented words and is asked to (1) identify two words among these that go together and (2) explain their relationship. We used the CLI (which is a standard score derived from age norms) in all statistical analyses reported below. The CELF-4 has been normed with a sample of 600 Scandinavian children aged 5;0–12;11 years to give the normal range 86–115.

Non-verbal IQ

General non-verbal IQ was measured by Raven's Colored Progressive Matrices for children aged 5;5–8;11 years and Raven's Standard Progressive Matrices (standard version or plus version) for children aged 9;0–12;11 years (Raven et al., 2004). Both tests consist of a series of visual patterns with one part of the pattern missing. The child is presented with a number of options and is instructed to select the correct part to complete the designs. The standard score for non-verbal IQ (derived from age norms) was used in the analyses reported below. The normal range is 86–115.

Memory Span

Memory span was measured by the digit span subtest from the Norwegian adaptation of the CELF-4 (Semel et al., 2003). The child is asked to repeat sequences of orally presented numbers of increasing length and difficulty, either in the order they are presented or backwards, starting with two numbers in sequence and ending (if stop criteria are not applied earlier) with a sequence of nine numbers. Stop criteria are set at two incorrect repetitions of sequences of the same length and

difficulty. A score of one was given for each correctly repeated sequence and a score of zero for each incorrectly repeated sequence. Scores obtained in the forward and the backwards repetition tasks were summed, and the highest possible score was 30 (16 points for the forward and 14 points for the backwards repetition). The raw score was transformed into a scaled score according to the age-based norm given in the CELF-4 manual. The scaled score from this test was used in the regression and correlation analyses reported below. The normal range was 7–13 with 10 as normative mean and 3 as + −1 standard deviation.

Analysis

The statistical analyses were carried out in SPSS for Windows v.25 (SPSS Inc., 2018). SiN perception scores were adjusted for age using the linear regression of age norms calculated in Myhrum et al. (2016). Standard scores derived from age norms were used for language ability and non-verbal IQ, and scaled scores were used for memory span. Speech perception of monosyllables and sentences in quiet were not adjusted for age as these measures are designed to be mastered by children aged 5–6 years, e.g., in Myhrum et al. (2016) normal hearing children of age from 5;6 years old were tested with HINT in quiet sentences and scored 100%. All variables used in the analyses were therefore expected to be independent of age, since the values represent performance compared to a norming sample of the same age. This means that if 6-year-old and a 10-year-old both obtain a standard score of e.g., 75 for language ability, they are both equally behind their age peers, but the actual language skills of the 10-year-old are more advanced than a the actual language skills of the 6-year-old. A parallel example for the age corrected HINT SRTs is that a 6-year-old with an age-corrected SRT of 2 dB will actually have a higher SRT than a 10-year-old with the equal age-corrected SRT of 2 dB, but the two children perform equally in comparison to their normal-hearing peers.

For monosyllable and sentence perception in quiet, there was a ceiling effect and a small range of variance in the two groups with typical hearing, the TD and DLD groups. One-way analyses of variance (ANOVAs) were used to test for differences between groups in SiN perception, language ability, non-verbal IQ and memory span. *Post hoc* comparisons used the Bonferroni correction. However, the *post hoc* comparisons were also carried out using the Hochberg's GT2 and Games-Howell tests to account for the differences in sample sizes and, on occasion, unequal variances (Field, 2013), but the use of those tests did not change any of the significant findings. Thus only the comparisons using the Bonferroni correction are reported.

To investigate which variables influence SiN perception (measured by SRTs adjusted for age), data were first analyzed in one regression model with all participants where group was added as one of the predictors (using groups as independent binary dummy variables). Pearson correlations were calculated to measure associations between SiN and the independent variables. In addition to investigating predictors to SiN in the full dataset, follow-up multiple regression models were fitted separately to data for children with CI, children with HA and children with TD. Simple linear regression is reported for

the DLD group as the sample size was too small to perform multiple regression analysis. Diagnostic statistics, such as Cook's distance, Mahalanobis distance and the DFBeta statistics, were used to assess the fit of the regression models and identify any influential points that were having any undue influence on the model (Barnett and Lewis, 1978; Cook and Weisberg, 1982; Stevens, 2002).

RESULTS

Group Differences

Speech Perception in Quiet (Monosyllables) for Children With CI, HA, DLD, and TD

Table 1 shows descriptive statistics of the monosyllable scores for children with CI, HA, DLD, and TD. In the TD group, 36 children (62%) scored 100% on the monosyllable perception test, 18 children scored 98% and 4 children scored between 92 and 96%. In the DLD group, the monosyllable perception scores were above 95% for all participants except for two participants who scored 70 and 80% respectively. Thus, the participants in the TD and DLD groups had a close to perfect recognition of monosyllables, compared to participants in the HA and CI groups who scored on average 88 and 87% respectively.

Speech Perception in Quiet (Sentences) for Children With CI, HA, DLD, and TD

Table 1 shows descriptive statistics for the speech perception of sentences in quiet for children with CI, HA, DLD, and TD. In the TD group, fifty children (86%) scored 100% on this test, 7 children scored 98% and 1 child scored 96%. In the DLD group, all scores were above 98% except for two participants who scored 66 and 42% (these were the same two participants who scored 70 and 80% on the monosyllables test). Thus except for those two children, all participants in the TD and DLD groups repeated the ten sentences in the speech perception in quiet test without errors or with only a single error.

In the CI group, 53% (*n* = 34) and in the HA group, 38% (*n* = 14) repeated all sentences correctly, and 89% of the participants in the HA and CI group had a score of 90% correct or better. This leaves only 11% (7 in the CI group and 4 in the HA group) with scores below 90%.

Speech-in-Noise Perception for Children With CI, HA, DLD, and TD

Figure 2 shows a boxplot of SiN perception scores for children with CI, HA, DLD, and TD with outliers displayed as dots. On average, children with CI had 2.0 dB higher SRTs than children with HA, 3.4 dB higher SRTs than children with DLD and 5.5 dB higher SRTs than children with TD (see **Table 1**). A one-way ANOVA showed a significant effect of group on SRT, $[F(3,171) = 59.4, p < 0.001, \eta^2 = 0.51]$. *Post hoc* comparisons, using the Bonferroni correction, revealed that there was no difference between the HA and DLD groups ($p = 0.20$) and significant differences in mean SRT between the other groups $[p < 0.001$ for all, except for between the TD and DLD group ($p = 0.01$)].

FIGURE 2 | Boxplot of age-adjusted HINT SRTs for children with CI, HA, DLD, and TD. The boxes go from the first quartiles to the third quartiles. Outliers were defined using Tukey (1977) i.e., as any data points more than 1.5 times the interquartile range from the quartiles. The whiskers go from the smallest non-outlier to the largest non-outlier.
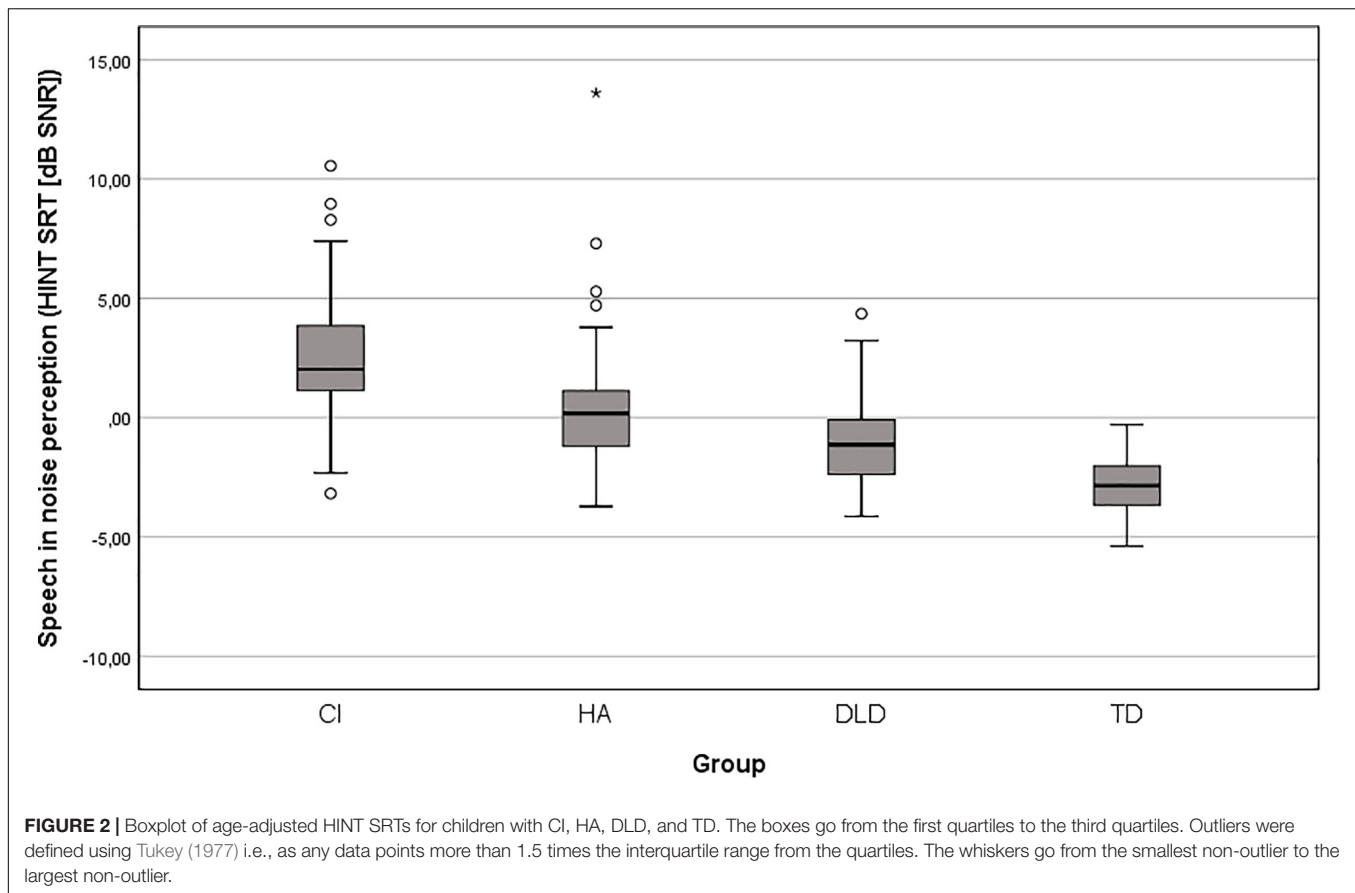
Language Ability for Children With CI, HA, DLD, and TD

Table 1 shows descriptive statistics for the CELF CLI for children with CI, HA, DLD, and TD. Seven children in the TD group scored below the normal range for the CLI, and two children in the DLD group scored within the normal range on this index. As the definition of DLD relies on a deficit in functional language ability, e.g., affecting everyday and school functioning and a complete clinical assessment, the TD children were not excluded despite being outside of the normal range, as they had been defined as 'typically developing' by their teachers. The result of one assessment is not sufficient to confirm language difficulties in these children. Similarly, the DLD children who scored within the normal range on the CLI were not excluded, as they still met criteria 6 and 7 (see section Characteristics of the Children in the DLD Group). Furthermore, in the HA group 4 children (11%) scored below -2 SD from the normative mean, and in the CI group 24 children (38%) scored below -2 SD from the normative mean. There was a significant effect of group on language ability $[F(3,171) = 34.80, p < 0.001, \eta^2 = 0.38]$. On average, children with TD scored within the normal range for the CLI (86–114) (Semel et al., 2003), children with HA scored just below this range, children with CI scored on average below -1 SD from the normative mean, and children with DLD scored on average below -2 SD from the normative mean. *Post hoc* comparisons, using the Bonferroni correction, revealed that there were significant

differences in language ability between all groups at the $p < 0.001$ level, except between the HA and CI groups ($p = 0.049$) and between the CI and DLD groups ($p = 0.29$).

Non-verbal IQ for Children With CI, HA, DLD, and TD

Table 1 shows descriptive statistics of the non-verbal IQ scores for children with CI, HA, DLD, and TD. There was a significant effect of group $[F(3,169) = 4.7, p = 0.004, \eta^2 = 0.08]$. *Post hoc* comparisons, using the Bonferroni correction, showed that there was a significant difference of 13.0 in non-verbal IQ scores between the TD and DLD groups ($p = 0.004$). There was no significant difference in mean non-verbal IQ scores between any other groups at $p < 0.05$ (TD and HA ($p = 0.42$), TD and CI ($p = 0.09$), HA and DLD ($p = 0.29$), HA and CI ($p = 1.0$), CI and DLD ($p = 0.36$)].

Memory Span for Children With CI, HA, DLD, and TD

Table 1 shows descriptive statistics for memory span for children with CI, HA, DLD, and TD. A one-way ANOVA showed a significant effect of group on digit span scores $[F(3,171) = 16.7, p < 0.001, \eta^2 = 0.23]$. *Post hoc* comparisons, using the Bonferroni correction, showed that there were significant differences ($p < 0.005$) in mean digit span scores in all pairwise group comparisons, except no significant differences between the DLD and CI groups ($p = 0.60$) or between the TD and HA groups ($p = 0.61$).

	Speech in noise perception (SiN)	Speech perception (sentences)	Speech perception (monosyllables)	Language ability	Non-verbal IQ	Memory span
Speech in noise perception (SiN)	1	−0.47** $p < 0.001$	−0.60** $p < 0.001$	−0.60** $p < 0.001$	−0.27** $p < 0.001$	−0.33** $p < 0.001$
Speech perception (sentences)	−0.47** $p < 0.001$	1	0.40** $p < 0.001$	0.45** $p < 0.001$	0.16* $p = 0.037$	0.23** $p = 0.002$
Speech perception (monosyllables)	−0.60** $p < 0.001$	0.40** $p < 0.001$	1	0.34** $p < 0.001$	0.19* $p = 0.011$	0.17* $p = 0.027$
Language ability	−0.60** $p < 0.001$	0.45** $p < 0.001$	0.34** $p < 0.001$	1	0.44** $p < 0.001$	0.58** $p < 0.001$
Non-verbal IQ	−0.270** $p < 0.001$	0.16* $p = 0.037$	0.19* $p = 0.011$	0.44** $p < 0.001$	1	0.29** $p < 0.001$
Memory span	−0.33** $p < 0.001$	0.23** $p = 0.002$	0.17* $p = 0.027$	0.58** $p < 0.001$	0.29** $p < 0.001$	1

*p < 0.05, **p < 0.01. Speech-in-noise perception measured by age-adjusted HINT SRT in noise [dB SNR)]; speech perception (sentences) measured by sentences in quiet (% words correct); speech perception (monosyllables) measured by % correct words; language ability measured by The Core Language Index of the Clinical Evaluation of Language Fundamentals (CELF-4) (standard score); non-verbal IQ measured by Raven's matrices (standard score); memory span measured by CELF Digit span subtest (scaled score of forward and backward digit span combined).

Regression Analyses

Since we used standardized scores derived from age-norms for the independent variables and age-corrected scores (based on age norms) for the dependent variable, age was not included in the regression models. The regression model was first run on the full sample of 175 children. Subsequently, the effect of the predictor variables were explored by running linear models for each of the separate groups.

Regression Analysis Using Data From All Groups

Table 3 reports correlations among all variables in the full sample. In the regression model predicting SiN perception with only group and language ability as predictors, the model explained 60% ($R^2 = 0.60$) of the variance ($F = 64.5$, $p < 0.001$). The mean values for SiN perception for each group changed when adjusting for language ability, and the mean values for SiN perception were no longer significantly different between the DLD group and the TD group ($p = 0.87$). The regression coefficient of language ability was −0.061 ($p < 0.001$) when included in the model together with the group variable. When speech perception in quiet (monosyllables) was added to the model, it explained 65% ($R^2 = 0.65$) of the variance in SiN perception, and the regression coefficients of language ability and monosyllable perception were significant predictors and equal to −0.057 and −0.094 respectively. This means that a 10 point increase in language ability was associated with a 0.6 dB decrease in SiN and that a 10 point increase in speech perception in quiet (monosyllables) was associated with a 0.9 dB decrease in SiN. When adding non-verbal IQ and memory span to the model, this did not explain any more of the variance in SiN perception. **Table 4** shows the results of the multiple linear regression model for SiN perception with all the explored variables included. The model accounted for 64% of the variance.

To further investigate the relationship between language ability and SiN perception, the data points were plotted in a scatter plot where each point represented an individual's language ability on the *x*-axis and SiN perception on the *y*-axis. **Figure 3** shows a scatterplot of SiN perception versus language ability with regression lines for the four groups. The scatterplot shows that SiN perception scores for the CI group were poorer than for the DLD group, but both groups showed a small-medium linear effect of language ability on SiN perception. The HA group showed a similar linear effect. However, the regression line may have been influenced by two outliers as described in section "Factors That Predict Speech in Noise Perception in Children With HA." The regression line for the TD group showed 0 slope which means that there was little to no effect of language ability on SiN perception. The scatterplot also indicates that language ability had more effect on speech perception in noise when language ability was lower than approximately 85, which is −1 SD below the normative mean, compared to when language ability was above 85. However, when modeling SiN perception in the linear regression model, we made the assumption that language ability linearly predicted SiN perception in the language ability interval of interest (from 42 to 135).

TABLE 4 | Linear model of predictors of speech-in-noise perception for the full group of participants (*n* = 175).

	B [95% CI]	*SE* B	β	*p*
Constant	12.26 [7.85, 16.7]	2.23		<0.001
Group difference from TD group	HA: 1.54 [0.57, 2.51]	0.49	0.19	0.002
	DLD: −0.29 [−1.57, 0.99]	0.65	−0.03	0.66
	CI: 2.87 [1.91, 3.83]	0.49	0.43	<0.001
Language ability	−0.057 [−0.075, −0.039]	0.009	−0.36	<0.001
Speech perception (monosyllables)	−0.094 [−0.14, −0.052]	0.021	−0.26	<0.001
Non-verbal IQ	−0.007 [−0.031, 0.016]	0.012	−0.031	0.550
Memory span	0.027 [−0.12, 0.17]	0.075	0.021	0.720
Adjusted R^2 = 0.64				

Forced entry method of regression was used.

FIGURE 3 | Scatterplot of HINT SRTs versus CELF-4 Core Language Index for children with CI (yellow), HA (red), DLD (blue) and TD (green). The solid lines are linear regression lines for groups CI, HA, and DLD ($p \leq 0.005$). The green dashed line is a non-significant linear regression line ($p = 0.29$) for the TD group.
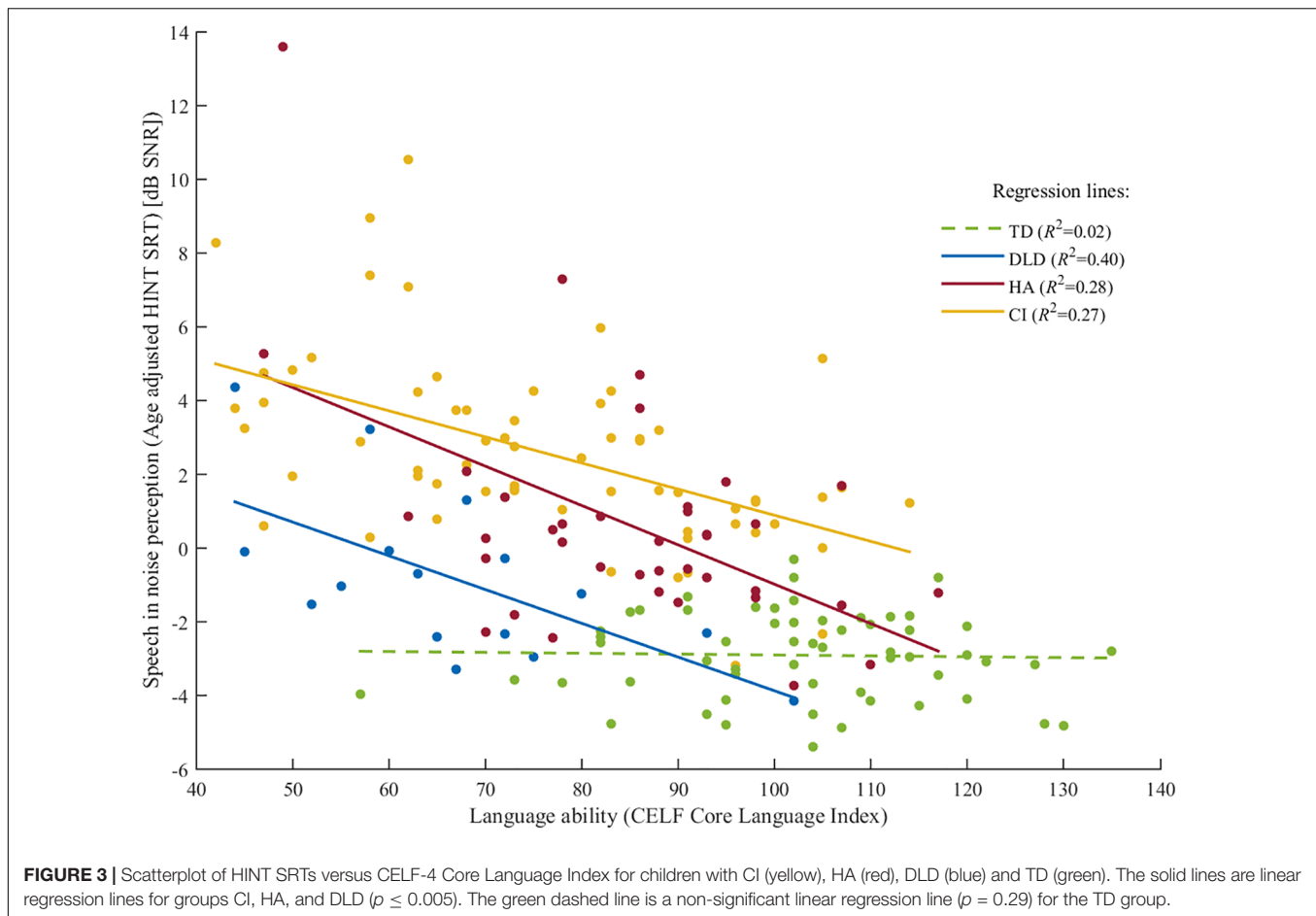
Factors That Predict Speech in Noise Perception in Children With CIs

For the CI group, the variable age at implantation was investigated in addition to the other independent variables for the subgroup of children who had not acquired oral language before implantation (see description in section Characteristics of the Children in the CI Group). **Table 5** reports correlations among variables for the whole CI group, and in addition correlation with implantation age for the subgroup who had not acquired oral language before implantation. In the full CI group, all variables except non-verbal IQ and memory span were significantly correlated with SiN perception.

A model was fitted for the CI group with SiN perception as the dependent variable including all the five independent variables. When non-verbal IQ and memory span were removed from the model (p-values 0.5 and 0.9 respectively) the model explained one percentage point less of the variance. **Table 6** shows the results of the latter multiple linear regression model which accounted for 50% of the variance. For the subgroup who had not acquired language before implantation, the same predictors were used in a model together with implantation age. This model is also reported in **Table 6**, and accounted for 55% of the variance in SiN.

There were significant correlations between some of the independent variables, e.g., between language ability and both non-verbal IQ and memory span (see **Table 5**), but correlations were weak to moderate, and the variance inflation factors for all predictors were below 2.1, which suggest no threat of multicolinearity (Hair et al., 1995). Diagnostic statistics revealed one potential influential data point. This data point had large leverage and Mahalanobis distance values, suggesting undue influence on the model (Barnett and Lewis, 1978; Stevens, 2002). Cook and Weisberg (1982) suggest that a Cook's distance value greater than 1 is of concern. This data point did not exceed 1. Further inspection of this data point revealed that this child had a low language ability score. However, this data point was not an outlier, i.e., not smaller than 1.5 times the interquartile range, and removing this data point did not substantially change the coefficients or the significance of the predictors. Without the data point the model accounted for 4 percentage points less of the variance. The data point was thus included in the multiple linear regression model presented in **Table 6**.

Factors That Predict Speech in Noise Perception in Children With HA

Table 7 reports correlations among all variables for children with HA. All variables except memory span correlated with SiN perception. However, all variables were initially included in a regression model which accounted for 59% of the variance

TABLE 5 | Correlations among variables for children with CI (*n* = 64) and correlations with implantation age for subgroup who had not acquired language before implantation (*n* = 46).

	Speech in noise perception (SiN)	Speech perception (sentences)	Speech perception (monosyllables)	Language ability	Non-verbal IQ	Memory span
Speech in noise perception (SiN)	1	−0.67** $p < 0.001$	−0.28* $p = 0.03$	−0.52** $p < 0.001$	−0.06 $p = 0.63$	−0.12 $p = 0.34$
Speech perception (sentences)	−0.67** $p < 0.001$	1	0.13 $p = 0.30$	0.51** $p < 0.001$	0.07 $p = 0.61$	0.08 $p = 0.54$
Speech perception (monosyllables)	−0.28* $p = 0.03$	0.13 $p = 0.30$	1	0.07 $p = 0.57$	−0.057 $p = 0.66$	−0.13 $p = 0.32$
Language ability	−0.52** $p < 0.001$	0.51** $p < 0.001$	0.07 $p = 0.57$	1	0.35** $p = 0.006$	0.52** $p < 0.001$
Non-verbal IQ	−0.06 $p = 0.63$	0.07 $p = 0.61$	−0.057 $p = 0.66$	0.35** $p = 0.006$	1	0.37** $p = 0.003$
Memory span	−0.12 $p = 0.34$	0.08 $p = 0.54$	−0.13 $p = 0.32$	0.52** $p < 0.001$	0.37** $p = 0.003$	1
Implantation age (*n* = 46)	0.40** $p = 0.006$	−0.08 $p = 0.59$	−0.04 $p = 0.77$	−0.50** $p < 0.001$	−0.31* $p = 0.04$	−0.30 $p = 0.04*$

*$p < 0.05$, **$p < 0.01$.

with speech perception in quiet as the only significant predictor. When non-verbal IQ and memory span were removed from the regression analyses, the model accounted for 3 percentage points less of the variance. In addition, the unique variance explained by speech perception in quiet decreased by 3%, while the unique variance explained by speech perception in quiet (monosyllables) and language ability increased by 9 and 7% respectively. **Table 8** shows the results of the multiple linear regression model of SiN perception with speech perception in quiet (monosyllables), speech perception in quiet (sentences), and language ability as predictors. The model accounted for 52% of the variance. All three predictors significantly influenced the model.

TABLE 6 | Linear model of predictors of speech in noise perception for children with CI.

	B [95% CI]	*SE B*	β	*p*
CI participants, whole group (*n* = 64)				
Constant	33.04 [23.91, 42.17]	4.47		<0.001
Speech perception (monosyllables)	−0.069 [−0.13, −0.01]	0.03	−0.20	0.033
Speech perception (sentences)	−0.23 [−0.32, −0.14]	0.05	−0.52	<0.001
Language ability	−0.033 [−0.06, −0.01]	0.01	−0.25	0.021
Adjusted R^2 = 0.50				
CI participants, subgroup who had not acquired language before implantation (*n* = 46)				
Constant	30.71 [6.00, 21.17]	4.74		0.001
Implantation age	0.83 [−0.24, 1.23]	0.30	0.33	0.009
Speech perception (monosyllables)	−0.074 [−0.17, −0.02]	0.03	−0.57	0.021
Speech perception (sentences)	−0.24 [−0.32, −0.14]	0.05	−0.34	<0.001
Language ability	−0.005 [−0.086, −0.013]	0.02	−0.04	0.782
Adjusted R^2 = 0.55				

Forced entry method of regression was used.

There were significant correlations between some of the independent variables, however only a positive weak to moderate correlation for language ability and both non-verbal IQ and memory span (see **Table 7**), and the variance inflation factors for all predictors were below 2.1 (largest and equal to 2.1 for language ability), which suggest no threat of multicolinearity (Hair et al., 1995).

Diagnostic statistics revealed two potential influential data points. Both data points had large leverage and Mahalanobis distance values. Further inspection showed that one child had a very low monosyllable perception score (48%, which was an outlier) despite having a good speech perception in quiet score for sentences (98%). The other child had very low scores for monosyllable perception (64%), speech perception in quiet (70%), language ability (49) and required a very high SRT (13.6 dB SNR) for SiN perception - all of these scores were outliers. The latter child had a value of Cook's distance that exceeded 1, suggesting undue influence on the model (Cook and Weisberg, 1982). When both of these data points were removed the predictors were no longer significant and the model accounted for only 13% of the variance. **Table 8** shows the impact of removing the two influential points on the regression model.

Factors That Predict Speech in Noise Perception in Children With DLD

Table 9 reports correlations among all variables for the DLD group. SiN perception was strongly related to language ability, but not significantly related to non-verbal IQ or memory span. There was little variation in the measures of speech perception of sentences and monosyllables in quiet [see **Table 1** and descriptions in sections Speech Perception in Quiet (Monosyllables) for Children With CI, HA, DLD, and TD and Speech Perception in Quiet (Sentences) for Children With CI, HA, DLD, and TD]. The ceiling effect and the small variance in results on these two tests prohibited their use as predictors in the regression model.

TABLE 7 | Correlations among variables for children with HA (*n* = 37).

	Speech in noise perception (SiN)	Speech perception (sentences)	Speech perception (monosyllables)	Language ability	Non-verbal IQ	Memory span
Speech in noise perception (SiN)	1	−0.66** *p* < 0.001	−0.46** *p* = 0.004	−0.53** *p* = 0.001	−0.41* *p* = 0.011	−0.23 *p* = 0.18
Speech perception (sentences)	−0.66** *p* < 0.001	1	0.36* *p* = 0.028	0.53** *p* = 0.001	0.19 *p* = 0.25	0.14 *p* = 0.40
Speech perception (monosyllables)	−0.46** *p* = 0.004	0.36* *p* = 0.028	1	0.071 *p* = 0.68	0.36* *p* = 0.027	−0.025 *p* = 0.89
Language ability	−0.53** *p* = 0.001	0.53** *p* = 0.001	0.071 *p* = 0.68	1	0.42** *p* = 0.009	0.47** *p* = 0.003
Non-verbal IQ	−0.41* *p* = 0.011	0.19 *p* = 0.25	0.36* *p* = 0.027	0.42** *p* = 0.009	1	0.32 *p* = 0.053
Memory span	−0.23 *p* = 0.18	0.14 *p* = 0.40	−0.025 *p* = 0.89	0.47** *p* = 0.003	0.32 *p* = 0.053	1

*p < 0.05, **p < 0.01.

Due to the small sample size, multiple regression was not conducted for the DLD group. Simple linear regression was carried out to investigate the relationship between SiN perception and language ability. There was a significant relationship between SiN perception and language ability with slope equal to −0.09 dB per unit change in language ability (*p* = 0.009), and 40% of the variance in SiN perception could be explained by the model containing only language ability. There were two children with DLD who did not have ceiling scores for monosyllable perception and sentence repetition in quiet (the same two children had low scores for both tests). When these two children were removed from the regression analysis, there was still a significant relation between SiN perception and language ability with slope equal to −0.0007 (*p* = 0.03), and 33% of the variance in SiN perception was explained by the model containing only language ability.

TABLE 8 | Linear model of predictors of speech in noise perception for children with HA.

	B [95% CI]	SE B	β	p
HA group (n = 37)				
Constant	32.99 [4.29,42.04]	5.95		0.001
Speech perception (sentences)	−0.21 [−0.38,0.03]	0.08	−0.39	0.012
Speech perception (monosyllables)	−0.085 [−0.15,0.04]	0.04	−0.30	0.024
Language ability	−0.061 [−0.11,−0.00]	0.03	−0.30	0.035
Adjusted *R²* = 0.52				
Two influential points excluded (n = 35)				
Constant	17.09 [0.30,30.84]	7.69		0.032
Speech perception (sentences)	−0.027 [−0.23,0.07]	0.10	−0.05	0.781
Speech perception (monosyllables)	−0.073 [−0.18,0.08]	0.03	−0.36	0.037
Language ability	−0.046 [−0.11,0.00]	0.03	−0.28	0.104
Adjusted *R²* = 0.13				

Forced entry method of regression was used.

Factors That Predict Speech in Noise Perception in Children With TD

Table 10 reports correlations among all variables for the TD group. None of the variables correlated significantly with SiN perception (all had *p* > = 0.25), and thus we did not carry out a regression analysis for the TD group. As **Table 1** shows, there was little variation in speech perception in quiet (monosyllable perception and sentence repetition). Therefore interpretation of correlations is valid only within the very limited range of values for these two scores. For the TD group, all except three participants had language ability standard scores above 80, and thus it should be kept in mind that the non-significant relationship between language ability and SiN perception was observed in this range of normal language ability.

Developmental Trend of Speech in Noise Perception in Children With TD

The HINT SRTs were corrected for (i) the acoustic environment, i.e., anechoic chamber or audiometric test room, and (ii) age (5;6–10;5 years) using the regression equation from Myhrum et al. (2016) with slope −0.69 dB/annum and with 95% CI (−0.84, −0.55). There is some evidence to suggest that children reach adult performance on the Norwegian HINT by 9–10 years of age in test conditions where speech and noise are collocated in front of the listener (Myhrum et al., 2016). However, the ages at which adult performance is reached vary slightly in similar studies using other HINT languages. For the HINT versions for American English (Nilsson et al., 1996) and French Canadian (Vaillancourt et al., 2008), significant differences were found between 10-year-olds and adults (1.5 and 1.0 dB SNR respectively), but not between 12-year-olds and adults, indicating that adult performance was reached between 10 and 12 years of age. In a study using the Words-in-noise test (with monosyllabic words as stimuli) Wilson et al. (2010) found that the recognition performance was stable between ages 9 and 12 years.

To investigate the age effect on HINT in the TD group in the current study, we examined the uncorrected HINT data, calculating the linear regression of HINT versus age. The slope

TABLE 9 | Correlations among variables for group with DLD (*n* = 16).

	Speech in noise perception (SiN)	Speech perception (sentences)	Speech perception (monosyllables)	Language ability	Non-verbal IQ	Memory span
Speech in noise perception (SiN)	1	−0.42 *p* = 0.11	−0.11 *p* = 0.69	−0.63** *p* = 0.009	−0.38 *p* = 0.15	−0.37 *p* = 0.16
Speech perception (sentences)	−0.42 *p* = 0.11	1	0.86** *p* < 0.001	0.38 *p* = 0.15	0.13 *p* = 0.62	0.34 *p* = 0.20
Speech perception (monosyllables)	−0.11 *p* = 0.69	0.86** *p* < 0.001	1	0.32 *p* = 0.23	0.11 *p* = 0.69	0.43 0.099
Language ability	−0.63** *p* = 0.009	0.38 *p* = 0.15	0.32 *p* = 0.23	1	0.62* *p* = 0.011	0.61* *p* = 0.013
Non-verbal IQ	−0.38 *p* = 0.15	0.13 *p* = 0.62	0.11 *p* = 0.69	0.62* *p* = 0.011	1	0.47 *p* = 0.067
Memory span	−0.37 *p* = 0.16	0.34 *p* = 0.20	0.43 0.099	0.61* *p* = 0.013	0.47 *p* = 0.067	1

*p < 0.05, **p < 0.01.

TABLE 10 | Correlations among variables for group with TD (*n* = 58).

	Speech in noise perception (SiN)	Speech perception (sentences)	Speech perception (monosyllables)	Language ability	Non-verbal IQ	Memory span
Speech in noise perception (SiN)	1	0.15 *p* = 0.25	−0.11 *p* = 0.40	−0.03 *p* = 0.83	−0.14 *p* = 0.29	0.08 *p* = 0.58
Speech perception (sentences)	0.15 *p* = 0.25	1	0.13 *p* = 0.32	0.28* *p* = 0.034	−0.05 *p* = 0.70	0.02 *p* = 0.91
Speech perception (monosyllables)	−0.11 *p* = 0.40	0.13 *p* = 0.32	1	0.18 *p* = 0.18	−0.04 *p* = 0.77	−0.04 *p* = 0.79
Language ability	−0.03 *p* = 0.83	0.28* *p* = 0.034	0.18 *p* = 0.18	1	0.32* *p* = 0.02	0.23 *p* = 0.08
Non-verbal IQ	−0.14 *p* = 0.29	−0.05 *p* = 0.70	−0.04 *p* = 0.77	0.32* *p* = 0.02	1	−0.09 *p* = 0.53
Memory span	0.08 *p* = 0.58	0.02 *p* = 0.91	−0.04 *p* = 0.79	0.23 *p* = 0.08	−0.09 *p* = 0.53	1

*p < 0.05.

of the HINT versus age regression for TD children of ages (5;6–10;5 years) was equal to −0.57 dB/annum (*p* < 0.001, explaining 29% of the variance). This is close to the slope used for the age norm correction (−0.69), and is within the confidence interval of the slope estimated in Myhrum et al. (2016). A second finding was that the slope of HINT versus age when including the children above 10 years (5;6–12;5 years) was less steep (−0.36). Furthermore, in the current study, the mean HINT SRTs for 10–12-year-olds in the TD group were approximately the same (10 years: *n* = 6, *m* = −2.76 dB, 11 years: *n* = 9, *m* = −2.19 dB, 12 years: *n* = 4, *m* = −2.49 dB). These sample sizes are too small to draw robust conclusions, but support claims that the developmental trend seen in the Norwegian HINT perception trails off by 10 years of age.

The age effect was further investigated by calculating the correlation between the age-adjusted HINT SRTs and age in the TD group. We found a weak positive correlation (*r* = 0.28, *p* = 0.03). A linear regression between age and the age-adjusted HINT gave a slope equal to 0.16 dB/annum. This weak positive correlation may indicate that in the TD group, the HINT SRTs of the younger children may have been somewhat overcorrected by applying the norm reported in Myhrum et al. (2016).

DISCUSSION

The present study investigated SiN perception in four groups of children: children with CI, HA, DLD, or TD. We aimed to identify the differences in performance on the HINT and to investigate which cognitive and linguistic factors predict SiN perception for these children.

Group Differences in Speech-in-Noise Perception

As we would expect, children with TD had, on average, the best SiN perception. There was a reliable difference between all groups in SiN perception ability except between the HA and DLD groups. Consistent with past literature, these findings show that children with permanent hearing loss and DLD exhibit speech perception deficits in noise (Ziegler et al., 2011; Misurelli and Litovsky, 2012; Nittrouer et al., 2013; McCreery et al., 2015). The finding that children with HA and CI require a higher SNR to perceive speech in noise should encourage educational settings to improve the SNR in the environment for these children. Through assistive listening device technology such as FM radio signal, infrared

light, and induction loop systems, children with HA and CI can better access speech in background noise. For children who do not use personal hearing devices, like children in the DLD and TD groups, classroom sound field systems can be used to help them listen in less-than-ideal conditions.

It should be kept in mind that the CI group included in the present study was not representative of the pediatric CI population as a whole, but was composed of those children who had relatively good speech perception in quiet. Thus, differences between children with CI and the other three groups, including children with HA, would likely have been substantially larger if children with CI with poorer performance on speech perception tasks in quiet had also been included. However, inclusion of this group of children with CI would have required the use of a different test to measure SiN perception, as the HINT standard adaptive procedure would likely have been too demanding.

While the DLD group was too small (n = 16) to draw robust conclusions, the findings represent preliminary evidence that some children with DLD may underperform not only on SiN perception tasks that assess fine phonological contrasts through monosyllable repetition (Ziegler et al., 2005, 2011), but also tasks that use simple sentences as stimuli. Our results differ from those of Ferguson et al. (2011) who found no difference between children with DLD and TD peers on a sentence repetition in noise task. One possible reason for this difference may be the scoring method. The sentences used by Ferguson et al. (2011) were based on materials and a scoring method reported by MacLeod and Summerfield (1990). Three of the words in each sentence were designated as keywords, and a correct score was given if these keywords were repeated. By contrast, in the present study, it was required that all words in the sentence (approximately 5 on average) were repeated correctly. Although almost all children with DLD were at ceiling when repeating the HINT sentences in quiet, the double demands of noisy conditions and the number of words that had to be repeated may have contributed to the deficit compared to TD peers on this task.

Factors Predicting Speech-in-Noise Perception

Whereas there was no relation between SiN perception and language ability in the TD group, language ability predicted unique variance in SiN perception for children in all three clinical groups, even when speech perception in quiet (monosyllable perception) was taken into account. This finding of a relation between SiN perception and language ability is in line with previous work on children with hearing loss (e.g., Ching et al., 2018). While the current study cannot determine the mechanisms responsible for the relation between SiN performance and language ability, we can think of several possible reasons for the observed association. One possibility is that the language demands posed by the HINT sentences were simply too high for the children with CI, HA, and DLD, despite efforts to keep the stimulus sentences at a level that was easily comprehensible and repeatable for 5-year-olds. However, approximately 90% of children in the HA and CI groups had a score of 90% correct or more on the sentence repetition task in quiet, and only two out

of 16 children with DLD had below-ceiling performance in quiet. Thus, it appears that for the great majority of children in all three clinical groups, the vocabulary, grammar and length of the HINT sentences were manageable under optimal listening conditions.

However, it is possible that more difficult listening conditions involving noise require more robust linguistic knowledge, as the matching between input and linguistic memory representations has to be completed with only partial information. If phonological or lexical representations are less detailed or unstable in children with hearing loss or DLD, or the ability to suppress competing lexical candidates is deficient, the degraded auditory input may not be sufficient for activating the correct lexical items. Grammatical skills may also be needed to supplement word recognition under difficult conditions, e.g., by proving information about the likely word class of an upcoming word. Additionally, if recognition processes are not able to settle on a word, this may have cascading effects for recognition of subsequent words in the speech stream. Thus, children with hearing loss and DLD may have language representations and processing mechanisms which suffice in optimal conditions (with simple sentences), but which are not robust enough to support efficient SiN perception. This interpretation is in line with previous studies suggesting that language knowledge can counteract the consequences of deficits in supra-threshold auditory processing tasks as it allows participants to better 'guess' the words in the sentences based on regularities and context (Bradlow et al., 2003; Sperling et al., 2005; Conway et al., 2014).

Another possible explanation for the observed association between SiN perception and language ability in children with hearing loss may be that children who have hearing-in-noise deficits get less and poorer quality language input in a number of everyday situations which are typically noisy, such as preschool and school, and therefore pick up less language. In other words, the hearing-in-noise deficit may be a cause of poor language skills. The difference in input between children with hearing loss and peers with typical hearing may be especially prominent in third-party learning situations, i.e., when the language is not addressed directly to the child, but to another person in the child's surroundings. A number of experimental studies of TD children show that they can learn words through listening in on others' conversations (for a review, see Akhtar et al., 2019), but this may be substantially more difficult for children with hearing loss, and especially under noisy conditions.

Although the language deficit displayed by children with CI or HA in this study may be traced back to poor audibility and phonetic discrimination, it still appears to pose an additional challenge when attending to speech in noise. Nittrouer et al. (2013) found that phonological sensitivity explained a significant amount of between-groups variance in SiN perception for children with HA, children with CI and children with typical hearing, and thus conclude that "it is not enough to focus only on ways to improve the acoustic environment; their language abilities also must be considered" (p. 523). Our findings are consistent with the view that interventions designed to help children with hearing loss develop good language skills could potentially be an effective way to improve their capabilities to handle noisy school environments. Examination of whether

gains in language skills resulting from language intervention are coupled with gains in SiN perception could also help determine whether better language skills are causally related to better SiN perception.

A clinical implication of the robust relation between HINT performance and language ability in children with hearing loss and DLD is that a full interpretation of HINT results for children in these groups should be made in conjunction with an assessment of the child's language skills.

The fact that language ability was not a significant predictor of SiN for the TD children in the present study may suggest that the linguistic load of the HINT sentences was low, even in demanding processing conditions, for TD school age children. However, language ability has been shown to predict speech perception in noise in other studies of normal hearing participants. For example, in a study by MacCutcheon et al. (2019), SiN perception was better in the 50% of participants who had the best expressive language scores. As sentence repetition is one of the best measures of individual differences in language ability (Klem et al., 2015), it is possible that language ability would have come out as a significant predictor also in the present study if sentences had been linguistically more challenging, e.g., using less frequent words or complex syntax.

For the full sample of participants, there was a moderate and significant correlation between SiN perception and memory span, as measured by the composite of forward and backward digit span, but this correlation was only about half the effect size of the correlation between SiN and language ability. Memory span was not a significant predictor of SiN perception when language ability and speech perception in quiet were taken into account. The finding of a significant association between memory span and SiN is in line with a number of previous studies of children (MacCutcheon et al., 2018, 2019, submitted). Still, our results suggest that for children with hearing loss and language disorder, general language ability may be more closely related to SiN perception. This is evidenced by the fact that for children with CI, HA and DLD, when seen as separate groups, there was a strong and significant bivariate correlation between SiN perception and language ability, but no significant correlation with memory span. This pattern of findings may partly be due to the language measure being more robust, as it represents a composite score from four comprehensive subtests, while the memory span was composed only of two subtests. Another possible reason for the relatively weak relationship between memory span and SiN performance in the present study, was that the sentences used in the HINT are relatively short. Moreover, the sentences mostly describe well known scenarios and thus allow the participant to use linguistic context and world knowledge to compensate for memory limitations. By contrast, the SiN task used by MacCutcheon et al. (2018; 2019, submitted) where all sentences follow the same template with some items (colors and numbers) replaced in each sentence, does not allow for use of linguistic context or world knowledge.

As expected, speech perception in quiet, measured by sentence repetition and monosyllable repetition, was related to SiN performance for children with CI and HA. However, most children, even in the two groups of children with hearing loss, had near ceiling performance on the HINT sentence repetition test. For children with CI and HA monosyllable repetition scores had a bell-shaped distribution around the average score of 87–88%, indicating that even if sentences could be repeated without errors, discrimination of monosyllabic words without a linguistic context was challenging. The ceiling effects for the monosyllable repetition in the TD and DLD groups can be explained by the fact that real (and frequent) words were used. Had non-words been used instead of real words, performance on these tests would possibly have explained more variance in SiN performance, especially for younger children.

For the subgroup of children who had a congenitally profound to severe hearing loss and who did not acquire spoken language before CI ($n = 46$), implantation age significantly predicted SiN perception above and beyond speech perception in quiet scores. This finding is in line with evidence from previous research on SiN perception in children with CI (Ching et al., 2018).

When using HINT SRTs which were not corrected for age, age was a significant predictor of speech perception in noise for the TD children, and the estimated developmental trend was quite similar to the developmental trend estimated in the paper presenting the Norwegian HINT normative data (Myhrum et al., 2016). The developmental trend in SiN perception is also consistent with previous research (Jamieson et al., 2004; Neuman et al., 2010; Wilson et al., 2010).

Limitations and Future Directions

In the current study we investigated predictors of SiN perception both in the full sample of 175 children and separately for each of the four participant groups. While the group-specific analyses were important for comparisons with previous studies of these groups, it should be acknowledged that a large number of statistical comparisons were carried out, thus increasing the probability of erroneous inferences. In addition, the study spanned a wide age-range (from 5;6 to 12;11 years), and while age norms were used, these norms may have been better suited for children with TD than for clinical samples, as the norming samples typically have few children at the lower tail of the distribution. The DLD sample in the present study was small, and thus we cannot draw robust conclusions about SiN perception in this group.

Another limitation of the study was the ceiling effects on the tests of speech perception in quiet in the TD, DLD (ceiling effects for both sentences and monosyllables) and, to some extent the HA and CI groups (ceiling effects for sentences), which made it difficult to assess the predictive value of speech perception in quiet for SiN perception. A nonsense word repetition test may have given a more fine-grained and better distributed measure of speech perception in quiet (for an overview of advantages of nonsense word repetition tests to assess speech perception, see Rødvik et al., 2018). Additionally, we only used OAEs (in combination with parent report of normal hearing) to assess hearing in the DLD and TD groups. OAEs do not give precise information of hearing thresholds. Thus, it is possible that subclinical differences in audiometric thresholds may have explained some of the observed variance within the two normal-hearing groups.

A limitation which applies to the CI and HA groups especially, was that the language ability and memory span tests were administered in the auditory modality. Although the tests were given in a quiet one-to-one setting, listening effort have likely been higher for the children with hearing loss (for a discussion of the interaction of perceptual and cognitive load, see e.g., Rönnberg et al., 2010). Listening effort may in turn have affected problem solving capacity and possibly led to fatigue in the children with hearing loss, and thus test results may not be entirely representative for their cognitive capacity.

In the current study, we used a HINT paradigm with notionally stationary[1] spectrally speech-shaped noise, presented from the front along with the target speech. In real world classrooms, noise will have an additional spatial and informational masking effect (for overview, see Brungart, 2001; MacCutcheon et al., 2019) as it will emanate from around the classroom and contain speech information. The SRT obtained by presenting target speech and noise from different directions would be more indicative of a real-life deficit than a score obtained for speech and noise coming from the same direction.

Furthermore, future studies could employ more ecologically valid tasks by simultaneously testing both SiN perception and another cognitive ability, e.g., by measuring differences in the outcomes of the cognitive tests when varying the difficulty of the speech perception task. Such simultaneous measures may contribute more knowledge about the interaction between deficits in SiN perception and cognitive abilities for children with hearing loss, as well as those with language disorders and typical development.

CONCLUSION

Results of the present study indicate that hearing-impaired children with HA and CI, but also some normal-hearing children with DLD, struggle with spoken language perception in noise compared to normal-hearing children with TD. The measure of SiN perception that was used in the present study, the HINT, was developed to have low linguistic demands and to be appropriate for children from 5 years upwards. Still, for the children with hearing loss and language disorder in the present study, language ability explained significant variance in results, even when taking into account speech perception abilities in quiet. Results on the HINT for children with hearing loss or language disorder should therefore be interpreted in light of their language profile.

Whilst technologies, such as directional microphones and FM systems, can improve the signal-to-noise ratio and thereby improve the recognition of speech in noise for young children with hearing loss, there may also be merit for parents, teachers and clinicians in focusing on language-specific early interventions to help improve children's capabilities to handle noisy classroom environments.

DATA AVAILABILITY STATEMENT

The datasets for this study are not publicly available at present, because a first publication from the larger project that this study is part of is currently in progress. Requests to access the datasets should be directed to OW: o.b.wie@isp.uio.no.

ETHICS STATEMENT

The study was reviewed and approved by the Regional Committees for Medical and Health Research Ethics, South-East Norway. Written informed consent to participate was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

JT and OW conceived and designed the study and collected the data together with MM, research assistants, and master's students. MM, AH, and JT performed and reported the analyses. MM made the figures. JT, AH, and MM wrote the manuscript with input from all authors. All authors contributed to interpreting all results and reviewing and editing the manuscript.

FUNDING

ACKNOWLEDGMENTS

---

[1] Background noise without superimposed amplitude modulation is often referred to as a stationary or steady noise. However, Stone et al. (2011, 2012) used the term notionally steady maskers since the maskers contain random amplitude fluctuations.

REFERENCES

Akeroyd, M. A. (2008). Are individual differences in speech reception related to individualdifferences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *Int. J. Audiol.* 47, S53–S71. doi: 10.1080/14992020802301142

Akhtar, N., Tolins, J., and Tree, J. E. F. (2019). "Young children's word learning through overhearing. Next steps," in *International Handbook of Language Acquisition*, eds J. S. Horst, and J. V. K. Torkildsen (London: Routledge), 427–441. doi: 10.4324/9781315110622-22

American Association on Intellectual and Developmental Disabilities, (2010). *Intellectual Disability: Definition, Classification, and Systems of Support**.

Available at: http://aaidd.org/intellectual-disability/definition#.WOy2GlLJIb1 doi: 10.4324/9781315110622-22 (accessed April 22, 2019).

American Psychiatric Association, (2013). *Diagnostic and Statistical Manual of Mental Disorders*, 5th Edn. Washington, DC: Author.

American Speech-Language-Hearing Association [ASHA], (1995). Position statement and guidelines for acoustics in educational settings. *ASHA* 37, 15–19.

Barnett, V., and Lewis, T. (1978). *Outliers in Statistical Data*. New York, NY: Wiley.

Bishop, D. V. M., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Westerveld, M. F., and Murphy, C. A. (2017). Phase 2 of CATALISE: a multinational and multidisciplinary Delphi consensus study of problems with language development: terminology. *J. Child Psychol. Psychiatry* 58, 1068–1080. doi: 10.1111/jcpp.12721

Bradlow, A. R., Kraus, N., and Hayes, E. (2003). Speaking clearly for children with learning disabilities: sentence perception in noise. *J. Speech Lang. Hear. Res.* 46, 80–97. doi: 10.1044/1092-4388(2003/007)

British Association of Teachers of the Deaf [BATOD], (2001). *Classroom Acoustics - Recommended Standards*. Available at: https://www.batod.org.uk/wp-content/uploads/2018/08/acoustic-standards.pdf (accessed April 22, 2019).

Bronkhorst, A. W. (2000). The cocktail party phenomenon: a review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica united with Acustica* 86, 117–128.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Caldwell, A., and Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *J. Speech Hear. Res.* 56, 13–30. doi: 10.1044/1092-4388(2012/11-0338)

Ching, T. Y., Zhang, V. W., Flynn, C., Burns, L., Button, L., Hou, S., et al. (2018). Factors influencing speech perception in noise for 5-year-old children using hearing aids or cochlear implants. *Int. J. Audiol.* 57, S70–S80. doi: 10.1080/14992027.2017.1346307

Ching, T. Y. C., Cupples, L., and Zhang, V. W. (2019). "Language abilities and language growth in children with hearing loss," in *International Handbook of Language Acquisition*, eds J. S. Horst, and J. V. K. Torkildsen (London: Routledge), 387–403. doi: 10.4324/9781315110622-20

Conti-Ramsden, G., Botting, N., and Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *J. Child Psychol. Psychiatry* 42, 741–748. doi: 10.1111/1469-7610.00770

Conway, C. M., Deocampo, J. A., Walk, A. M., Anaya, E. M., and Pisoni, D. B. (2014). Deaf children with cochlear implants do not appear to use sentence context to help recognize spoken words. *J. Speech Lang. Hear. Res.* 57, 2174–2190. doi: 10.1044/2014_jslhr-l-13-0236

Cook, R. D., and Weisberg, S. (1982). *Residuals and Influence in Regression*. New York, NY: Chapman & Hall.

Cord, M. T., Walden, B. E., and Atack, R. M. (1992). *Speech Recognition in Noise Test (SPRINT) for H-3 Profile*. Bethesda: Walter Reed Army Medical Center.

Crandell, C. C., and Smaldino, J. J. (1994). An update of classroom acoustics for children with hearing impairment. *Volta Rev.* 96, 291–306. doi: 10.4103/1463-1741.104894

Crandell, C. C., and Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Lang. Speech Hear. Ser. Sch.* 31, 362–370. doi: 10.1044/0161-1461.3104.362

Davidson, L. S., Geers, A. E., Hale, S., Sommers, M. M., Brenner, C., and Spehar, B. (2019). Effects of Early auditory deprivation on working memory and reasoning abilities in verbal and visuospatial domains for pediatric cochlear implant recipients. *Ear Hear.* 40, 517–528. doi: 10.1097/AUD.0000000000000629

Dunn, L. M., Dunn, L. M., Whetton, C., and Burley, J. (1997). *British Picture Vocabulary Scale 2nd Edition (BPVS-II)*. Windsor, CA: NFER-Nelson.

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *J. Acoust. Soc. Am.* 66, 651–653. doi: 10.1121/1.383691

Engdahl, B., Tambs, K., Borchgrevink, H. M., and Hoffman, H. J. (2005). Otoacoustic emissions in the general adult population of Nord-Trøndelag, Norway: III. relationships with pure-tone hearing thresholds. *Int. J. Audiol.* 44, 15–23. doi: 10.1080/14992020400022504

Fallon, M., Trehub, S. E., and Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *J. Acoust. Soc. Am.* 108, 3023–3029. doi: 10.1121/1.1323233

Ferguson, M. A., Hall, R. L., Riley, A., and Moore, D. R. (2011). Communication, listening, cognitive and speech perception skills in children with auditory processing disorder (APD) or specific language impairment (SLI). *J. Speech Lang. Hear. Res.* 54, 211–227. doi: 10.1044/1092-4388(2010/09-0167)

Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics*, 4th Edn. Thousand Oaks, CA: Sage, 458–459.

Füllgrabe, C., and Rosen, S. (2016). On the (un) importance of working memory in speech-in-noise processing for listeners with normal hearing thresholds. *Front. Psychol.* 7:1268. doi: 10.3389/fpsyg.2016.01268

Furnes, B., and Samuelsson, S. (2009). Preschool cognitive and language skills predicting Kindergarten and Grade 1 reading and spelling: a cross-linguistic comparison. *J. Res. Read.* 32, 275–292. doi: 10.1111/j.1467-9817.2009.01393.x

Gathercole, S. E., Willis, C. S., Baddeley, A. D., and Emslie, H. (1994). The children's test of nonword repetition: a test of phonological working memory. *Memory* 2, 103–127. doi: 10.1080/09658219408258940

Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., and Cowan, N. (2000). The development of auditory attention in children. *Front. Biosci.* 5:d108–d120.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate Data Analysis with Readings*. Tulsa, OK: Prentice Hall.

Hall, J. W. III, and Grose, J. H. (1990). The masking-level difference in children. *J. Am. Acad. Audiol.* 1, 81–88.

Harianawala, J., Galster, J., and Hornsby, B. (2019). Psychometric comparison of the hearing in noise test and the american english matrix test. *J. Am. Acad. Audiol.* 30, 315–326. doi: 10.3766/jaaa.17112

Jamieson, D. G., Kranjc, G., Yu, K., and Hodgetts, W. E. (2004). Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *J. Am. Acad. Audiol.* 15, 508–517. doi: 10.3766/jaaa.15.7.5

Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E., and Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Dev. Sci.* 18, 146–154. doi: 10.1111/desc.12202

Lucertini, M., Moleti, A., and Sisto, R. (2002). On the detection of early cochlear damage by otoacoustic emission analysis. *J. Acoust. Soc. Am.* 111, 972–978. doi: 10.1121/1.1432979

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., and Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Dev.* 75, 1357–1372. doi: 10.1111/j.1467-8624.2004.00745.x

Lyster, S. A. H., Horn, E., and Rygvold, A. L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British picture vocabulary scale, (BPVS II). *Spesialpedagogikk* 9, 35–43.

Lyxell, B., Sahlén, B., Wass, M., Ibertsson, T., Larsby, B., Hällgren, M., et al. (2008). Cognitive development in children with cochlear implants: relations to reading and communication. *Int. J. Audiol.* 47, S47–S52. doi: 10.1080/14992020802307370

MacCutcheon, D., Pausch, F., Fels, J., and Ljung, R. (2018). The effect of language, spatial factors, masker type and memory span on speech-in-noise thresholds in sequential bilingual children. *Scand. J. Psychol.* 59, 567–577. doi: 10.1111/sjop.12466

MacCutcheon, D., Pausch, F., Füllgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., et al. (2019). The contribution of individual differences in memory span and language ability to spatial release from masking in young children. *J. Speech Lang. Hear. Res.* 62, 3741–3751. doi: 10.1044/2019_JSLHR-S-19-0012

MacLeod, A., and Summerfield, Q. (1990). A procedure for measuring auditory and audiovisual speech-reception thresholds for sentences in noise: rationale, evaluation, and recommendations for use. *Br. J. Audiol.* 24, 29–43. doi: 10.3109/03005369009077840

McCreery, R. W., Walker, E. A., Spratford, M., Oleson, J., Bentler, R., Holte, L., et al. (2015). Speech recognition and parent-ratings from auditory development questionnaires in children who are hard of hearing. *Ear Hear.* 36, 60S–75S. doi: 10.1097/AUD.0000000000000213

Misurelli, S. M., and Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry. *J. Acoust. Soc. Am.* 132, 380–391. doi: 10.1121/1.4725760

Moore, D. R., Hogan, S. C., Kacelnik, O., Parsons, C. H., Rose, M. M., and King, A. J. (2001). Auditory learning as a cause and treatment of central dysfunction. *Audiol. Neurotol.* 6, 216–220. doi: 10.1159/000046836

Myhrum, M., Tvete, O. E., Heldahl, M. G., Moen, I., and Soli, S. D. (2016). The Norwegian hearing in noise test for children. *Ear Hear.* 37, 80–92. doi: 10.1097/AUD.0000000000000224

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1097/AUD.0b013e3181d3d514

Nilsson, M., Soli, S. D., and Gelnett, D. (1996). *Development of the Hearing in Noise Test for Children (HINT-C)*. Los Angeles, CA: House Ear Institute, 1–9.

Nilsson, M., Soli, S. D., and Sullivan, J. A. (1994). Development of the Hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. *J. Acoust. Soc. Am.* 95, 1085–1099. doi: 10.1121/1.408469

Nittrouer, S., Caldwell-Tarr, A., Tarr, E., Lowenstein, J. H., Rice, C., and Moberly, A. C. (2013). Improving speech-in-noise recognition for children with hearing loss: potential effects of language abilities, binaural summation, and head shadow. *Int. J. Audiol.* 52, 513–525. doi: 10.3109/14992027.2013.792957

Øygarden, J. (2009). *Norwegian Speech Audiometry*. PhD thesis. Norwegian University of Science and Technology, Norway.

Picard, M., and Bradley, J. S. (2001). Revisiting speech interference in classrooms. *Audiology* 40, 221–244. doi: 10.3109/00206090109073117

Pisoni, D. B., and Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear Hear.* 24(1 Suppl), 106S. doi: 10.1097/01.aud.0000051692.05140.8e

Plomp, R. (1978). Auditory handicap of hearing impairment and the limited benefit of hearing aids. *J. Acoust. Soc. Am.* 63, 533–549. doi: 10.1121/1.381753

Plomp, R. (1986). A signal-to-noise ratio model for the speech-reception threshold of the hearing impaired. *J. Speech Lang. Hear. Res.* 29, 146–154. doi: 10.1044/jshr.2902.146

Raven, J., and Raven, J. (eds) (2008). *Uses and Abuses of Intelligence: Studies Advancing Spearman and Raven's Quest for Non-Arbitrary Metrics*. Unionville, NY: Royal Fireworks Press.

Raven, J., Raven, J. C., and Court, J. H. (2004). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.

Rødvik, A. K., von Koss Torkildsen, J., Wie, O. B., Storaker, M. A., and Silvola, J. T. (2018). Consonant and vowel identification in cochlear implant users measured by nonsense words: a systematic review and meta-analysis. *J. Speech Lang. Hear. Res.* 61, 1023–1050. doi: 10.1044/2018_JSLHR-H-16-0463

Rönnberg, J., Rudner, M., Lunner, T., and Zekveld, A. A. (2010). When cognition kicks in: working memory and speech understanding in noise. *Noise Health* 12:263. doi: 10.4103/1463-1741.70505

Semel, E., Wiig, E. H., and Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals (CELF-4)*. San Antonio, TX: The Psychological Corporation.

Shield, B. M., and Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Soli, S. D., and Wong, L. L. N. (2008). Assessment of speech intelligibility in noise with the hearing in noise test. *Int. J. Audiol.* 47, 356–361. doi: 10.1080/14992020801895136

Sperling, A. J., Lu, Z. L., Manis, F. R., and Seidenberg, M. S. (2005). Deficits in perceptual noise exclusion in developmental dyslexia. *Nat. Neurosci.* 8:862. doi: 10.1038/nn1474

Stach, B. A. (2010). *Clinical Audiology: An introduction*. New York, NY: Cengage Learning.

Stevens, J. (2002). *Applied Multivariate Statistics for the Social Sciences*, 4th Edn. Hillsdale, NJ: Erlbaum.

Stone, M. A., Füllgrabe, C., Mackinnon, R. C., and Moore, B. C. J. (2011). The importance for speech intelligibility of random fluctuations in "steady" background noise. *J. Acoust. Soc. Am.* 130, 2874–2881. doi: 10.1121/1.3641371

Stone, M. A., Füllgrabe, C., and Moore, B. C. J. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317–326. doi: 10.1121/1.4725766

Stuart, A. (2005). Development of auditory temporal resolution in school-age children revealed by word recognition in continuous and interrupted noise. *Ear Hear.* 26, 78–88. doi: 10.1097/00003446-200502000-00007

Sullivan, J. R., Osman, H., and Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *J. Speech Lang. Hear. Res.* 58, 1043–1051. doi: 10.1044/2015_JSLHR-H-14-0204

Talarico, M., Abdilla, G., Aliferis, M., Balazic, I., Giaprakis, I., Stefanakis, T., et al. (2007). Effect of age and cognition on childhood speech in noise perception abilities. *Audiol. Neurotol.* 12, 13–19. doi: 10.1159/000096153

Tomblin, J. B., Harrison, M., Ambrose, S. E., Walker, E. A., Oleson, J. J., and Moeller, M. P. (2015). Language outcomes in young children with mild to severe hearing loss. *Ear Hear* 36, 76S–91S. doi: 10.1097/AUD.0000000000000219

Tukey, J. W. (1977). *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.

Vaillancourt, V., Laroche, C., Giguere, C., and Soli, S. D. (2008). Establishment of age-specific normative data for the Canadian French version of the hearing in noise test for children. *Ear Hear.* 29, 453–466. doi: 10.1097/01.aud.0000310792.55221.0c

Vandewalle, E., Boets, B., Ghesquiere, P., and Zink, I. (2012). Auditory processing and speech perception in children with specific language impairment: relations with oral language and literacy skills. *Res. Dev. Disabil.* 33, 635–644. doi: 10.1016/j.ridd.2011.11.005

Wilson, R. H. (2003). Development of a speech-in-multitalker-babble paradigm to assess word-recognition performance. *J. Am. Acad. Audiol.* 14, 453–470.

Wilson, R. H., Farmer, N. M., Gandhi, A., Shelburne, E., and Weaver, J. (2010). Normative data for the words-in-noise test for 6-to 12-year-old children. *J. Speech Lang. Hear Res.* 53, 1111–1121. doi: 10.1044/1092-4388(2010/09-0270)

Ziegler, J. C., Pech-Georgel, C., George, F., Alario, F. X., and Lorenzi, C. (2005). Deficits in speech perception predict language learning impairment. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14110–14115. doi: 10.1073/pnas.0504446102

Ziegler, J. C., Pech-Georgel, C., George, F., and Lorenzi, C. (2011). Noise on, voicing off: speech perception deficits in children with specific language impairment. *J. Exp. Child Psychol.* 110, 362–372. doi: 10.1016/j.jecp.2011.05.001

Influence of Classroom Acoustics on Noise Disturbance and Well-Being for First Graders

Arianna Astolfi[1]*, Giuseppina Emma Puglisi[1], Silvia Murgia[1], Greta Minelli[1], Franco Pellerey[2], Andrea Prato[3] and Tiziana Sacco[4]

[1] Department of Energy, Polytechnic University of Turin, Turin, Italy, [2] Department of Mathematical Sciences, Polytechnic University of Turin, Turin, Italy, [3] INRiM – National Institute of Metrological Research, Turin, Italy, [4] Department of Neurosciences, University of Turin, Turin, Italy

Several studies have shown so far that poor acoustics inside classrooms negatively affects the teaching and learning processes, especially at the lowest grades of education. However, the extent to which noise exposure or excessive reverberation affect well-being of children at school in their early childhood is still unanswered, as well as their awareness of noise disturbance. This work is a pilot study to investigate to which extent classroom acoustics affects the perceived well-being and noise disturbance in first graders. About 330 pupils aged from 6 to 7 years participated in the study. They belonged to 20 classes of 10 primary schools located in Torino (Italy), where room acoustic measurements were performed and where noise level was monitored during classes. The school buildings and the classrooms were balanced between socioeconomic status and acoustic conditions. Trained experimenters administered questionnaires in each class, where pupils answered all together during the last month of the school year (May). Questions included the happiness scale, subscales assessing self-esteem, emotional health, relationship at home and with friends, enjoyment of school, intensity and noise disturbance due to different sound sources, and quality of voice. The findings of the study suggest that long reverberation times, which are associated with poor classroom acoustics as they generate higher noise levels and degraded speech intelligibility, bring pupils to a reduced perception of having fun and being happy with themselves. Furthermore, bad classroom acoustics is also related to an increased perception of noise intensity and disturbance, particularly in the case of traffic noise and noise from adjacent school environments. Finally, happy pupils reported a higher perception of noise disturbance under bad classroom acoustic conditions, whereas unhappy pupils only reported complaints in bad classroom acoustics with respect to the perception of pleasances with himself or herself and of fitting in at school. Being a mother tongue speaker is a characteristic of children that brings more chances of attending classes in good acoustics, of being less disturbed, and of having more well-being, and richer districts presented better acoustic conditions, in turn resulting in richer districts also revealing a greater perception of well-being.

Keywords: well-being, noise disturbance, classroom acoustics, reverberation, first graders, happiness

INTRODUCTION

According to the (World Health Organisation [WHO], 2014), the physical environment in schools is one of the major elements of health promotion, and among the stressful environmental factors, high levels of noise can cause irritation, encourage aggressiveness, reduce physical and mental performance, and cause discomfort and headaches. Furthermore, children with learning difficulties, which are usually included in regular classes, are particularly dependent on a good acoustic (GA) environment (Winblad and Dudley, 1997).

Research has widely focused so far on the effects that classroom acoustics has on the teaching and learning processes, even at the lowest grades of education, but few studies have investigated the perception of noise disturbance at school and the influence of bad acoustics (BA), i.e., both excessive noise and reverberation, on the pupil's well-being. Particularly, no study has investigated, with in-classroom surveys, children well-being at school. Another important lack in the literature is the investigation of fundamental aspects of school life at the lowest grade of primary education, i.e., for most of the countries in Europe from 6 to 7 years. It is in the early childhood that the neuroplasticity of the human brain cortex is still high, and interventions can produce more positive effects. At the cortical level, various sensory and cognitive systems interact and adjust functional properties based upon experience and learning (Cardon et al., 2012).

In the achievement-oriented context of a classroom, well-being is a necessary precondition for learning (Hascher, 2008). School well-being is a multidimensional phenomenon in which occur school conditions, social relationships, means for self-fulfillment, and health status (Konu et al., 2002). Moreover, well-being of primary school pupils is positively influenced by learning skills (Epstein and McPartland, 1976; Tobia et al., 2019), which in turn are negatively influenced by BA (Puglisi et al., 2017).

Effect of Bad Classroom Acoustics on Learning Attainments of Children

On the one side, with very high reverberation time, primary school teachers raise their voice in order to be understood by pupils (Bottalico and Astolfi, 2012; Puglisi et al., 2017). This is mainly due to the effect of amplification of indoor noise due to excessive sound reflection. High noise levels can bring dysphonia or other vocal pathologies for teachers (Astolfi et al., 2012a; Bottalico et al., 2017a,b), which in turn can determine increased listening difficulties for children (Rudner et al., 2018). On the other side, thanks to voice monitoring performed with professional dosimeters (Carullo et al., 2013; Bottalico et al., 2018) and the uncertainty estimation of the vocal parameters (Castellana et al., 2017; Astolfi et al., 2018a), research has shown detrimental effects in speech production either in low reverberation (Astolfi et al., 2019) or in high reverberation (Astolfi et al., 2015) in the absence of noise, and optimal reverberation times for speaking have been proposed (Pelegrín-García et al., 2014; Calosso et al., 2017; Puglisi et al., 2017).

Unfavorable acoustics in classroom determines challenging environments for children, who are more sensitive than adults or older peers to noise and reverberation when performing tasks that involve listening comprehension and non-auditory features such as short-term memory, reading, and writing (Klatte et al., 2013). As a result, BA brings lower speech intelligibility scores, mostly for first graders (Astolfi et al., 2012b; Prodi et al., 2013; Puglisi et al., 2015b); degradation of the accuracy in identifying and producing newly learned words (Riley and McGregor, 2012); reduced reading speed of second graders (Puglisi et al., 2018); and lower scores in the standardized tests of literacy, mathematics, and science for pupils aged 7–11 years (Shield and Dockrell, 2008).

Children Perception of the Sound Environment at School

Together with the presence of noise sources as distractors for children's ability to understand, the subjective perception of the sound environment makes it very different the way listeners experience their everyday living spaces. Brännström et al. (2017) investigated on the 9- to 13-year-old children's personal ratings of perceived noise in order to improve the classrooms' design. They found that children were more annoyed to noise in tasks where the demands of verbal processing are higher. Dockrell and Shield (2004) administered a survey on the perception of noise in schools to children of grades 2 and 6. The more the external noise level increased, the more they were annoyed and the less they reported to be able to hear the teacher speaking inside the classroom. Astolfi and Pellerey (2008) assessed the subjective and objective environmental quality in classrooms involving 1,006 high school students with an average age of 16.1 years. They found that students reported to be strongly annoyed by noise sources internal in the classroom, i.e., other students talking in the classroom, and as a side effect of poor classroom acoustics on the overall perception of the school environment, students reported a decrease in their ability to concentrate.

Association Between Noise Annoyance and Well-Being

Given the above evidences, it is clear that the learning process is affected by the sound environment. Extensive literature on the subjective perception of the sound environment, especially with surveys in-field, is anyway lacking so far on the possible comorbidities that go beyond the students' performance at school. Particularly, research should thus focus on the assessment of the association between the perception of sound environment at school, in terms of noise annoyance, and of the wide concept of personal well-being. In fact, as the concept of well-being is composed of three main aspects, that is, subjective, psychological, and social well-being (Ryff and Singer, 1998; Diener, 2006; Keyes, 2013), it can be assumed to be strongly linked to the experience of everyday life situations and circumstances.

Some researchers found no association between noise from airplane and traffic and children well-being. In particular, chronic aircraft noise exposure resulted associated with high levels of noise annoyance but not to mental health problems in over 300 children aged 8–11 years (Haines et al., 2001). Mental health is

defined by World Health Organisation [WHO] (2014) as "a state of well-being in which every individual realizes his or her own potential, can cope with the normal stresses of life, can work productively and fruitfully, and is able to make a contribution to her or his community." Also, Stansfeld et al. (2009) reported no association between either aircraft or road traffic noise exposure and the Strengths and Difficulties Questionnaire (SDQ) total score, in more than 2,000 children aged 9–11 years. This questionnaire is a largely used psychometrically valid instrument to assess mental health of children aged 3–16 years, but the drawback of this tool is that for younger children, it has to be filled in by parents at home (Goodman, 2001).

On the opposite, Crombie et al. (2011) showed association between noise and mental health. Particularly, they found a relationship between road- and aircraft-generated noise and the incidence of mental problems in 9- to 10-year-old children. Lim et al. (2018) have found that high noise levels and high noise sensitivity determine less mental health of children and adolescents and that the effects of these noise-related variables depend on socioenvironmental factors. At school, Klatte and Hellbrück (2010) found that the children from classrooms with poor acoustics reported a higher burden of indoor noise in the classrooms and judged their relationships to their peers and teachers less positively than children from classrooms with GA. As a further study, Stansfeld et al. (2000) reviewed a set of studies on the relationship between noise and mental health. They found, as a general result, that noise was associated with stress-related factors of mental health (e.g., self-reported stress, sociability, and behaviors) and with well-being in children.

Other researchers investigated physiological symptoms of mental health, such as Evans et al. (2001) who found that children who lived in noisier areas had elevated resting systolic blood pressure and 8-h overnight urinary cortisol. Moreover, under laboratory conditions, they found that children from noisier neighborhoods showed elevated heart rate reactivity to reading tests. Similarly, Wålinder et al. (2007) studied the physiological and psychological stress reactions of children in relation to classroom noise. They reported that equivalent noise levels in classrooms, in the range between 59 and 87 dB(A), were significantly related to an increased prevalence of symptoms of fatigue and headache and to a reduced diurnal cortisol variability, indicating that noise should be focused on as a risk factor for children's well-being in the school environment.

As comorbidities, the exposure to road traffic noise in both residential and school areas was found to be associated in children aged between 7 and 11 years with emotional symptoms (Dreger et al., 2015) and hyperactivity and inattention (Forns et al., 2016; Hjortebjerg et al., 2016).

Need to Investigate the Effect of BA at School on Noise Disturbance and Well-Being for First Graders

Going beyond the available knowledge is therefore necessary, especially to understand if BA and other factors at school, where children spend most of their time, directly influence their harmonic growth, learning, and well-being. The research

should be focused at the lowest grades of education, when interventions can be more effective for children. In addition, subjective outcomes should be acquired by children themselves at school, while they are having classes, in order to catch their feeling while they are immersed in the classroom environment. Most of the previous studies on well-being for children aged less than 11 years are instead based on questionnaires administrated to parents or filled in by parents or children at home, and to the authors' knowledge, only few works are available so far in literature that carried out well-being and noise disturbance surveys with first graders at school. And few studies compared results from questionnaires with acoustic measures of noise, room acoustics, and speech intelligibility. This comparison is essential for planning future interventions and increase learning and well-being of children.

This work is thus a pilot study to investigate to which extent classroom acoustics affects the perceived well-being and noise disturbance in children. Noise disturbance has been used in this study, instead of noise annoyance, as noise annoyance is a multifaceted concept that could be more complicated for children as it includes, beside noise disturbance, interferences with some activities, nuisance, unpleasantness, and other factors (Guski et al., 1999; Di Blasio et al., 2019).

The purposes of the adopted approach, which was based on the combination of objective and subjective measurements, can be summarized in three main points: (i) assessment of classroom acoustics of first graders at school; (ii) assessment of the perceived well-being and noise disturbance at school; and (iii) association between classroom acoustics, perceived well-being, and noise disturbance. To reach the proposed objectives, room acoustic measurements were performed, and noise levels were monitored during classes. Then, based on the work by Sabri et al. (2015), questions on the perceived well-being inside the classroom environment were designed, containing information on emotional well-being (self-esteem, emotional health, and resilience), friends and family (quality of the relationship), satisfaction of school, and life satisfaction. To build a questionnaire oriented to investigating a multidimensional measure of school-related quality of life, based on Dockrell and Shield (2004) and Astolfi and Pellerey (2008), the abovementioned information was integrated with questions on the perceived noise disturbance.

MATERIALS AND METHODS

Participants

During a meeting for the presentation of the research project, teachers and parents have been informed about the project goal and the scientific evidence of the relationship between acoustics and subjective perceptions and well-being. Only pupils with parental consent were involved in this study, resulting in 367 students from 20 first-grade classes belonging to 10 primary schools in Turin. In the classes, a total number of pupils in the range between 8 and 25 was present. Subjects aged 6 represented 62% of the total, while 37% was aged 7 and only 1% was 8 years old. There were more males

(54%) than females (45%). The overall sample consisted of 77% Italian mother tongue (MT) pupils, while 23% used other primary languages in their family context (e.g., Romanian, Moroccan, English, German, Spanish, Albanian, and Arabic). Furthermore, some subjects declared to speak a second language besides Italian. **Table 1** shows the sociodemographic characteristics of the considered sample. Twenty-seven subjects were excluded from the data analysis because of cognitive or hearing deficits proved by the school administration through an official medical certificate or due to incorrect completion of the questionnaire.

Schools and Classrooms

Table 1 reports information about the school buildings and classes involved in the research project. Schools differed in location, period of construction, and architectural features. They are scattered in the city of Turin, in neighborhoods characterized by "low" or "medium" volume of traffic depending on the road typology as defined by the UK Department of Transport (2012), i.e., local road for local traffic or road intended to connect different city areas but not intended to provide large-scale transport links. In the case of medium volume of traffic, classrooms were not directly facing the road, but a courtyard or a corridor was present in between. In the case of low volume of traffic, most of the classrooms faced the road. The presence or absence of an acoustic treatment (AT) in the classrooms

("yes" or "no") was based on the presence or absence of whatever sound-absorbing material.

Acoustic Measurements
Acoustic Parameters and Adopted Protocol

Acoustic measurements were performed within 1 day approximately in the last 2 months of the school year. Measurements were carried out under occupied conditions, with the number of children inside being on average 18 across all classes. The adopted protocol regarded the acquisition of acoustic parameters that are useful to characterize the classroom's response to easy listening (Minelli et al., 2019). Below, the equipment, the parameters, and their measurement procedures are described.

Figure 1 shows the standard measurement setup used in each classroom. All the classrooms showed a traditional distribution of the pupils over the seating area, with the teacher's desk parallel to one of the shorter sides of the room, so that the source position S has been chosen and several microphone positions were considered. In particular, a fixed reference position that was common across all the classrooms was selected, being placed at 1 m from the source and at the same height of the source, i.e., 1.5 m from the ground, and that was named REF. With this microphone position being the same for every classroom, this measure well describe the difference across classrooms of changes in the vocal output, expressed as A-weighted equivalent sound

TABLE 1 | Description of the classrooms considered in terms of students and building features.

ID	Classroom volume (m³)	Classes		Schools			
		Male pupils (%)	Italian mother tongue (%)	Year of construction	Acoustic treatment	District real estate value (€)	Traffic volume
A1	194	60	67	1846	Yes	€€€€€	low
A2	261	50	61	1846	Yes	€€€€€	low
A3	283	58	79	1846	Yes	€€€€€	low
B1	203	50	91	1904	Yes	€€€€€	low
C1	123	58	33	1966	No	€	low
D1	255	67	72	1891	No	€€€€	low
D2	252	57	76	1891	No	€€€€	low
E1	236	68	74	1882	No	€€€	medium
E2	236	62	67	1882	No	€€€	medium
F1	279	61	67	1913	No	€€€€	medium
F2	261	56	44	1913	No	€€€€	medium
G1	136	75	75	1975	No	€€€	low
G2	106	58	95	1975	No	€€€	low
H1	133	57	76	1968	No	€€€€	low
H2	132	48	87	1968	No	€€€€	low
H3	140	50	90	1968	No	€€€€	low
H4	132	50	85	1968	No	€€€€	low
I1	237	13	100	1909	No	€€€	low
L1	241	45	86	1921	No	€€€	low
L2	264	52	96	1921	No	€€€	low

Schools were identified based on their neighborhood quality [i.e., "district real estate value," from Turin Real Estate Market Observatory (OICT, 2019) 2019, which was € for the property value interval of €1.000–1.500/m²; €€ for €1.500–2.000/m²; €€€ for €2.000–2.500/m²; €€€€ for €2.500–3.000/m²; and €€€€€ for €3.000–3.500/m²], the presence of acoustic treatment, the year of construction, and the traffic volume. IDs refer to schools (capital letters) and classrooms (number).

FIGURE 1 | Measurement setup in a typical classroom.
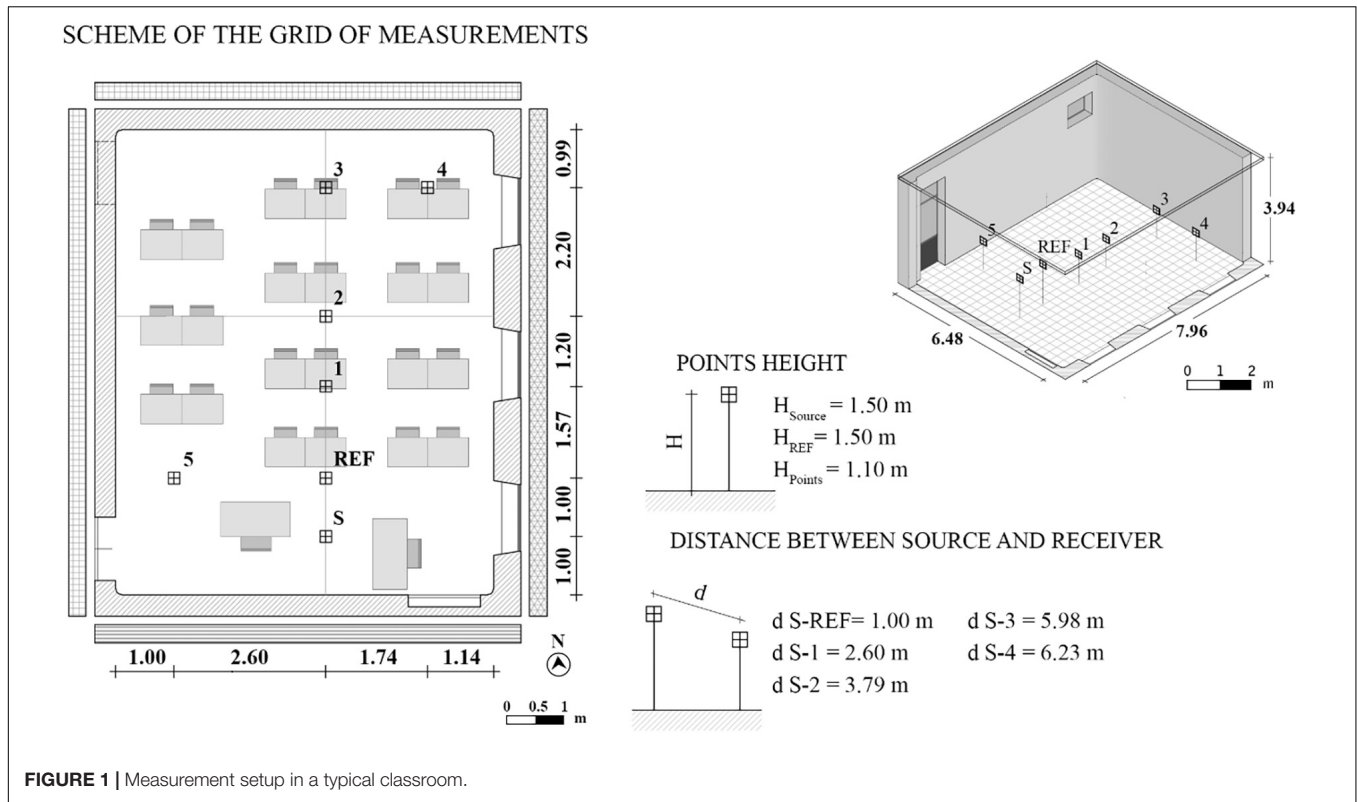
pressure level measured at 1 m from the speaker's mouth (ISO 9921, 2015), due to the reverberant sound field. Then, positions 1, 2, 3, 4, and 5 were selected to cover the whole seating area, with a fixed height of 1.1 m from the ground and with varying distances from the source depending on the geometrical characteristics of the classrooms. As a source, an acoustic stimuli generator, namely, a TalkBox (by NTi Audio, Schaan, Liechtenstein), which has the directivity pattern of the human voice, was positioned in the representative place that is typically covered by the teacher depending on the classroom's dimension and distribution, at the height of 1.5 m from the floor. With the aim of acquiring acoustic stimuli for the extraction of several parameters, a calibrated class 1 sound level meter (SLM, model XL2 by NTi Audio, Schaan, Liechtenstein) was located 1.5 m from the floor in the case of the REF position and 1.1 m from the floor for the other microphone positions, which were distributed over the pupils' seating area. Overall, both the instruments were positioned at least at a distance of 1 m from any surface.

Reverberation time (T20, s) and speech clarity (C50, dB) were measured according to ISO 3382-2 (2008) and ISO 3382-1 (2009), respectively. Room impulse responses were acquired from three repeated exponential sine sweep signals, which were emitted by the TalkBox and recorded by the SLM at each measurement point. The sweep signals were generated with a sample frequency of 44.1 kHz and a resolution of 16 bits. They were designed to cover a range of 0.05–2 kHz and to have a duration of 3 s each, based on the assumption that in cases with moderate background noise, as in the classrooms under study, it is normally a safe practice to use sweeps with a length of two to four times the expected longest

reverberation time (ISO 18233, 2006). Particularly, the SLM was moved along the main axes represented in **Figure 1** in the receiver positions REF, 1, 2, 3, and 6. Frequency averages were calculated according to Din 18041 (2016) in the range 0.25–2 kHz for T20 and according to ISO 3382-1 (2009) in the range 0.5–1 kHz for C50. A range of optimal occupied T20 was set between 0.5 and 0.8 s, according to a number of recent studies (0.7 s in Yang and Bradley, 2009; 0.8 s in Bottalico and Astolfi, 2012; 0.5–0.6 s in Pelegrín-García et al., 2014; 0.7 s in Puglisi et al., 2017; and 0.8 s in Calosso et al., 2017), which concerned both speech and listening performance in typical primary and secondary school classrooms. Therefore, for the subsequent analyses, classes were split into GA and BA whether they were in or out of such optimal range, respectively. As far as C50 is concerned, it represents an index related to speech intelligibility in the classroom in the presence of low noise. As given in Bradley (1986), it is assumed that an optimal value of C50 for small classrooms with an optimal reverberation time of 0.8 s at middle frequencies should be greater than around 3 dB.

T20 in unoccupied conditions, T20_e, where the subscript "e" is for empty, was measured with a wooden clapper, i.e., two wooden boards hinged together, according to the ISO 3382-2 (2008), as described in Puglisi et al. (2017). The measurement method in the empty condition was different compared to the one in the occupied condition; however, the two procedures can be considered equivalent in the frequency range of interest as described in Puglisi et al. (2018). Optimal values of T20_e, which are proportional to the classroom volume, were derived from Italian

technical standard UNI 11367 (2010) as an average value between 0.5 and 1 kHz.

Background noise level (L_N, dBA) was considered in terms of indoor A-weighted equivalent sound pressure level (L_{Aeq}). Repeated measurements were performed based on 3-min acquisitions (Puglisi et al., 2015a), with windows and doors closed to ensure that no external unforeseen noises influenced them, except the typical ones related to traffic and roads. For this measure, the SLM was located in two or three positions in each classroom, which alternatively corresponded to positions 2, 4, and 5, as presented in **Figure 1**. Noise measurements were carried out with children in silence, L_{N_sil}, and with the children performing group activities, L_{N_gr}, in order to be representative of typical classroom scenarios. Both the mentioned conditions were guaranteed with the help of the teacher who asked the children to keep silent and then to speak as they were in a traditional group lesson. According to Shield and Dockrell (2008) and BB93 (WSP, 2015), the L_{N_sil} recommended value must be less than or equal to 35 dBA.

The speech signal (L_S, dBA) was measured to characterize the propagation of a voice signal in each classroom. The TalkBox was positioned in S according to **Figure 1**, and then it was set to emit a voice signal with a "normal" vocal effort, i.e., corresponding to 60 dBA at 1 m in anechoic conditions to comply with ANSI S3.5 (1997). The speech signals were acquired as an equivalent A-weighted continuous sound pressure level at the SLM, having the source switched on for the receiving positions REF, 1, 2, 3, and 4.

The ratio of useful to detrimental energy (U50, dB) was calculated to consider both the effect of the acoustics of the environment and the effect of the signal-to-noise ratio on speech intelligibility (Bradley et al., 1999). It is obtained in terms of ratio between useful and detrimental energy, i.e., noise and reverberation. Particularly, it is obtained for each position, from C50, L_S, and L_{N_sil} values. U50 was then averaged between 500 and 1,000 Hz ($U_{50.0.5-1\ kHz}$), as it is derived from a C50 value that exhibits the same span of frequency averaging. As given in Bradley et al. (1999), it is assumed that an optimal value of U50 for small classrooms should be greater than 1 dB.

Overall, the acoustic parameters that are distance dependent (i.e., C50, L_S, and U50) were measured point by point as described above, but then the acquired measures were processed to have single values that are useful for an effective comparison across classes. In particular, for C50 and U50, the values measured across the classroom were averaged together to have a spatial mean, excluding the reference point (REF in **Figure 1**), being hereafter reported to as "M" in the parameters' symbols. Such a value was found to be not so different from the central value, i.e., the value measured at position 2 in **Figure 1**, which is hereafter reported as "ctr" as subscript for C50 and U50, as underlined in Puglisi et al. (2018). Furthermore, L_S values, which were measured on the axis in front of S (i.e., acquisitions in points REF, 1, 2, and 3) were associated to obtain its slope per double distance (in decibels per double distance) (Astolfi et al., 2008), which is hereafter referred to as "m" in the parameters symbols.

Measurement Results

Table 2 shows the acoustic parameters grouped for GA and BA, i.e., with occupied reverberation time in or out of the optimal range of 0.5–0.8 s, respectively. A good agreement is shown between occupied, T20, and unoccupied reverberation times, T20_e, even though in the case of the unoccupied condition, optimal values are too low compared to the measured values. For the classrooms A2, A3, C1, E1, and I1, T20_e is surprisingly lower or equal to T20. This strange behavior is due to the different frequency range assumed for the averaging, which is only confined to 0.5–1 kHz in the case of the empty room.

Optimal values are shown for speech intelligibility expressed by the parameters C50 and U50, in the classroom with GA, as well as lower values of noise level during group activity, L_{N_gr}, that is strictly related to reverberation time. In contrast, noise level in silence, L_{N_sil}, does not show decreased values in GA.

No such clear tendency is shown for the parameters related to the speech level, L_{S_REF} and mL_S, with reverberation time, even though the expected values should be lower in GA and higher in BA. In the case of the slope of the speech signal, mL_S, higher values (or lower if the slope is taken in absolute value) mean values closer to zero, i.e., absence of slope due to a more uniform acoustic field.

Questionnaires

Two questionnaires have been designed in order to evaluate children subjective perception of well-being and noise disturbance while in classroom. In both cases, questions have been adapted for children of 6 years old, taking into account readability, comprehension and ease of administration of the document. An Italian translation of the questionnaire developed by Sabri et al. (2015), suitable for young people with special educational needs (SENs), was done and then used as the questionnaire for perceived well-being assessment, while the questionnaire on noise disturbance was adapted on the base of the work by Dockrell and Shield (2004) and by Astolfi and Pellerey (2008). The translation of instructions and items was made using a back-to-back translation, that is, this procedure was used to validate that the translation content did not deviate significantly from the original. Initially, the English version was translated into Italian, then it was translated back to English and this back translation checked with the original for inconsistencies. The back translation was judged to be consistent with the original English version.

The final version of the whole questionnaire is shown in **Table 3**. It was administrated inside the classrooms where the didactic activity takes place every day, while researchers of the Polytechnic of Turin and teachers assisted the compilation: in particular each question was read by a child in turn who was then asked to explain its meaning. If it was unclear or not understood by all, an intervention by the teacher was required in order to clarify the question. Questionnaires were filled in during sessions of about 40 min each, with a break in between, along 1 day either

TABLE 2 | Acoustic parameters for classrooms grouped for BA and GA. L_{N_sil} is the noise acquired with students in silence, L_{N_gr} is the noise acquired with students doing group activities; L_{S_REF} is the signal measured at the reference point.

ID	Acoustic parameters									
	T20_e (s)	T20 (s)	L_{N_sil} (dBA)	L_{N_gr} (dBA)	L_{S_REF} (dBA)	mL_S (dBA/dd)	C50_ctr (dB)	C50_M (dB)	U50_ctr (dB)	U50_M (dB)
Bad acoustics										
A1	1.0 (0.04)	0.9 (0.04)	51.7	NA	61.3	−1.9	1.0	1.3 (1.2)	−1.1	−1.2 (1.0)
A2	**0.8** (0.06)	0.9 (0.18)	49.0	64.7	61.2	−2.4	0.0	2.2 (1.8)	−1.3	0.2 (1.1)
D1	1.4 (0.06)	1.2 (0.31)	51.2	68.0	63.0	−1.8	−0.6	0.0 (0.9)	−2.1	−1.5 (1.0)
D2	1.4 (0.01)	1.3 (0.25)	52.0	NA	62.7	−2.1	0.0	−0.3 (1.1)	−1.6	−1.8 (1.4)
E1	1.2 (0.06)	1.2 (0.11)	54.0	66.6	62.1	−1.4	0.7	1.1 (0.9)	−1.5	−1.2 (0.6)
E2	1.0 (0.01)	0.9 (0.08)	54.3	73.7	61.5	−1.9	3.8	2.7 (1.0)	0.0	−0.9 (0.8)
F1	1.5 (0.01)	1.2 (0.12)	52.0	75.1	62.1	−1.7	1.1	−0.3 (1.8)	−0.9	−2.2 (1.6)
F2	1.7 (0.03)	1.4 (0.33)	52.0	73.8	62.9	−1.8	−1.1	−0.1 (1.2)	−2.7	−1.8 (1.3)
G1	1.2 (0.07)	0.9 (0.05)	51.5	72.2	62.3	−2.1	3.3	2.6 (1.0)	1.3	0.9 (0.9)
I1	1.3 (0.07)	1.4 (0.07)	45.7	59.9	61.9	−1.6	−2.2	−2.2 (0.2)	−2.6	−2.6 (0.2)
L1	NA	1.0 (0.05)	47.9	67.4	61.1	−1.9	1.2	1.6 (1.0)	0.0	0.4 (1.1)
L2	1.2 (0.02)	1.1 (0.07)	46.0	71.6	62.0	−2.2	0.6	0.5 (2.0)	−0.1	−0.2 (2.1)
Good acoustics										
A3	**0.8** (0.05)	**0.8** (0.05)	38.4	61.8	60.3	−2.0	**5.1**	**4.1** (0.9)	**4.8**	**3.8** (0.9)
B1	**0.6** (0.07)	**0.5** (0.05)	49.3	66.3	60.8	−2.1	**7.3**	**7.6** (1.5)	**2.9**	**4.0** (1.9)
C1	0.8 (0.03)	**0.8** (0.12)	49.3	62.2	62.8	−1.6	2.8	**3.3** (0.8)	1.9	**2.2** (0.6)
H1	0.8 (0.03)	**0.7** (0.20)	51.6	71.9	61.5	−1.1	**3.8**	**3.6** (0.2)	1.4	**1.5** (0.2)
H2	0.9 (0.04)	**0.6** (0.08)	55.9	68.1	62.4	−2.2	**5.4**	**5.3** (0.2)	−1.0	−0.8 (0.5)
H3	1.0 (0.03)	**0.7** (0.04)	45.5	63.9	62.9	−1.8	**3.8**	**3.8** (0.3)	**3.0**	**3.2** (0.3)
H4	0.8 (0.04)	**0.7** (0.07)	53.1	65.5	62.9	−2.1	**3.5**	**4.1** (0.6)	−0.1	0.6 (0.6)
G2	0.9 (0.01)	**0.6** (0.01)	51.9	65.3	60.7	−0.8	**3.5**	2.9 (0.9)	**1.4**	0.8 (0.7)

The slope of a parameter per double distance is referred to with "m" (e.g., mL_S); the mean distribution of a parameter in the classroom is referred to with "M" (e.g., U50_M); the value of a parameter measured in the center of the classroom is referred to with "ctr" (e.g., U50_ctr). Standard deviations are indicated in brackets when available. Not available measures are marked as NA. Values in bold are compliant with the optimal values in occupied conditions.

in May 2017 or in May 2018. Approval to conduct the present study was granted by the Politecnico di Torino Ethics Committee.

The well-being questionnaire started with an introductory section that consists in five items on sociodemographic information such as age, gender, number of people living at home, the quietest place known by the individual, primary language spoken in family. Finally, the last item of the section and the questionnaire is an open question where the child is asked to report an opinion about the feelings of their school sound environment. Then the questionnaire presents five sections: (1) self-esteem; (2) emotional health; (3) relationship at home and with friends; (4) enjoyment of school; (5) scale of happiness. Sections 1–4 consist in three questions each, where a three-point ordinal scale allows to choose the accordance among the options (a) yes, (b) not sure or (c) no. For section 5 the evaluation scale consists in a 11-point scale, where pupils had to put a cross on the number of an illustrated stair corresponding to their perceived level of happiness. A visual feedback with sketches and emoticons helped in the compilation of the questionnaire.

The noise questionnaire contained three sections: (1) perceived disturbance of specific noise sources (i.e., traffic, car sirens, internal noise, and natural noise), (2) perceived intensity and disturbance of noise during school activities performed either in silence or in group, (3) perceived voice quality under two situations, that is, while a classmate asks a question or while the teacher explains. The described sections were associated with three-point ordinal scales of evaluation, in which the judging items typically varied from the less (left) to the most (right) disturbing/annoying response. As in the well-being questionnaire, figures and emoticons were used to make it easier to identify the type of noise being investigated and a symbology that facilitated the indication of the perceived disturbance. Finally, the last item of the questionnaire consisted in an open question where the child was asked to optionally add comments.

Statistical Analysis

The statistical analysis was carried out with SPSS (IBM Statistics 20, IBM, Armonk, NY, United States). In order to detect and eliminate outliers from the original sample, two different methodologies have been applied, one that refers to the well-being answers and the other based on the noise answers. As far as the well-being answers, pupils have been divided in two groups based on the Q13_WB answer, i.e., the happiness scale, through a 2-means cluster analysis. Particularly, unhappy children have been considered for answers from 0 to 6, and happy children for answers from 7 to 10. Then, a logistic regression has been carried out considering the membership in the group as the response variable and the others well-being answers as explicative variables, and the Cook's distance, for every pupil has been obtained. The Cook's distance measures inconsistence between the response variables and the explicative ones. All the pupils with a

TABLE 3 | Administered questionnaires on the perceived well-being and noise disturbance, with scales and labels question by question.

Topic	ID	Questions	Scale	Labels
Well-being				
Self-esteem	Q1_WB	I am proud of myself	1–3	Yes (1) Not sure (2) No (3)
	Q2_WB	I'm serene	1–3	
	Q3_WB	I have a lot of fun	1–3	
Emotional health	Q4_WB	Lots of things about me are good	1–3	
	Q5_WB	I feel pleased with myself	1–3	
	Q6_WB	I am a cheerful child	1–3	
Relationship at home and with friends	Q7_WB	My home is a good place to relax	1–3	
	Q8_WB	I enjoy myself with my friends	1–3	
	Q9_WB	My friends help me if I need it	1–3	
Enjoyment of school	Q10_WB	I like being in school	1–3	
	Q11_WB	I feel safe at school	1–3	
	Q12_WB	I feel like I fit in at school	1–3	
Scale of happiness	Q13_WB	Put a cross on the number corresponding to your degree of happiness	0–10	Not happy (0) Happy (10)
Noise				
Traffic disturbance	Q1_N	How much traffic noise disturb you?	1–3	A little (1) Quite (2) A lot (3)
Siren disturbance	Q2_N	How much ambulance, firemen and police sirens disturb you?	1–3	
Disturbance from internal noises	Q3_N	How much sounds of radios or recorders coming from other classrooms or from the corridor disturb you?	1–3	
Rain disturbance	Q4_N	How much rain noise, if it's raining outside, disturb you?	1–3	
Intensity and disturbance during silent task	Q5_N	The noise present when you perform a task in silence seems you.	1–3	Low (1) Medium (2) High (3)
	Q6_N	How much does this noise disturb you?	1–3	A little (1) Quite (2) A lot (3)
Intensity and disturbance during group activity	Q7_N	The noise present when you're working in a group seems to you.	1–3	Low (1) Medium (2) High (3)
	Q8_N	How much does this noise disturb you?	1–3	A little (1) Quite (2) A lot (3)
Quality of voice	Q9_N	How do you hear the teacher' voice if you are in silence while she is talking?	1–3	Good (1) Quite good (2) Bad (3)
	Q10_N	How do you hear your classmate's voice if he or she is answering to the teacher?	1–3	

Cook's distance higher than 0.15 have been recognized as outliers. As far as the noise answers is concerned, outliers have been automatically recognized for each class, by the corresponding box-plots related to the means of the answers from Q1_N to Q10_N.

In the current sample, the Cronbach's α values of 0.69 and 0.71 have been obtained from the answers to the two questionnaires on well-being and noise disturbances, respectively, thus showing the internal consistency of both the questionnaires, i.e., a good intercorrelations among test items in each set of questions. But, being the two values not sufficiently close to unity, they also reveal that the two sets of items actually measure several unrelated latent constructs. This fact confirms that both questionnaires are suitable to capture different aspects of well-being and of noise disturbances, in the form they were designed.

Non-parametric methods have been used to analyze data obtained with ordinal scales, as in the case of this study (Sigel and Castellan, 1988). The significance of the differences between happy and unhappy children in good and bad classroom acoustics, related to several factors concerning well-being and noise, as well as the differences between males and females, was assessed with the Mann–Whitney U Test (MWU), a test that is used for two groups of independent observations.

The relationships between the subjective outcomes and the issues concerning the school context, the MT percentage and the classrooms characteristic acoustic parameters, as well as between the well-being and noise disturbance scores and the acoustic parameters, were also investigated through the non-parametric and non-linear correlation estimator Spearman's rho (Croux and Dehon, 2010).

In order to perform a more robust correlation analysis, the acoustic parameters have been previously analyzed with respect to their expected relationships with the parameter Reverberation time, T20. All the acoustic parameters for each classroom, except L_{Aeq_sil}, have been related to T20 through a linear regression, and the classrooms that did show evident anomalous tendencies have been considered outliers and thus canceled from the database. From zero to three classroom values for each parameter have been canceled from the original database. The anomalies were due to non-uniform reverberant sound field in the classroom caused by concentration due to vaulted ceilings, shape of the room, causalities in the measurements and so on.

RESULTS

After the removal of outliers, a final sample of 326 questionnaires was used for the subjective analyses. The students were almost equally subdivided between males (54%) and females (46%), and 79.0% were Italian. A reduced sample of 296 students, corresponding to the happy children, were also used for the correlation between the subjective and objective acoustic data and between the subjective data and the classes and schools' characteristics. Unhappy children have been removed from this analysis in order to have a more homogeneous sample.

A preliminary analysis has been made in order to get differences between males and females on the single questions, considering the whole sample of children after the removal of the outliers. Apart from the questions Q5_WB (pleasances with himself or herself) and Q6_WB (cheerfulness), for which a statistical significant difference has been found between males and females according to the MWU test, and lower average scores have been gathered by males for both the questions, no difference has been found for the other well-being and noise disturbance aspects. Females have been found less pleased with themselves and less cheerful than males.

The main statistical analyses have then been carried out and the results commented in the paragraphs below. In particular, the research questions were explained through the following outcomes:

- Comparison between subjective responses on well-being and noise disturbance in good and bad classroom acoustics for all the students (**Table 4**), for happy students (**Table 5**) and for unhappy students (**Table 6**);
- Comparison between subjective responses on well-being and noise disturbance for happy and unhappy students in classrooms with GA (**Table 7**) and BA (**Table 8**);
- Correlation matrix of the perceived well-being and noise disturbance scores of happy children with the classes and schools characteristics (**Table 9**);
- Correlation matrix of the acoustic parameters (**Table 10**);
- Correlation matrix of the perceived well-being and noise disturbance scores of happy children with the acoustic parameters (**Table 11**).

Standard deviations in **Tables 4–7** are quite high considering the 1–3 points scale of the adopted questionnaires. However, they are proportionally comparable to those reported in literature for children aged from 6–7 to 10–11 years in Dockrell and Shield (2004), and in the range 9–13 years in Brännström et al. (2017), which adopted 1–5 points scales.

As far as the correlation analysis is concerned, some of the correlation coefficients shown in **Tables 10, 11** are low in absolute value, leading anyway to weak but statistically significant relationships according to the size of the samples.

Influence of Bad Classroom Acoustics on Well-Being and Noise Disturbance for Happy and Unhappy Children

Table 4 shows higher mean values, which corresponds to worse conditions, in BA compared to GA, with significant differences according to the MWU Test, for the answers to the question Q12_WB related to the enjoyment of school, and particularly to the fitting in at school. The same occurs for the answers to the questions Q7_N, Q8_N, and Q9_N, which relate to noise disturbance during silence tasks, and to intensity and disturbance during group activity, respectively. The results highlight as noise in classrooms with higher reverberation, either from outdoor or indoor, can be perceived as more disturbing being more amplified by reverberation.

The same results that concern noise disturbance have also been found for happy students, as shown in **Table 5**, while **Table 6** shows as for unhappy students the main problem in BA only concern serenity (Q2_WB), pleasances with himself or herself (Q5_WB) and fitting in at school (Q12_WB).

Tables 7, 8 show the comparison between happy and unhappy students' subjective responses on the perceived well-being and noise disturbance, in good and bad classroom acoustics, respectively. Seven out of twelve well-being items have been judged significantly worse by unhappy students in GA, compared to eleven out of twelve in BA, thus suggesting as bad classroom acoustics determines less well-being that GA.

Influence of Classes and School Characteristics on Well-Being and Noise Disturbance

Table 9 shows the most significant correlations between the perceived well-being and noise disturbance answers related to the happy children and the classes and schools characteristics. To this analysis' aim, correlations were run between individual subjective responses (i.e., on perceived well-being and noise disturbance) and overall classes and schools characteristics, as specific objective data did not differ from pupil to pupil. Only Spearman correlation coefficients with p-values lower than 0.01 are shown. All the relationships are coherent among them, but the most interesting ones deserve some comments that are reported below:

- The percentage of Italian MT children is negative related to the traffic volume (TV), to the reverberation time in the empty classrooms, T20_e, and to the background noise level with the children in silence, L_{N_sil}, thus showing as schools with a higher percentage of MT children have better acoustics. This is also confirmed by the lower perception of intensity and disturbance of noise in quiet (Q5_N and Q6_N, respectively), when the MT are higher. Moreover, MT is negatively related to the feeling of being pleased with themselves (Q5_WB), thus suggesting as a higher MT is related to higher well-being.
- The district real estate value (DV) is negatively related to the TV and to the acoustic parameters T20_e and L_{N_sil}, and it is also positively related to the presence of AT. These

TABLE 4 | All students' subjective responses on the perceived well-being and noise disturbance, with the sample being grouped per classroom acoustics (i.e., good and bad).

Topic	ID	Scale	Good acoustics			Bad acoustics			MWU p-value
			(N = 139)			(N = 187)			
			M	SD	95% CI	M	SD	95% CI	
Well-being									
Self-esteem	Q1_WB	1–3	1.24	0.50	1.15–1.32	1.25	0.55	1.17–1.33	0.86
	Q2_WB	1–3	1.43	0.71	1.32–1.55	1.56	0.76	1.45–1.67	0.11
	Q3_WB	1–3	1.19	0.52	1.11–1.28	1.23	0.54	1.16–1.31	0.35
Emotional health	Q4_WB	1–3	1.30	0.55	1.21–1.39	1.34	0.60	1.26–1.43	0.64
	Q5_WB	1–3	1.21	0.47	1.13–1.29	1.32	0.61	1.23–1.41	0.12
	Q6_WB	1–3	1.20	0.50	1.12–1.28	1.24	0.55	1.16–1.32	0.56
Relationship at home and with friends	Q7_WB	1–3	1.28	0.59	1.18–1.38	1.28	0.62	1.19–1.37	0.72
	Q8_WB	1–3	1.12	0.34	1.06–1.17	1.16	0.46	1.09–1.22	0.70
	Q9_WB	1–3	1.32	0.63	1.22–1.43	1.30	0.56	1.23–1.38	0.86
Enjoyment of school	Q10_WB	1–3	1.25	0.56	1.15–1.34	1.41	0.75	1.30–1.51	0.09
	Q11_WB	1–3	1.41	0.65	1.30–1.52	1.40	0.63	1.31–1.49	0.91
	Q12_WB	1–3	1.26	0.56	1.16–1.35	1.42	0.70	1.32–1.53	**0.03**
Scale of happiness	Q13_WB	0–10	9.24	1.83	8.94–9.55	9.28	1.99	9.00–9.57	0.51
Noise									
Traffic disturbance	Q1_N	0–3	1.88	0.87	1.65–2.10	1.89	0.84	1.67–2.11	0.92
Siren disturbance	Q2_N	0–3	1.90	0.83	1.72–2.08	1.88	0.78	1.72–2.05	0.93
Disturbance from internal noises	Q3_N	0–3	1.83	0.81	1.66–2.00	1.75	0.75	1.60–1.90	0.56
Rain disturbance	Q4_N	0–3	1.72	0.74	1.57–1.88	1.83	0.91	1.67–1.99	0.58
Intensity and disturbance during silent task	Q5_N	1–3	1.42	0.67	1.31–1.53	1.53	0.74	1.42–1.64	0.18
	Q6_N	1–3	1.28	0.57	1.19–1.38	1.52	0.76	1.41–1.63	**0.01**
Intensity and disturbance during group activity	Q7_N	1–3	1.93	0.80	1.79–2.06	2.14	0.77	2.03–2.25	**0.02**
	Q8_N	1–3	1.74	0.78	1.61–1.88	1.97	0.83	1.85–2.09	**0.02**
Quality of voice	Q9_N	1–3	1.29	0.61	1.19–1.40	1.25	0.55	1.17–1.33	0.65
	Q10_N	1–3	1.58	0.71	1.46–1.70	1.53	0.74	1.43–1.64	0.37

Mean (M), standard deviation (SD) and the 95% confidence interval (95% CI) values are provided, and the p-values of significance for the differences between the two acoustic conditions, according to the MWU Test. Any statistically significant differences, with p-value < 0.05, are reported in bold.

results prove as richer districts allows for schools with better acoustics. DV is also negatively related to the goodness of the children' s relationships at home and with friends (Q7_WB and Q9_WB), generally indicating that growing in poorer districts can bring to less well-being.

- High TV determines lower serenity and increases noise disturbance, as shown from the positive relationships of TV with the feeling of being serene (Q2_WB) and with the perception of disturbance of noise from outdoor and indoor (Q1_N, Q2_N, Q3_N, and Q6_N).
- Acoustic treatments is negatively related to the reverberation time in the empty classrooms, T20_e, and to the background noise level with the children in silence, L_{N_sil}, as expected.

Relationships Between Acoustic Parameters and Subjective Outcomes
Correlations Between Acoustic Parameters
Table 10 shows the correlations between the acoustic parameters measured in the classrooms. Only Spearman correlation coefficients with p-value less than 0.01 are shown. Very high

relationships are shown between the acoustic quantities, as expected. Particularly most of the quantities are very well related to the reverberation time T20, both empty and occupied.

The parameters L_{N_sil}, L_{N_gr}, L_{S_REF}, and mL_S are positive related to the reverberation time, that is when T20 is higher they are higher, as expected, and negatively related to C50 and U50. A very tight connection is shown between central and mean values of the quantities C50 and U50, thus bringing to the practical conclusion that only one measurement in the center of the room can well describe the behavior of the whole classroom in terms of speech intelligibility, as already shown by Puglisi et al. (2017). U50_ctr is also well related to C50_ctr, which suggests the use of only one quantity instead of two to represent speech intelligibility.

Correlations Between Acoustic Parameters and Well-Being and Noise Disturbance Scores
In order to choose among the most important objective quantities to be measured in a classroom it is useful to look at the results in **Table 11**, which shows the correlation matrix of the perceived well-being and noise disturbance

TABLE 5 | Happy students' subjective responses on the perceived well-being and noise, with the sample being grouped per classroom acoustics (i.e., good and bad).

Topic	ID	Scale	Good acoustics (N = 124)			Bad acoustics (N = 172)			MWU p-value
			M	SD	95% CI	M	SD	95% CI	
Well-being									
Self-esteem	Q1_WB	1–3	1.22	0.49	1.13–1.30	1.22	0.53	1.16–1.91	0.80
	Q2_WB	1–3	1.43	0.73	1.30–1.56	1.50	0.73	1.81–2.59	0.31
	Q3_WB	1–3	1.14	0.45	1.06–1.22	1.19	0.48	1.24–2.19	0.15
Emotional health	Q4_WB	1–3	1.26	0.52	1.17–1.35	1.33	0.59	1.21–1.86	0.35
	Q5_WB	1–3	1.22	0.49	1.13–1.30	1.28	0.57	1.29–2.18	0.35
	Q6_WB	1–3	1.10	0.31	1.05–1.16	1.15	0.41	1.82–2.71	0.43
Relationship at home and with friends	Q7_WB	1–3	1.23	0.54	1.14–1.33	1.24	0.57	1.25–2.22	0.89
	Q8_WB	1–3	1.09	0.31	1.03–1.14	1.12	0.39	1.18–2.02	0.68
	Q9_WB	1–3	1.27	0.57	1.17–1.37	1.24	0.47	1.53–2.47	0.77
Enjoyment of school	Q10_WB	1–3	1.22	0.53	1.12–1.31	1.35	0.70	1.55–2.59	0.14
	Q11_WB	1–3	1.31	0.59	1.20–1.41	1.30	0.54	2.21–2.73	0.81
	Q12_WB	1–3	1.23	0.52	1.13–1.32	1.33	0.61	2.20–3.00	0.11
Noise									
Traffic disturbance	Q1_N	0–3	1.87	0.87	0.62–1.01	1.94	0.83	0.00–0.44	0.65
Siren disturbance	Q2_N	0–3	1.86	0.84	0.86–1.25	1.90	0.77	0.00–0.94	0.68
Disturbance from internal noises	Q3_N	0–3	1.79	0.83	0.93–1.31	1.74	0.76	0.46–1.54	0.78
Rain disturbance	Q4_N	0–3	1.74	0.75	0.92–1.29	1.84	0.91	1.11–2.17	0.67
Intensity and disturbance during silent task	Q5_N	1–3	1.38	0.63	1.27–1.49	1.49	0.72	1.49–2.38	0.19
	Q6_N	1–3	1.26	0.56	1.16–1.36	1.50	0.76	1.33–2.14	**0.01**
Intensity and disturbance during group activity	Q7_N	1–3	1.91	0.80	1.77–2.05	2.12	0.79	2.14–2.66	**0.03**
	Q8_N	1–3	1.72	0.79	1.58–1.86	1.94	0.83	1.86–2.67	**0.03**
Quality of voice	Q9_N	1–3	1.27	0.59	1.16–1.37	1.23	0.53	1.14–1.79	0.74
	Q10_N	1–3	1.50	0.65	1.39–1.62	1.52	0.73	1.28–2.15	0.78

Mean (M), standard deviation (SD) and the 95% confidence interval (95% CI) values are provided, and the p-values of significance for the differences between the two acoustic conditions, according to the MWU Test. Any statistically significant differences, with p-value < 0.05, are reported in bold.

scores for the happy children with the acoustic parameters. Most of the results are coherent and meaningful and can be useful to understand more about the relationships between environmental factors and children perception and behavior. The following interesting outcomes can be drawn:

- Generally, a positive relationship between disturbance from noise coming from outside the classroom, and in particular from traffic (Q1_N) and from adjacent rooms and corridor (Q3_N), and the noise level measured when the children were in silence, L_{N_Sil}, has been found, well underling as poor sound insulation is related to higher sound levels in classrooms and consistent higher noise disturbance.
- The higher is the slope of the speech level along the central line of the classroom, mL_S, or the lower is the slope when taken in absolute value, the higher is the perceived noise intensity and noise disturbance when the children are doing a task in silence (Q5_N and Q6_N, respectively). This is explainable considering that with higher values of speech slope the room is more reverberant and the speech level at the back of the classroom less clear, as well the noise is higher compared to less reverberant rooms.

- The higher the values of speech Clarity, C50, and Useful-to-detrimental-energy ratio, U50, both at the central position and average across the classroom, the lower are the noise intensity and disturbance scores Q3_N, Q5_N, Q6_N, Q7_N, and Q8_N. This means that good speech intelligibility corresponds to low noise disturbance both in the case of working in silence and in the case of group activity.
- As far as well-being is concerned, it is shown that the feeling of having fun is well related to higher values of speech intelligibility, in terms of U50_ctr, and the feeling of being happy of herself or himself is well related to lower values of mL_S, i.e., in rooms with higher slope of L_S, which means with less reverberation, children are happier with themselves.
- A particular remark is related to the parameter L_{N_gr}, which is unexpected negative related to the well-being aspect "proudness of myself." It means that the higher the level of noise during the group activity the higher is the proudness of the children. This result can be explained from the psychological point of view, as children were very happy while doing a lot of noise and they could be proud of having done it.

TABLE 6 | Unhappy students' subjective responses on the perceived well-being and noise disturbance, with the sample being grouped per classroom acoustics (i.e., good and bad).

Topic	ID	Scale	Good acoustics			Bad acoustics			MWU *p*-value
			(*N* = 15)			(*N* = 15)			
			M	SD	95% CI	*M*	SD	95% CI	
Well-being									
Self-esteem	Q1_WB	1–3	1.40	0.63	1.08–1.72	1.53	0.74	1.16–1.91	0.64
	Q2_WB	1–3	1.47	0.64	1.14–1.79	2.20	0.77	1.81–2.59	**0.01**
	Q3_WB	1–3	1.67	0.82	1.25–2.08	1.71	0.91	1.24–2.19	0.96
Emotional health	Q4_WB	1–3	1.67	0.62	1.35–1.98	1.53	0.64	1.21–1.86	0.52
	Q5_WB	1–3	1.13	0.35	0.96–1.31	1.73	0.88	1.29–2.18	**0.03**
	Q6_WB	1–3	2.00	0.93	1.53–2.47	2.27	0.88	1.82–2.71	0.54
Relationship at home and with friends	Q7_WB	1–3	1.67	0.82	1.25–2.08	1.73	0.96	1.25–2.22	0.96
	Q8_WB	1–3	1.33	0.49	1.09–1.58	1.60	0.83	1.18–2.02	0.47
	Q9_WB	1–3	1.80	0.86	1.36–2.24	2.00	0.93	1.53–2.47	0.55
Enjoyment of school	Q10_WB	1–3	1.50	0.76	1.10–1.90	2.07	1.00	1.55–2.59	0.12
	Q11_WB	1–3	2.27	0.46	2.04–2.50	2.47	0.52	2.21–2.73	0.26
	Q12_WB	1–3	1.50	0.76	1.10–1.90	2.62	0.77	2.20–3.00	**0.00**
Noise									
Traffic disturbance	Q1_N	0–3	2.00	1.00	0.87–3.00	1.00	0.00	/	0.12
Siren disturbance	Q2_N	0–3	2.22	0.67	1.79–2.66	1.50	1.00	0.52–2.48	0.14
Disturbance from internal noises	Q3_N	0–3	2.09	0.54	1.77–2.41	1.88	0.64	1.43–2.32	0.42
Rain disturbance	Q4_N	0–3	1.56	0.73	1.08–2.03	1.77	0.93	1.27–2.27	0.66
Intensity and disturbance during silent task	Q5_N	1–3	1.73	0.88	1.29–2.18	1.93	0.88	1.49–2.38	0.52
	Q6_N	1–3	1.47	0.64	1.14–1.79	1.73	0.80	1.33–2.14	0.36
Intensity and disturbance during group activity	Q7_N	1–3	2.07	0.80	1.66–2.47	2.40	0.51	2.14–2.66	0.24
	Q8_N	1–3	1.93	0.70	1.58–2.29	2.27	0.80	1.86–2.67	0.21
Quality of voice	Q9_N	1–3	1.53	0.74	1.16–1.91	1.47	0.64	1.14–1.79	0.89
	Q10_N	1–3	2.20	0.86	1.76–2.64	1.71	0.83	1.28–2.15	0.13

Mean (M), standard deviation (SD), and the 95% confidence interval (95% CI) values are provided and the p-values of significance for the differences between the two acoustic conditions, according to the MWU test. Any statistically significant differences, with a p-value < 0.05, are reported in bold.

With the aim to further clarify the relationships between subjective and objective parameters, **Figures 2**, **3** show, as example, the regression lines between the average values across classes of the answers to the questions Q3_N and Q5_N and the acoustic parameters L_{N_sil} and mL_S, respectively. These relationships have been chosen as they show two of the highest correlation coefficients in **Table 11**.

DISCUSSION

In-field studies on the effects of classroom acoustics on the perception of noise disturbance are only a few in the available literature. Furthermore, no study has investigated the perception of well-being at school and its relationships with acoustics. The present study has aimed at investigating three main aspects that are lacking in the available literature so far: (i) the influence of bad classroom acoustics on perceived well-being and noise disturbance for happy and unhappy children; (ii) the influence of class and school characteristics on perceived well-being and noise disturbance; (iii) the relationships between classroom

acoustic parameters and perceived well-being and noise disturbance scores.

Role of Bad Classroom Acoustics on Well-Being and Noise Disturbance for Happy and Unhappy Children

The classrooms involved in the study are representative of the typical acoustic quality available in the majority of Italian schools, with reverberation times under occupied conditions ranging from 0.5 to 1.4 s. Therefore, they could be grouped under GA and BA labels depending on their performance in agreement with recent studies, which suggest an optimal reverberation time range between 0.5 and 0.8 s to account for both speaking and listening premises at the same time. Based on the GA and BA clustering of classrooms, the data from subjective questionnaires administered to the pupils were analyzed to understand the extent to which GA and BA have an influence on the perception of well-being and noise disturbance at school. The presented results highlight that outdoor and indoor noise in BA is perceived as more disturbing than in GA, being significantly amplified by the higher reverberation. These outcomes corroborate previous

TABLE 7 | Subjective responses on the perceived well-being and noise disturbance of students in good acoustics only, with the sample being grouped per happiness (i.e., happy and unhappy).

Topic	ID	Scale	Happy subjects			Unhappy subjects			MWU p-value
			(N = 124)			(N = 15)			
			M	SD	95% CI	M	SD	95% CI	
Well-being									
Self-esteem	Q1_WB	1–3	1.22	0.49	1.13–1.30	1.40	0.63	1.08–1.72	0.18
	Q2_WB	1–3	1.43	0.73	1.30–1.56	1.47	0.64	1.14–1.79	0.58
	Q3_WB	1–3	1.14	0.45	1.06–1.22	1.67	0.82	1.25–2.08	**0.00**
Emotional health	Q4_WB	1–3	1.26	0.52	1.17–1.35	1.67	0.62	1.35–1.98	**0.00**
	Q5_WB	1–3	1.22	0.49	1.13–1.30	1.13	0.35	0.96–1.31	0.59
	Q6_WB	1–3	1.10	0.31	1.05–1.16	2.00	0.93	1.53–2.47	**0.00**
Relationship at home and with friends	Q7_WB	1–3	1.23	0.54	1.14–1.33	1.67	0.82	1.25–2.08	**0.01**
	Q8_WB	1–3	1.09	0.31	1.03–1.14	1.33	0.49	1.09–1.58	**0.00**
	Q9_WB	1–3	1.27	0.57	1.17–1.37	1.80	0.86	1.36–2.24	**0.00**
Enjoyment of school	Q10_WB	1–3	1.22	0.53	1.12–1.31	1.50	0.76	1.10–1.90	0.07
	Q11_WB	1–3	1.31	0.59	1.20–1.41	2.27	0.46	2.04–2.50	**0.00**
	Q12_WB	1–3	1.23	0.52	1.13–1.32	1.50	0.76	1.10–1.90	0.10
Noise									
Traffic disturbance	Q1_N	0–3	1.87	0.87	0.62–1.01	2.00	1.00	0.87–3.00	0.79
Siren disturbance	Q2_N	0–3	1.86	0.84	0.86–1.25	2.22	0.67	1.79–2.66	0.19
Disturbance from internal noises	Q3_N	0–3	1.79	0.83	0.93–1.31	2.09	0.54	1.77–2.41	0.18
Rain disturbance	Q4_N	0–3	1.74	0.75	0.92–1.29	1.56	0.73	1.08–2.03	0.47
Intensity and disturbance during silent task	Q5_N	1–3	1.38	0.63	1.27–1.49	1.73	0.88	1.29–2.18	0.11
	Q6_N	1–3	1.26	0.56	1.16–1.36	1.47	0.64	1.14–1.79	0.11
Intensity and disturbance during group activity	Q7_N	1–3	1.91	0.80	1.77–2.05	2.07	0.80	1.66–2.47	0.47
	Q8_N	1–3	1.72	0.79	1.58–1.86	1.93	0.70	1.58–2.29	0.24
Quality of voice	Q9_N	1–3	1.27	0.59	1.16–1.37	1.53	0.74	1.16–1.91	0.07
	Q10_N	1–3	1.50	0.65	1.39–1.62	2.20	0.86	1.76–2.64	**0.00**

Mean (M), standard deviation (SD), and the 95% confidence interval (95% CI) values are provided and the p-values of significance for the differences between the two happiness conditions, according to the MWU test. Any statistically significant differences, with a p-value < 0.05, are reported in bold.

studies that proved outdoor and indoor noise sources to be the most annoying for children of grades 2 and 6 (Dockrell and Shield, 2004) and high school students (Astolfi and Pellerey, 2008), respectively.

This study has shown that the 90.8% of the involved pupils reported to be happy. The subgroup of 9.2% of unhappy pupils has shown that the significant detrimental effect of BA was on the perception of serenity, self-pleasances, and fitting in at school, and not on increased noise disturbance. Overall, unhappy pupils have judged worse a higher number of well-being items in BA compared to GA, confirming the need of GA quality in classrooms to enhance well-being at school. This is consistent with a previous study by Klatte and Hellbrück (2010), who found that children attending classes under BA conditions judged their relationships to their peers and teachers less positively than did children from classrooms with GA.

Influence of Classes and School Characteristics on Well-Being and Noise Disturbance Perception

School and classroom characteristics need to be accounted for in a study that investigates the subjective perception as

it may be influenced by the subjects' background, origins, and features. To the aim of the present work, schools and classrooms had to present different characteristics in order to provide objective outcomes that were not affected by formal bias. Therefore, objective information was collected that consisted in student and building features, the presence of AT, the year of construction, and the TV.

In classrooms with better acoustics, a higher percentage of mother-tongue-speaking children was present. Such a result was corroborated by the lower perception of noise intensity and disturbance in quiet. Moreover, the students' feature of being speakers of the MT resulted in a higher feeling of being pleased with themselves. These outcomes suggest that being speakers of the MT is a child characteristic that brings more chances of attending classes in GA, being less disturbed, and having more well-being.

As far as school features are concerned, richer districts presented better acoustic conditions, resulting in turn in the fact that richer districts also revealed a greater perception of well-being.

Higher TVs were found to determine lower serenity and increase in noise disturbance. This outcome agrees with many studies, among which Klatte and Hellbrück (2010),

TABLE 8 | Subjective responses on the perceived well-being and noise disturbance of students in bad acoustics only, with the sample being grouped per happiness (i.e., happy and unhappy).

Topic	ID	Scale	Happy subjects (N = 172)			Unhappy subjects (N = 15)			MWU p-value
			M	SD	95% CI	M	SD	95% CI	
Well-being									
Self-esteem	Q1_WB	1–3	1.22	0.53	1.16–1.91	1.53	0.74	1.16–1.91	**0.03**
	Q2_WB	1–3	1.50	0.73	1.81–2.59	2.20	0.77	1.81–2.59	**0.00**
	Q3_WB	1–3	1.19	0.48	1.24–2.19	1.71	0.91	1.24–2.19	**0.01**
Emotional health	Q4_WB	1–3	1.33	0.59	1.21–1.86	1.53	0.64	1.21–1.86	0.12
	Q5_WB	1–3	1.28	0.57	1.29–2.18	1.73	0.88	1.29–2.18	**0.02**
	Q6_WB	1–3	1.15	0.41	1.82–2.71	2.27	0.88	1.82–2.71	**0.00**
Relationship at home and with friends	Q7_WB	1–3	1.24	0.57	1.25–2.22	1.73	0.96	1.25–2.22	**0.01**
	Q8_WB	1–3	1.12	0.39	1.18–2.02	1.60	0.83	1.18–2.02	**0.00**
	Q9_WB	1–3	1.24	0.47	1.53–2.47	2.00	0.93	1.53–2.47	**0.00**
Enjoyment of school	Q10_WB	1–3	1.35	0.70	1.55–2.59	2.07	1.00	1.55–2.59	**0.00**
	Q11_WB	1–3	1.30	0.54	2.21–2.73	2.47	0.52	2.21–2.73	**0.00**
	Q12_WB	1–3	1.33	0.61	2.20–3.00	2.62	0.77	2.20–3.00	**0.00**
Noise									
Traffic disturbance	Q1_N	0–3	1.94	0.83	0.00–0.44	1.00	0.00	/	0.05
Siren disturbance	Q2_N	0–3	1.90	0.77	0.00–0.94	1.50	1.00	0.52–2.48	0.28
Disturbance from internal noises	Q3_N	0–3	1.74	0.76	0.46–1.54	1.88	0.64	1.43–2.32	0.53
Rain disturbance	Q4_N	0–3	1.84	0.91	1.11–2.17	1.77	0.93	1.27–2.27	0.81
Intensity and disturbance during silent task	Q5_N	1–3	1.49	0.72	1.49–2.38	1.93	0.88	1.49–2.38	**0.04**
	Q6_N	1–3	1.50	0.76	1.33–2.14	1.73	0.80	1.33–2.14	0.19
Intensity and disturbance during group activity	Q7_N	1–3	2.12	0.79	2.14–2.66	2.40	0.51	2.14–2.66	0.22
	Q8_N	1–3	1.94	0.83	1.86–2.67	2.27	0.80	1.86–2.67	0.14
Quality of voice	Q9_N	1–3	1.23	0.53	1.14–1.79	1.47	0.64	1.14–1.79	0.05
	Q10_N	1–3	1.52	0.73	1.28–2.15	1.71	0.83	1.28–2.15	0.34

Mean (M), standard deviation (SD), and the 95% confidence interval (95% CI) values are provided and the p-values of significance for the differences between the two happiness conditions, according to the MWU test. Any statistically significant differences, with a p-value < 0.05, are reported in bold.

Crombie et al. (2011), and Lim et al. (2018), who proved a significant association between noise exposure to noise and high levels of noise annoyance, as well as decreased well-being, but all the studies took into account children older than 6–7 years old and were mostly based on questionnaires that were not filled in directly by children at school.

Relationships Between Classroom Acoustic Parameters and Well-Being and Noise Disturbance Scores

Overall, very high relationships were found between the acoustic quantities, corroborating past results (Bradley, 1986; Sato and Bradley, 2008; Yang and Bradley, 2009). The perceived noise intensity and disturbance were found to be strongly positively related to the slope of speech level along a classroom's main axis (mL$_S$) and strongly negatively related to speech clarity (C50) and the ratio of useful to detrimental energy (U50), both at the central position and on average across the classroom. These coherent outcomes are dependent on the reverberation time, which directly affects the uniformity of the sound energy in the room and increases the noise level: the higher is the reverberation time

in the room, the higher is the speech slope along the main axis (or lower when it is considered in absolute values) and the lower are the values of speech clarity and the ratio of useful to detrimental energy. Much should then be done in the acoustic design of classrooms to guarantee optimal distribution of parameters across all the listening positions. Promising results are achieved by combining the use of absorptive and diffusive surfaces (Choi, 2014), even though the perceptive effect of diffusive linings is still under research (Shtrepi et al., 2015, 2016, 2017).

The very high correlation between speech clarity and the ratio of useful to detrimental energy and the tight connection between central and mean values suggest the adoption of only one of the two indexes measured in the central position of the room to evaluate the classroom speech intelligibility. This outcome is in agreement with the results found by Puglisi et al. (2017), who suggested to measure clarity in the middle of the classroom to characterize its average behavior in terms of speech intelligibility.

U50 in the central position of the classroom was also strongly related to the feeling of having fun. A strong positive relationship was found for mL$_S$ and the feeling of children of being happy with themselves; that is, the higher the slope (or the lower in absolute value), the higher is the feeling of

TABLE 9 | Correlation matrix of the perceived well-being and noise disturbance scores of the happy children with the class and the school characteristics (with MT being mother tongue, DV being district real estate value, TV being traffic volume, and AT being acoustic treatment) and the reverberation time and the background noise level in the empty classrooms, i.e., T20_e and L_{N_sil}, respectively.

ID	Question or parameter	Class and school characteristics			
		MT	DV	TV	AT
MT	Mother tongue percentage	1			
DV	District real estate value		1		
TV	Traffic volume	−0.54	−0.22	1	
AT	Acoustic treatment		0.75	−0.27	1
T20_e (s)	Reverberation time in empty classroom	−0.31	−0.44	0.56	−0.63
L_{N_sil} (dBA)	Noise level in silence	−0.36		0.55	−0.35
Q1_WB	Proudness				
Q2_WB	Serenity			0.21	
Q3_WB	Fun				
Q4_WB	Holding many positive things about himself/herself				
Q5_WB	Gladness	−0.18			
Q6_WB	Cheerfulness				
Q7_WB	Relaxing at home		−0.18		
Q8_WB	Enjoyment with friends				
Q9_WB	Getting help from friends		−0.21		
Q10_WB	Pleasure of being at school				
Q11_WB	Safety at school				
Q12_WB	Fitting in at school				
Q1_N	Traffic disturbance			0.28	
Q2_N	Siren disturbance			0.22	
Q3_N	Disturbance from internal noises			0.21	
Q4_N	Rain disturbance				
Q5_N	Intensity during silent task	−0.20			
Q6_N	Disturbance during silent task	−0.19		0.16	
Q7_N	Intensity during group activity				
Q8_N	Disturbance during group activity				
Q9_N	Quality of the teacher' voice				
Q10_N	Quality of the classmates' voice				

Spearman correlation coefficients are given for significant relationships with a p-value < 0.01.

TABLE 10 | Correlation matrix of the acoustic parameters measured in the classrooms.

	Acoustic parameters									
	T20_e (s)	T20 (s)	L_{N_sil} (dBA)	L_{N_gr} (dBA)	L_{S_REF} (dBA)	mL_S (dBA/dd)	C50_ctr (dB)	C50_M (dB)	U50_ctr (dB)	U50_M (dB)
T20_e (s)	1									
T20 (s)	0.86	1								
L_{N_sil} (dBA)	0.27		1							
L_{N_gr} (dBA)	0.64	0.48	0.49	1						
L_{S_REF} (dBA)	0.73	0.59	0.35	0.40	1					
mL_S (dBA/dd)	0.57	0.51			0.36	1				
C50_ctr (dB)	−0.84	−0.90		−0.36	−0.55	−0.46	1			
C50_M (dB)	−0.88	−0.96		−0.56	−0.48	−0.48	0.91	1		
U50_ctr (dB)	−0.71	−0.87	−0.36	−0.46	−0.65	−0.30	0.90	0.82	1	
U50_M (dB)	−0.84	−0.90	−0.37	−0.67	−0.62	−0.46	0.80	0.93	0.89	1

Spearman correlation coefficients are given for significant relationships with a p-value < 0.01.

being unpleased with oneself. To the authors' knowledge, no studies are available in literature that considered the influence of speech intelligibility parameters and the effect of reverberation on the perceived well-being of young children. Anyway, the obtained results on the influence of acoustics on well-being at school are in line with those from Wålinder et al. (2007), who found that noise can be considered as a risk factor for the children's well-being in the school environment because it has primary consequences on the increased prevalence of symptoms of fatigue and headache.

TABLE 11 | Correlation matrix of the perceived well-being and noise disturbance scores for the happy children with the acoustic parameters.

ID	Question	Acoustic parameters								
		T20 (s)	L_{N_sil} (dBA)	L_{N_gr} (dBA)	L_{S_REF} (dBA)	mL_S (dBA/dd)	C50_ctr (dB)	C50_M (dB)	U50_ctr (dB)	U50_M (dB)
Q1_WB	Proudness			−0.18						
Q2_WB	Serenity									
Q3_WB	Fun								−0.16	
Q4_WB	Hold many positive things about himself/herself									
Q5_WB	Gladness					0.18				
Q6_WB	Cheerfulness									
Q7_WB	Relax at home									
Q8_WB	Enjoinment with friends									
Q9_WB	Get help from friends									
Q10_WB	Pleasure of being at school									
Q11_WB	Safety at school									
Q12_WB	Fit in at school									
Q1_N	Traffic disturbance	0.30		0.33						
Q2_N	Siren disturbance									
Q3_N	Disturbance from internal noises	0.26								−0.22
Q4_N	Rain disturbance									
Q5_N	Intensity during silent task					0.28				−0.17
Q6_N	Disturbance during silent task					0.22	−0.17	−0.17	−0.16	−0.18
Q7_N	Intensity during group activity	0.16			0.18		−0.17		−0.24	−0.17
Q8_N	Disturbance during group activity								−0.19	
Q9_N	Quality of the teacher's voice									
Q10_N	Quality of the classmates' voice									

Spearman correlation coefficients are given for significant relationships with a p-value < 0.01.

Furthermore, the noise levels measured inside the classrooms were found to be significantly correlated with the perception of outdoor noise sources, in particular from traffic and from adjacent rooms and corridor. This relationship supports the need of an accurate acoustic design of classrooms with a focus on the enhancement of sound insulation as suggested in Secchi et al. (2017), as it is related to higher sound levels in classrooms and consistent with higher noise disturbance.
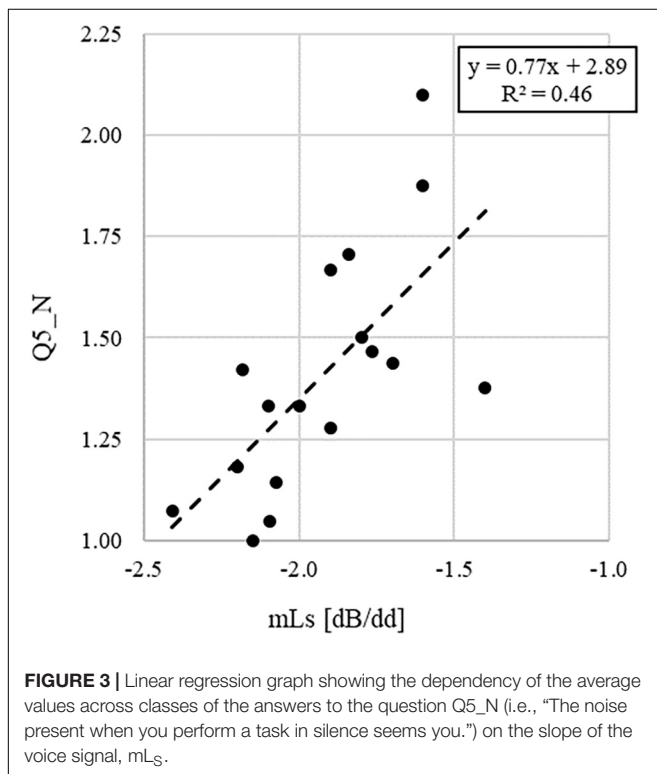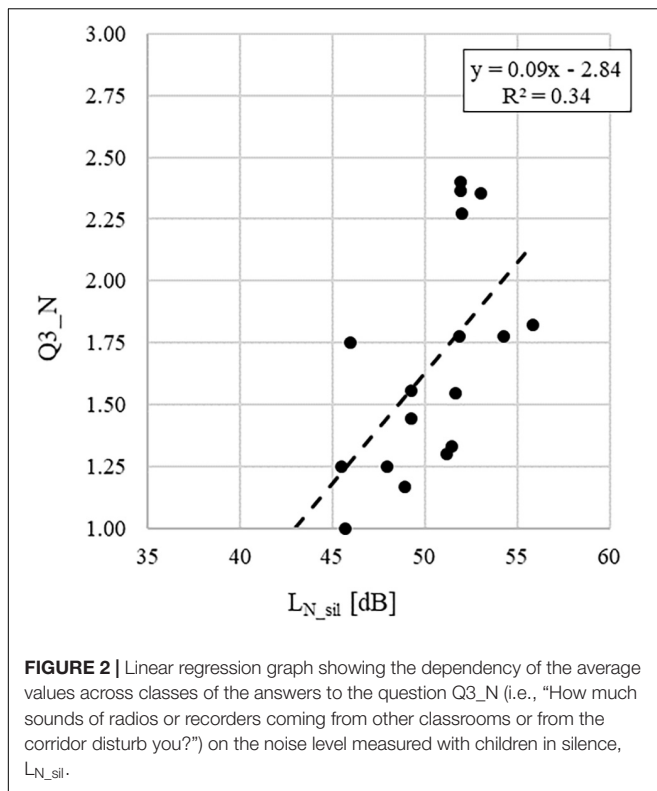
Strengths and Limitations of the Study

The present study confirms the association between bad classroom acoustics and noise disturbance for children at school. The main strength of the study is that the acoustic parameters needed to discriminate BA and GA were measured accurately, on the basis of referred protocols, and considers all the classroom acoustic aspects. In addition, this is the first study that examines the combined effect of noise disturbance and well-being perception of first graders at school. On the one side, the wide range of classroom acoustic conditions chosen for the study, ranging from 0.5 to 1.4 s of reverberation time, has ensured the covering of most of the current classrooms' features. On the other side, the school characteristics sufficiently covered most of the school buildings of a typical town.

There are also several limitations in the present study. The selection of the schools depended on their willingness to participate in the research, which was part of a wider study on the assessment and strengthening of the reading abilities of first graders (Astolfi et al., 2018b). Furthermore, the teachers of first graders volunteered to participate in the study, probably due to an already existing interest in the strengthening of the reading abilities of the pupils. The influence of the acoustic environment on the children's abilities was only proposed as a secondary outcome to the teachers. All the invited children of the classrooms (with the consent of their legal guardians) agreed to participate in the study.

Well-being was measured using an Italian translation of the questionnaire developed by Sabri et al. (2015), suitable for young people with SENs. It was selected as it was validated for children aged 6–7 years old and because it was especially designed for well-being investigations at school. The questionnaire included many items which may have affected their completion rate, especially for first graders. The items were *a priori* explained to the children by an experimenter, who was always the same, but this could have biased the answers, involuntarily.

In addition, the questionnaires filled in by children with hearing impairment (HI) or with SENs identified by the teachers were excluded, and these data have not been used for further

FIGURE 2 | Linear regression graph showing the dependency of the average values across classes of the answers to the question Q3_N (i.e., "How much sounds of radios or recorders coming from other classrooms or from the corridor disturb you?") on the noise level measured with children in silence, L_{N_sil}.

FIGURE 3 | Linear regression graph showing the dependency of the average values across classes of the answers to the question Q5_N (i.e., "The noise present when you perform a task in silence seems you.") on the slope of the voice signal, mL_S.

analyses. It is well known that children with HI or SENs are more susceptible to the adverse effects of BA, and future studies need to survey this fragile group of pupils. Moreover, because this is

a cross-sectional study, the causality of the association cannot be univocally determined. To investigate temporal changes in noise disturbance and well-being also for particular groups of children, as well as to infer causality, longitudinal studies are needed based on a larger population. Additional information on parent's and family conditions, both related to economic and educational issues and on well-being and noise disturbance perceived at home, should have been acquired.

An unexpected result consisted in the strong relationship that was found between L_{N_gr} and the well-being aspect "proudness of myself," as it was found that pupils were prouder of themselves as the level due to activity noise was higher in the classroom. This outcome should be deepened from a psychological point of view rather than an acoustic point of view, as it may be that due to the excitement of children when they were asked to be noisy, the measurement condition was biased. Thus, a proposal for the future is to measure other noise levels in the classrooms, related to real noise situations during lessons; however, it is needed to adopt an effective measurement procedure that avoids external influences on the behavior of pupils.

CONCLUSION

In the present study, subjective outcomes on the assessment of perceived well-being and noise disturbance at school for first graders were related to classroom acoustics characteristics and to classes and schools features.

The findings of this pilot study, which involved 326 first graders in their own classrooms, suggest that long reverberation times, which are associated with poor classroom acoustics as they generate higher noise levels and degrade speech intelligibility, bring pupils to a reduced perception of having fun and being happy with themselves. Furthermore, bad classroom acoustics is also related to an increased perception of noise intensity and disturbance, particularly in the case of traffic noise and noise from adjacent school environments.

Finally, an analysis of the perception of well-being and noise disturbance depending on the self-judgment of pupils on being happy or unhappy was performed. Happy pupils reported a higher perception of noise disturbance under bad classroom acoustic conditions, whereas unhappy pupils reported only complaints in bad classroom acoustics with respect to the perception of pleasances with himself or herself and of fitting in at school.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Politecnico di Torino.

Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

AA and GP contributed in the conception and design of the study. AA, GP, AP, SM, and GM participated in the in-field measurements and surveys. SM and GM organized the database. GP and AP drafted the "Introduction" section. GM and GP drafted the "Materials and Methods" section. AA, FP, and SM performed the statistical analysis. AA, GP, GM, and AP wrote the first draft of the manuscript. AA directed the research, wrote the "Results," "Discussion," and "Conclusion" sections, and drafted the final version of the manuscript. TS provided research assistance and commented on earlier versions of the manuscript.

REFERENCES

ANSI S3.5, (1997). *Methods for Calculation of the Speech Intelligibility Index.* New York, NY: Acoustical Society of America.

Astolfi, A., and Pellerey, F. (2008). Subjective and objective assessment of acoustical and overall environmental quality in secondary school classrooms. *J. Acoust. Soc. Am.* 123, 163–173. doi: 10.1121/1.2816563

Astolfi, A., Bottalico, P., Accornero, A., Garzaro, M., Nadalin, J., and Giordano, C. (2012a). "Relationship between vocal doses and voice disorders on primary school teachers," in *Proceedings of the Conference on Noise Control – EuroNoise2012*, Prague.

Astolfi, A., Bottalico, P., and Barbato, G. (2012b). Subjective and objective speech intelligibility investigations in primary school classrooms. *J. Acoust. Soc. Am.* 131, 247–257. doi: 10.1121/1.3662060

Astolfi, A., Carullo, A., Pavese, L., and Puglisi, G. E. (2015).). Duration of voicing and silence periods of continuous speech in different acoustic environments. *J. Acoust. Soc. Am.* 137, 565–579. doi: 10.1121/1.4906259

Astolfi, A., Castellana, A., Carullo, A., and Puglisi, G. E. (2018a). Uncertainty of speech level parameters measured with a contact-sensor-based device and a headworn microphone. *J. Acoust. Soc. Am.* 143, EL496–EL502. doi: 10.1121/1.5042761

Astolfi, A., Castellana, A., Puglisi, G. E., Fugiglando, U., and Carullo, A. (2019). Speech level parameters in very low and excessive reverberation measured with a contact-sensor-based device. *J. Acoust Soc. Am.* 145, 2540–2551. doi: 10.1121/1.5098942

Astolfi, A., Corrado, V., and Griginis, A. (2008). Comparison between measured and calculated parameters for the acoustical characterization of small classrooms. *Appl. Acoust.* 69, 966–976. doi: 10.1016/j.apacoust.2007.08.001

Astolfi, A., Puglisi, G. E., Minelli, G., Shtrepi, L., Prato, A., and Sacco, T. (2018b). "Effects of classroom acoustics on potential struggling first-grade readers," in *Proceedings of the Conference on Noise Control – EuroNoise2018*, Crete.

Bottalico, P., and Astolfi, A. (2012). Investigations into vocal doses and parameters pertaining to primary school teachers in classrooms. *J. Acoust. Soc. Am.* 131, 2817–2827. doi: 10.1121/1.3689549

Bottalico, P., Astolfi, A., and Hunter, E. J. (2017a). Teachers' voicing and silence periods during continuous speech in classrooms with different reverberation times. *J. Acoust. Soc. Am.* 141, El26–El31. doi: 10.1016/S0892-1997(03)00033-X

Bottalico, P., Graetzer, S., Astolfi, A., and Hunter, E. J. (2017b). Silence and voicing accumulations in Italian primary school teachers with and without voice disorders. *J. Voice* 31, 260.e11–260.e20. doi: 10.1016/j.jvoice.2016.05.009

Bottalico, P., Ipsaro Passione, I., Astolfi, A., Carullo, A., and Hunter, E. J. (2018). Accuracy of the quantities measured by four vocal dosimeters and its uncertainty. *J. Acoust. Soc. Am.* 143, 1591–1602. doi: 10.1121/1.5027816

Bradley, J. S. (1986). Speech intelligibility studies in classrooms. *J. Acoust. Soc. Am.* 80, 846–854. doi: 10.1121/1.393908

Bradley, J. S., Reich, R. D., and Norcross, S. G. (1999). On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility. *J. Acoust. Soc. Am.* 106, 1820–1828. doi: 10.1121/1.427932

Brännström, K. J., Johansson, E., Vigertsson, D., Morris, D. J., Sahlén, B., and Lyberg-Åhlander, V. (2017). How children perceive the acoustic environment of their school. *Noise Health* 19, 84–94. doi: 10.4103/nah.NAH-33-16

Calosso, G., Puglisi, G. E., Astolfi, A., Castellana, A., Carullo, A., and Pellerey, F. (2017). A one-school year longitudinal study of secondary school teachers' voice parameters and the influence of classroom acoustics. *J. Acoust. Soc. Am.* 142, 1055–1066. doi: 10.1121/1.4998707

Cardon, G., Campbell, J., and Sharma, A. (2012). Plasticity in the developing auditory cortex: evidence from children with sensorineural hearing loss and auditory neuropathy spectrum disorder. *J. Am. Acad. Audiol.* 23, 396–495. doi: 10.3766/jaaa.23.6.3

Carullo, A., Vallan, A., and Astolfi, A. (2013). "A low-cost platform for voice monitoring," in *Proceedings of the International Instrumentation and Measurement Technology Conference (I2MTC)*, Minneapolis.

Castellana, A., Carullo, A., Astolfi, A., Puglisi, G. E., and Fugiglando, U. (2017). Intra-speaker and inter-speaker variability in speech sound pressure level across repeated readings. *J. Acoust. Soc. Am.* 141, 2353–2363. doi: 10.1121/1.4979115

Choi, Y.-J. (2014). The application of diffusers for classroom acoustical design. *Noise Vib. Worldwide* 45, 8–16. doi: 10.1260/0957-4565.45.5.8

Crombie, R., Clark, C., and Stansfeld, S. A. (2011). Environmental noise exposure, early biological risk and mental health in nine to ten years old children: a cross-sectional field study. *Environ. Health* 10, 1–8. doi: 10.1186/1476-069X-10-39

Croux, C., and Dehon, C. (2010). Influence functions of the spearman and kendall correlation measures. *Statist. Methods Appl.* 19, 497–515. doi: 10.1007/s10260-010-0142-z

Di Blasio, S., Shtrepi, L., Puglisi, G. E., and Astolfi, A. (2019). A cross-sectional survey on the impact of irrelevant speech noise on annoyance, mental health and well-being, performance and occupants' behavior in shared and open-plan offices. *Int. J. Environ. Res. Public Health* 16:280. doi: 10.3390/ijerph16020280

Diener, E. (2006). Guidelines for national indicators of subjective wellbeing and ill-being. *J. Happiness Stud.* 7, 397–404. doi: 10.1007/s10902-006-9000-y

Din 18041. (2016). *Hörsamkeit in Räumen–Anforderungen Empfehlungen und Hinweise Für Die Planung (Acoustic Quality in Rooms–Specifications and Instructions for the Room Acoustic Design)*. Berlin: German Institute for Standardisation.

Dockrell, J. E., and Shield, B. (2004). Children's perceptions of their acoustic environment at school and at home. *J. Acoust. Soc. Am.* 115, 2964–2973. doi: 10.1121/1.1652610

Dreger, S., Meyer, N., Fromme, H., Bolte, G., Hendrowarsito, L., Frieß-Hesse, T., et al. (2015). Environmental noise and incident mental health problems: a prospective cohort study among school children in Germany. *Environ. Res.* 143, 49–54. doi: 10.1016/j.envres.2015.08.003

Epstein, J. L., and McPartland, J. M. (1976). The concept and measurement of the quality of school life. *Am. Educ. Res. J.* 13, 15–30. doi: 10.3102/00028312013001015

Evans, G. W., Lercher, P., Meis, M., Ising, H., and Kofler, W. W. (2001). Community noise exposure and stress in children. *J. Acoust. Soc. Am.* 109, 1023–1027. doi: 10.1121/1.1340642

Forns, J., Dadvand, P., Foraster, M., Alvarez-Pedrerol, M., Rivas, I., López-Vicente, M., et al. (2016). Traffic-Related air pollution, noise at school, and behavioral problems in Barcelona schoolchildren: a cross-sectional study. *Environ. Health Perspect.* 124, 529–535. doi: 10.1289/ehp.1409449

Goodman, R. (2001). Psychometric properties of the strengths and difficulties questionnaire (SDQ). *J. Am. Acad. Child Adolesc. Psychiatry* 40, 1337–1345. doi: 10.1097/00004583-200111000-00015

Guski, R., Felscher-Suhr, U., and Schuemer, R. (1999). The concept of noise annoyance: how international experts see it. *J. Sound Vib.* 223, 513–527. doi: 10.1006/jsvi.1998.2173

Haines, M. M., Stansfeld, S. A., Job, R. F., Berglund, B., and Head, J. (2001). Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children. *Psychol. Med.* 31, 265–277. doi: 10.1017/S0033291701003282

Hascher, T. (2008). Quantitative and qualitative research approaches to assess student well-being. *Int. J. Educ. Res.* 47, 84–96. doi: 10.1016/j.ijer.2007.11.016

Hjortebjerg, D., Andersen, A. M. N., Christensen, J. S., Ketzel, M., Raaschou-Nielsen, O., Sunyer, J., et al. (2016). Exposure to road traffic noise and behavioural problems in 7-year-old children: a cohort study. *Environ. Health Perspect.* 124, 228–234. doi: 10.1289/ehp.1409430

ISO 18233, (2006). *Acoustics: Application of New Measurement methods in Building and Room Acoustics*. Geneva: International Organization for Standardization.

ISO 3382-1, (2009). *Acoustics - Measurement of Room Acoustic Parameters – Part 1: Performance Spaces*. Genève: International Organization for Standardization.

ISO 3382-2, (2008). *Acoustics: Measurement of Room Acoustic Parameters– Reverberation Time in Ordinary Rooms*. Geneva: International Organization for Standardization.

ISO 9921, (2015). *Ergonomics: Assessment of Speech Communication*. Geneva: International Organization for Standardization.

Keyes, C. L. (ed.) (2013). *Mental Wellbeing. International Contributions to the Study of Positive Mental Health*. Georgia: Springer.

Klatte, M., and Hellbrück, J. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: a field study. *Environ. Behav.* 42, 659–692. doi: 10.1177/0013916509336813

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4:578. doi: 10.3389/fpsyg.2013.00578

Konu, A., Alanen, E., Lintonen, T., and Rimpelä, M. (2002). Factor structure of the school well-being model. *Health Educ. Res.* 17, 732–742. doi: 10.1093/her/17.6.732

Lim, J., Kweon, K., Kim, H. W., Cho, S. W., Park, J., and Sim, C. S. (2018). Negative impact of noise and noise sensitivity on mental health in childhood. *Noise Health* 20, 199–211. doi: 10.4103/nah.NAH-9-18

Minelli, G., Astolfi, A., Puglisi, G. E., Murgia, S., Shtrepi, L., Prato, A., et al. (2019). "Measuring classroom acoustics with a systematic approach," in *Proceedings of Internoise 2019*, Madrid.

OICT, (2019). *Turin Real Estate Market Observatory*. Available at: http://www.oict.polito.it/en/ (accessed May, 2019).

Pelegrín-García, D., Brunskog, J., and Rasmussen, B. (2014). Speaker-oriented classroom acoustics design guidelines in the context of current regulations in European countries. *Acta Acust. Acust.* 100, 1073–1089. doi: 10.3813/AAA.918787

Prodi, N., Visentin, C., and Feletti, A. (2013). On the perception of speech in primary school classrooms: Ranking of noise interference and of age influence. *J. Acoust. Soc. Am.* 133, 255–268. doi: 10.1121/1.4770259

Puglisi, G. E., Astolfi, A., Cantor Cutiva, L. C., and Carullo, A. (2015a). "Assessment of indoor ambient noise level in school classrooms," in *Proceedings of the Conference on Noise Control – EuroNoise2015*, Maastricht.

Puglisi, G. E., Cantor Cutiva, L. C., Astolfi, A., and Carullo, A. (2017). Four-day-follow-up study on the voice monitoring of primary school teachers: relationships with conversational task and classroom acoustics. *J. Acoust. Soc. Am.* 141, 441–452. doi: 10.1121/1.4973805

Puglisi, G. E., Prato, A., Sacco, T., and Astolfi, A. (2018). Influence of classroom acoustics on the reading speed: a case study on Italian second-graders. *J. Acoust. Soc. Am.* 141, EL144–EL149. doi: 10.1121/1.5051050

Puglisi, G. E., Warzybok, A., Hochmuth, S., Visentin, C., Astolfi, A., Prodi, N., et al. (2015b). An Italian matrix sentence test for the evaluation of speech intelligibility in noise. *Int. J. Audiol.* 54, 44–50. doi: 10.3109/14992027.2015.1061709

Riley, K. G., and McGregor, K. K. (2012). Noise hampers children's expressive word learning. *Lang. Speech Hear. Serv. Sch.* 43, 325–337. doi: 10.1044/0161-1461

Rudner, M., Lyberg-Åhlander, V., Brännström, J., Nirme, J., Pichora-Fuller, M. K., and Sahlén, B. (2018). Listening comprehension and listening effort in the primary school classroom. *Front. Psychol.* 9:1193. doi: 10.3389/fpsyg.2018.01193

Ryff, C. D., and Singer, B. (1998). The contours of positive human health. *Psychol. Inq.* 9, 1–28. doi: 10.1207/s15327965pli0901-1

Sabri, F., Rotheroe, A., and Kazimirski, A. (2015). *Measuring the Well-Being of Young People with Special Educational Needs*. London: New Philanthropy Capital.

Sato, H., and Bradley, J. S. (2008). Evaluation of acoustical conditions for speech communication in working elementary school classrooms. *J. Acoust. Soc. Am.* 123, 2064–2077. doi: 10.1121/1.2839283

Secchi, S., Astolfi, A., Calosso, G., Casini, D., Cellai, G., Scamoni, F., et al. (2017). Effect of outdoor noise and façade sound insulation on indoor acoustic environment of Italian schools. *Appl. Acoust.* 126, 120–130. doi: 10.1016/j.apacoust.2017.05.023

Shield, B. M., and Dockrell, J. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Shtrepi, L., Astolfi, A., D'antonio, G., and Guski, M. (2016). Objective and perceptual evaluation of distance-dependent scattered sound effects in a small variable-acoustics hall. *J. Acoust. Soc. Am.* 140, 3651–3662. doi: 10.1121/1.4966267

Shtrepi, L., Astolfi, A., Pelzer, S., Vitale, R., and Rychtáriková, M. (2015). Objective and perceptual assessment of the scattered sound field in a simulated concert hall. *J. Acoust. Soc. Am.* 138, 1485–1497. doi: 10.1121/1.4929743

Shtrepi, L., Astolfi, A., Puglisi, G. E., and Masoero, M. C. (2017). Effects of the distance from a diffusive surface on the objective and perceptual evaluation of the sound field in a small simulated variable-acoustics hall. *Appl. Sci.* 224, 1–23. doi: 10.3390/app7030224

Sigel, S., and Castellan, N. J. (1988). *Non Parametric Statistics for the Behavioral Sciences*, 2nd Edn, New York, NY: McGraw-Hill.

Stansfeld, S. A., Clark, C., Cameron, R. M., Alfred, T., Head, J., Haines, M. M., et al. (2009). Aircraft and road traffic noise exposure and children's mental health. *J. Environ. Psychol.* 29, 203–207. doi: 10.1016/j.jenvp.2009.01.002

Stansfeld, S. A., Haines, M. M., Burr, M., Berry, B., and Lercher, P. (2000). A review of environmental noise and mental health. *Noise Health* 2, 1–8.

Tobia, V., Greco, A., Steca, P., and Marzocchi, G. M. (2019). Children's wellbeing at school: a multi-dimensional and multi-informant approach. *J. Happiness Stud.* 20, 841–861. doi: 10.1007/s10902-018-9974-2

UK Department of Transport, (2012). Available at: https://www.gov.uk/government/publications/guidance-on-road-classification-and-the-primary-route-network (accessed October 11, 2019).

UNI 11367 (2010). "Acustica in edilizia - classificazione acustica delle unità immobiliari - procedura di valutazione e verifica in opera," in *Building Acoustics - Acoustic Classification of Living Units - Evaluation and in Field Verification procedures*. Milan, Italy: Ente Italiano di Normazione.

Wålinder, R., Gunnarsson, K., Runeson, R., and Smedje, G. (2007). Physiological and psychological stress reactions in relation to classroom noise. *Scand. J. Work Environ. Health* 33, 260–266. doi: 10.5271/sjweh.1141

Winblad, U., and Dudley, E. (1997). *Primary School Physical Environment and Health: WHO Global School Health Initiative*. Geneva: World Health Organization.

World Health Organisation [WHO], (2014). Available at: https://www.who.int/features/factfiles/mental_health/en/ (accessed May, 2019).

WSP, (2015). *Acoustic Design in Schools: Performance Standards Building Bulletin, BB93*. London: Educational Funding Agency.

Yang, W., and Bradley, J. S. (2009). Effects of room acoustics on the intelligibility of speech in classrooms for young children. *J. Acoust. Soc. Am.* 125, 922–933. doi: 10.1121/1.3058900

Investigating the Effect of One Year of Learning to Play a Musical Instrument on Speech-in-Noise Perception and Phonological Short-Term Memory in 5-to-7-Year-Old Children

Douglas MacCutcheon[1,2]*, Christian Füllgrabe[3], Renata Eccles[2,4], Jeannie van der Linde[2,4], Clorinda Panebianco[2] and Robert Ljung[1]

[1] Department of Building, Energy and Environmental Engineering, Högskolan i Gävle, Gävle, Sweden, [2] Department of Music, University of Pretoria, Pretoria, South Africa, [3] School of Sport, Exercise and Health Sciences, Loughborough University, Loughborough, United Kingdom, [4] Department of Speech-Language Pathology and Audiology, University of Pretoria, Pretoria, South Africa

The benefits in speech-in-noise perception, language and cognition brought about by extensive musical training in adults and children have been demonstrated in a number of cross-sectional studies. Therefore, this study aimed to investigate whether one year of school-delivered musical training, consisting of individual and group instrumental classes, was capable of producing advantages for speech-in-noise perception and phonological short-term memory in children tested in a simulated classroom environment. Forty-one children aged 5–7 years at the first measurement point participated in the study and either went to a music-focused or a sport-focused private school with an otherwise equivalent school curriculum. The children's ability to detect number and color words in noise was measured under a number of conditions including different masker types (speech-shaped noise, single-talker background) and under varying spatial combinations of target and masker (spatially collocated, spatially separated). Additionally, a cognitive factor essential to speech perception, namely phonological short-term memory, was assessed. Findings were unable to confirm that musical training of the frequency and duration administered was associated with a musicians' advantage for either speech in noise, under any of the masker or spatial conditions tested, or phonological short-term memory.

Keywords: speech in noise, phonological short-term memory, musical training, children, cognition

INTRODUCTION

Children receive their education in acoustic environments in which background noise is nearly always present. Classroom noise is known to cause distraction and annoyance in children, but its primary effect is a reduction in speech intelligibility (for reviews, see Shield and Dockrell, 2003; Klatte et al., 2013), with a consequently negative impact on academic achievement (Shield and Dockrell, 2008). In typically developing children, the ability to cope with speech in noise (SiN) has been linked to individual differences in cognitive and language

abilities (Nelson et al., 2005; Strait et al., 2012; MacCutcheon et al., 2019), age (Corbin et al., 2016), gender (Prodi et al., 2019), and supra-threshold auditory processing abilities (Lorenzi et al., 2000), as well as environmental factors, including reverberation and the spatial, spectral and temporal characteristics of the background noise (MacCutcheon et al., 2018, 2019; McCreery et al., 2019).

Many studies have focused on how manipulating the acoustic environment can improve children's attention to verbal instructions, self-rated ability to cope with noise, speech reception thresholds (SRTs) and cognitive performance (DiSarno et al., 2002; Purdy et al., 2009; Dockrell and Shield, 2012; Prodi et al., 2019). Contrastingly, the aim of the present study is to investigate whether musical training can improve individual characteristics of the listener that contribute to speech perception (e.g., auditory, linguistic and cognitive abilities) and thereby mitigate speech-intelligibility challenges posed by noise.

Musical training has been suggested as a possible candidate for improving auditory, linguistic and cognitive abilities (Patel, 2011; Tallal, 2014) because a multitude of studies indicate that adults and children with musical training show greater motor, cognitive, linguistic and auditory skills (for a review, see Benz et al., 2016), referred to as the "musicians' advantage" (Başkent and Gaudrain, 2016; Talamini et al., 2017). Indeed, a musicians' advantage for SiN perception has been reported by a number of studies in adults and children (Parbery-Clark et al., 2009; Strait et al., 2012, 2013; Bidelman et al., 2014; Kraus et al., 2014; Slater et al., 2015; Başkent and Gaudrain, 2016). However, there are also a substantial number of studies that failed to find strong evidence in favor of advantages in musicians (Strait et al., 2012; Fuller et al., 2014; Ruggles et al., 2014; Boebinger et al., 2015; Fleming et al., 2019; Zendel et al., 2019).

Despite diverging findings, there is a compelling theoretical basis for the possibility that musical training could improve speech perception. Indeed, due to the similarity of the acoustic features of music and speech, these stimuli are processed by the same brain networks (Patel, 2011). For example, both music and speech perception require the processing of fluctuations in the amplitude envelope of the acoustic signal (Patel, 2011) to discriminate musical notes and phrases and segments of syllables and words, respectively. Additionally, pitch processing (the ability to perceptually discriminate between frequencies) is both an essential aspect of the emotional and linguistic content of speech as well as the harmonic and melodic content of music.

How and why abilities developed through musical training might lead to improvements in SiN processing is currently still unknown. In this study, we consider three possibilities. The first is that musical training confers benefits for dealing with energetic and/or informational maskers; the second is that musical training improves spatial listening; and the third is that musical training confers benefits for SiN perception by improving mediating cognitive processes.

Noise presents a challenge for speech perception as a consequence of the acoustic and spatial characteristics of the masker. Energetic maskers reduce speech intelligibility, while informational maskers reduce speech perception due to acoustic similarity with the target speech, resulting in perceptual confusion (Brungart, 2001; Wightman and Kistler, 2005; Wightman et al., 2006; MacCutcheon et al., 2019), and informational interference (Dole et al., 2012; Stone et al., 2012). Meanwhile, localization cues provided by the spatial separation of the target speech from the masker can improve intelligibility because timing and level differences between the two ears assist with sound segregation (Litovsky, 2005; Johnstone and Litovsky, 2006); referred to as "spatial release from masking" (Freyman et al., 1999; Hawley et al., 2004). However, assessments of the potential for musical training to help speech perception under these acoustic and spatial conditions have produced mixed results (Parbery-Clark et al., 2009; Strait et al., 2012; Swaminathan et al., 2015) and there is a dearth of longitudinal studies in children in the literature.

The development of SiN perception occurs in conjunction with cognitive development (Hall et al., 2002; Bradley and Sato, 2008; Neuman et al., 2010). According to the Ease of Language Understanding model (Rönnberg et al., 2008), noise places demands on cognitive processing of speech as working memory resources are required for assisting with the matching of incoming phonological information with phonological representations stored in long term memory. Meanwhile, explicit processing resources are also used for making guesses (informed by prior knowledge and experience as well as contextual factors) that might provide clues as to the nature of the missing input. This turns a relatively automatic task into a cognitively demanding, effortful task. Both cross-sectional and longitudinal studies have shown musical-training-induced improvements in cognitive functioning in adults and children (Benz et al., 2016). In particular, phonological short-term memory processes essential for SiN perception seem to be higher in child and adult musicians than in non-musician controls (Chan et al., 1998; Lee et al., 2007; Franklin et al., 2008; Strait et al., 2012, 2013; Bergman Nutley et al., 2014; Roden et al., 2014).

The present study builds longitudinally on a previous cross-sectional study by MacCutcheon et al. (2019). The study investigated whether individual differences in linguistic and cognitive abilities contribute to SiN perception in a variety of listening conditions, composed of different masker types and spatial configurations of the target speech and masker. Participants were typically developing children in early stages of development that are critical to the co-development of language (Rhyner, 2009) and speech perception (Johnstone and Litovsky, 2006). The results of MacCutcheon et al. (2019) indicated that, under certain listening conditions, memory span and expressive language provided benefits for SiN perception. The present study adds to these findings by longitudinally assessing the effect of 1 year of musical training on SiN perception and phonological short-term memory. Children attended one of two schools with equivalent academic curriculums, except that one school offered additional music lessons as part of the school curriculum while the

other school offered additional sports activities. Based on the published literature, it was hypothesized that musical training minimizes the effect of energetic and/or informational masking on speech perception and maximizes the use of spatial cues, resulting in improved speech perception relative to the control group. An additional hypothesis was that musical training improves speech perception via improvements in phonological short-term memory.

Previous studies reporting evidence for a musicians' advantage provided a higher frequency and longer duration of musical training for their participants than the present study. For example, Kraus et al. (2014)'s and Slater et al. (2015)'s children received up to 4 h of musical training per week for up to 2 years before a musicians' advantage was discernible. Although lesson frequencies and lengths for beginners learning an instrument are by no means standardized, norms suggest that children who show an interest in music will initially receive a lesson in their primary instrument once per week. Beginner instrumental lesson times for young children are generally 30–60 min depending on the child's innate musical abilities and attentional capacity as well as practicalities such as parental preferences and resources. As this range is more representative of what the majority of children engaging in musical activities at that age receive under "normal" circumstances, the present study hoped to ascertain a musicians' advantage within a shorter timeframe and with a lower intensity of musical training than previous studies.

MATERIALS AND METHODS

Participants

A total of 41 typically developing male school children participated in the study. On average, they were aged 6.3 years (standard deviation = 0.5 years, range: 5–7 years) at the start of the study, and had no history of cognitive, sensory or behavioral deficits, according to parental report. Parents of children in the participating schools received an information letter through the schoolteacher and agreed for their children to participate by providing written consent. Ethical approval for the study was granted by the University of Pretoria Research Ethics Committee, Approval 25071999 (GW20171130HS).

Prior to participation, all children were screened for hearing deficits. Normal hearing function was established using the smartphone hearing-screening application hearScreen[TM] that detects hearing losses in excess of 20 dB Hearing Level at 1, 2, and 4 kHz with 97.8% reliability compared to standard manual audiometric procedures (Swanepoel et al., 2014). The application was run on Samsung Galaxy J2 mobile phones connected to Sennheiser HD280 Pro headphones.

Musical Training and Control Groups

Twenty-six participants attended a music-focused school (the musical-training group) where they received up to 1 h per week of instrumental training over the course of a 38-week school year. The training was delivered by a qualified music teacher

who used a combination of Kodaly and Orff methodologies.[1] All children attended a 30-min group recorder lesson, and twelve (29%) children received a further 30-min individual piano or violin lesson. The remaining fifteen participants attended a sports-focused school (the control group) where they participated in extra-curricular sports (e.g., football, cricket, hockey and swimming) for 2–5 h per week. Both schools otherwise followed an equivalent Independent Examinations Board academic curriculum. As part of this curriculum, all children attended a weekly 30-min general group music lesson that did not involve instrumental training. None of the participants received additional musical training outside school.

The musical-training and control groups did not differ in age [$t(39) = 1.38$, $p = 0.177$, two-tailed], and socio-economic status as measured by maternal education level [$t(39) = 0.39$, $p = 0.695$, two-tailed]. Both groups were tested on the SiN and FDS tasks twice: once at the first assessment point (T1) when none of the participants had received any formal musical training, and then again at the second assessment point (T2) after attending their respective schools for 1 year. Between-group differences in language ability were also measured using the Renfew Action Picture Test (RAPT; Renfrew, 1980). This test consists of 10 pictures that must be verbally described (e.g., a girl hugging a teddy-bear), and the information and grammar content of the responses are scored out of 40 and 35 points, respectively. No group differences in language ability were detected at T1 [$t(39) = -0.10$, $p = 0.922$, two-tailed].

Design

A 2 Groups (musical training vs. control) × 2 Assessment points (T1 vs. T2) × 2 Masker types [speech-shaped noise (SSN) vs. single talker] × 2 Spatial locations (collocated vs. spatially separated) mixed design was used. Speech-in-noise intelligibility was analyzed separately for each group at the two assessment points in each of the four listening conditions obtained by combining masker type and spatial location, as well as averaged across listening conditions.

Tasks

Speech-in-Noise Perception

The SiN test was run on a DELL Latitude E6430 laptop, and the auditory stimuli were presented to the participants through a Focusrite Scarlett 2i2 audio interface and Sennheiser HD 650 headphones. All stimuli were pre-recorded and acoustics were simulated in a virtual classroom with a mean mid-frequency reverberation time T_{30} of 0.6 s using the software Room Acoustics for Virtual Environments (RAVEN; Schröder, 2011). Binaural room impulse responses were simulated based on a head-related transfer function measured from a child dummy head so that the

---

[1]The musical training taught the following musical concepts: pitch (identify and produce high and low pitches, identify and produce pitch contours), duration (identify and produce long and short sounds), beat (keeping steady beat to music through movement and instrumental play), timbre (identify sounds through aural cues, identify instrument families), dynamics (getting louder and softer), form structure (introducing common form structures including AB, ABA, and Rondo form), rhythm (producing crotchet and quaver rhythmic patterns, creating own rhythmic patterns) and creativity (creating a "sound story").

virtually simulated environment was appropriate for the sample under investigation (Fels et al., 2004). Further details about the masker and the simulation of the virtual acoustic environment are reported in MacCutcheon et al. (2019). Speech identification was assessed using an adaptation of the "Children's Coordinate Response Measure" software described in Vickers et al. (2016). The task was to identify two target words in the carrier sentence "show the dog where the [number word] [color word] is," spoken by an adult male with an English accent. The color word was one of six colors (black, red, green, white, blue or pink) and the number word was a number between one and nine, with the exception of the disyllabic number seven. The location of the target talker was simulated to be at 0° azimuth. The target speech was accompanied by either a single male adult talker reading fictitious news items, or SSN with the same long-term average speech spectrum as the masking talker. The masker started and ended with the target sentence. Within the simulated virtual environment, each masker was either collocated with the target talker, or spatially separated to the right of the target talker, at +90° azimuth. SRTs for identifying the two target words correctly 50% of the time were assessed. The presentation level of the masker was fixed at 55 dB(A) while the presentation level of the target speech, initially set to 68 dB(A), was adaptively varied, using a 1-up, 1-down procedure (Levitt, 1971). Until the first incorrect response, the presentation level for the target speech was decreased by 8 dB. Then, a step size of 4 dB was used until the second incorrect response occurred. Thereafter, the step remained fixed to 2 dB. Each threshold run was composed of 48 sentences, corresponding to all possible color-number combinations. The SRT was computed as the mean of the final four reversals for a given threshold run.

Phonological Short-Term Memory Capacity

The "Number Repetition – Forward" subtest from the Clinical Evaluation of Language Fundamentals (CELF-4; Semel et al., 2003) was used to assess phonological short-term memory capacity. This version of a forward digit span (FDS) test required the participant to recall number sequences of varying length

(from two to nine digits) in serial order. Initially, the sequence was composed of two digits and the sequence length was increased by one digit after two sequences of the same length were presented. The test was terminated once the participant incorrectly recalled two sequences of the same sequence length in a row, or completed all the lists. Each correctly recalled sequence was awarded a point, resulting in a maximum score of 16 points. Raw scores were converted to age-normed standard scores provided in the CELF-4 manual and all further analyses were conducted using standard scores.

Experimental Procedure

Testing was conducted in a sound-isolated music room of one of the participating schools in the presence of an experimenter. For the SiN test, the graphical user interface showed a photograph of a dog beside six colored panels, each subdivided into nine numbered buttons representing all possible number and color combinations. Given their young age, participants were asked to repeat verbally the number and color they had heard, and the experimenter entered the responses for them by clicking the appropriate buttons on the user interface. The order of the four listening conditions was counterbalanced using a Latin square design. The FDS test was administered according to the protocol provided in the manual of the CELF-4.

RESULTS

Results for the two groups on the short-term memory task and the speech-perception task in the four different listening conditions and on average are given in **Table 1** for the first and second assessment point.

Baseline Performance

At the start of the study (i.e., at T1), the two groups did not differ significantly in SRTs averaged across the four listening conditions [$t(39) = 0.017$, $p = 0.987$, two-tailed]. However, there

TABLE 1 | Group summary statistics in terms of mean, standard deviation (SD), and the lower and upper range of the 95% confidence interval (CI 95%) for performance on the forward digit span (FDS) test and speech-in-noise perception (SiN) test in each listening condition and on average.

Test	Assessment point	Control group			Musical training group		
		Mean (dB)	SD	CI 95% [lower, upper]	Mean (dB)	SD	CI 95% [lower, upper]
FDS	T1	6.4	2.1	[5.2, 7.6]	7.9	2.0	[7.0, 8.7]
	T2	7.1	1.9	[6.1, 8.2]	8.2	1.5	[7.6, 8.8]
SiN: Collocated * SSN	T1	−3.1	3.9	[−5.3, −0.9]	−3.1	4.7	[−5.0, −1.2]
	T2	−4.1	6.2	[−7.6, −0.7]	−5.6	2.8	[−6.7, −4.4]
SiN: Separated * SSN	T1	−4.7	4.1	[−7.0, −2.4]	−3.8	5.3	[−5.9, −1.6]
	T2	−5.5	3.3	[−7.3, −3.7]	−7.1	2.5	[−8.1, −6.1]
SiN: Collocated * single talker	T1	4.6	3.7	[2.5, 6.6]	5.2	3.6	[3.8, 6.7]
	T2	1.8	2.4	[0.5, 3.1]	1.8	3.8	[0.3, 3.4]
SiN: Separated * single talker	T1	0.9	5.0	[−1.8, 3.7]	−0.7	5.5	[−3.0, 1.5]
	T2	−4.4	4.0	[−6.7, −2.2]	−4.4	3.6	[−5.9, −3.0]
SiN: Average	T1	−0.6	2.5	[−2.0, 0.8]	−0.6	3.3	[−1.9, 0.8]
	T2	−3.1	2.5	[−4.5, −1.7]	−3.8	1.9	[−4.6, −3.1]

was a significant group difference on the FDS task [$t(39) = -2.49$, $p = 0.013$, two-tailed].

Effect of Musical Training, Noise-Type, Spatial Factors and Time on Speech-in-Noise Perception

To determine whether additional musical training over 1 year yielded improvements in SiN perception, a repeated-measures analysis of variance (ANOVA) was conducted on the SRTs, with Group as the between-subjects factor, and Assessment point, Masker type and Spatial location as within-subjects factors. Estimated marginal means for all main effects and interactions are provided in **Table 2**.

The main effect of Assessment point indicated that both groups' SiN perception was significantly better by 2.9 dB after 1 year [$F(1,39) = 33.54$, $p < 0.001$, $\eta_p^2 = 0.46$] consistent with findings that SiN perception improves with age (Hall et al., 2002). The significant main effect of Masker type [$F(1,39) = 123.68$, $p < 0.001$, $\eta_p^2 = 0.76$] indicated that the presence of a single talker led to an increase in SRTs by 5.2 dB compared to spectrally matched noise, across both groups and assessment points. The relative increase in perceptual difficulty experienced when the masker was a single talker is attributable to the acoustic similarity of the target and the masker with resulting informational interference (Dole et al., 2012; Stone et al., 2012), as well as the audible semantic content of the masker, which effectively captures attention in children (Cowan et al., 1999). The significant main effect of Spatial location [$F(1,39) = 59.25$, $p < 0.001$, $\eta_p^2 = 0.60$] indicated that across Group, Assessment point and Masker type factors, the average SRT in the collocated listening conditions was 3.4-dB higher compared to spatially separated listening conditions. This corroborates studies with adults and children indicating a benefit of spatially separating target and maskers (Litovsky, 2005; Johnstone and Litovsky, 2006).

TABLE 2 | Estimated marginal mean speech-reception thresholds (SRTs), standard error (SE) and the lower and upper range of the 95% confidence intervals (CI 95%) for the main effects and interactions.

Factors	Mean (dB)	SE	CI 95% [lower, upper]
T1	−0.6	0.5	[−1.6, 0.4]
T2	−3.5	0.4	[−4.2, −2.7]
SSN	−4.6	0.4	[−5.5, −3.8]
Single talker	0.6	0.4	[−0.2, 1.4]
Collocated	−0.3	0.4	[−1.1, 0.5]
Separated	−3.7	0.5	[−4.6, −2.8]
Collocated * SSN	−4.0	0.6	[−5.1, −2.8]
Collocated * Single talker	3.4	0.5	[2.4, 4.3]
Separated * SSN	−5.3	0.5	[−6.3, −4.3]
Separated * Single talker	−2.2	0.6	[−3.4, −0.9]
SSN * T1	−3.7	0.6	[−4.9, −2.4]
SSN * T2	−5.6	0.4	[−6.5, −4.7]
Single talker * T1	2.5	0.5	[1.4, 3.6]
Single talker * T2	−1.3	0.4	[−2.2, −0.4]

The interaction between Assessment point and Group was not significant [$F(1,39) = 0.59$, $p = 0.448$, $\eta_p^2 = 0.018$], suggesting that the two groups did not differ in SiN perception, neither at baseline nor after providing additional musical training to one of the groups.

An interaction between Masker type and Spatial location and subsequent simple-effects analysis indicated that when the masker was SSN, speech in the collocated condition was significantly harder to perceive by 1.3 dB than in the spatially separated condition. When the masker was a single talker, this difference increased to 5.3 dB. This 4-dB difference in spatial release from masking shows that spatial cues are more helpful for children's speech perception when dealing with realistic changing-state maskers that would often be present in the classroom environment. Furthermore, SRTs for the collocated condition were 7.3 dB higher in the presence of a single talker than in SSN, indicative of the burden that masker-target similarity and attention capture place on auditory stream segregation in children.

A significant interaction was found between Masker type and Spatial location [$F(1,39) = 15.38$, $p < 0.001$, $\eta_p^2 = 0.28$]. A simple-effect analysis revealed that spatially separating the masker from the target resulted in better SiN perception regardless of the type of masker: when the masker was SSN, speech in spatially separated conditions was significantly easier to perceive by 1.3 dB than when collocated [$F(1,39) = 4.12$, $p = 0.05$, $\eta_p^2 = 0.095$], but when the masker was a single talker, this increase between separated and collocated conditions grew to 5.5 dB [$F(1,39) = 54.61$, $p < 0.001$, $\eta_p^2 = 0.5$]. Furthermore, under both spatial conditions, speech masked by SSN was more intelligible than when masked by the single talker: when the masker was spatially separated, speech perception masked by SSN was 7 dB easier to discern than the single talker [$F(1,39) = 21.39$, $p < 0.001$, $\eta_p^2 = 0.35$], but this difference decreased to 3 dB when the masker was collocated but remained significant [$F(1,39) = 94.91$, $p < 0.001$, $\eta_p^2 = 0.71$].

Another significant interaction was found between Masker type and Assessment point [$F(1,39) = 7.79$, $p = 0.008$, $\eta_p^2 = 0.17$]. The simple effects analysis indicated that at both assessment points, SSN was the less challenging masker: SRTs at T1 were 6.2 dB better for SSN than for the single talker masker [$F(1,39) = 102.02$, $p < 0.001$, $\eta_p^2 = 0.72$], and at T2, the difference was reduced to 4.3 dB but remained significant [$F(1,39) = 62.43$, $p < 0.001$, $\eta_p^2 = 0.62$]. Furthermore, the improvement between the two assessment points was greater for the single talker than SSN: when the masker was SSN, the significant increase from T1 to T2 was almost 2 dB [$F(1,39) = 9.04$, $p = 0.005$, $\eta_p^2 = 0.19$], and this increase between assessment points grew to 3.8 dB when the masker was a single talker [$F(1,39) = 47.41$, $p < 0.001$, $\eta_p^2 = 0.55$]. This suggests that there are different developmental trajectories for coping with energetic and informational maskers. While the effect of the energetic masker (SSN) takes place in the auditory periphery, the effect of the informational masker (single talker) is located more centrally and probably involves cognitive processes. That the developmental effect was larger in the single-talker masker indicates that cognitive abilities which assist with SiN

perception develop faster than those attributable to peripheral auditory processing.

Effect of Musical Training on Phonological Short-Term Memory

A repeated-measures ANOVA, with the between-subjects factor Group and the within-subjects factor Assessment point, was conducted on the FDS scores to determine whether additional musical training yielded improvements in phonological short-term memory. There was a significant effect of Group $[F(1,39) = 9.54, p = 0.004, \eta_p^2 = 0.197]$, with higher FDS score in the musical training group at both baseline and T2 $[t(39) = -1.84, p = 0.022$, two-tailed]. Within-subject effects indicated that, relative to T1, the average FDS score increased from 10.2 (SD = 3.1) to 10.4 points (SD = 2.5) at T2, but this increase was not significant $[F(1,39) = 0.17, p = 0.684, \eta_p^2 = 0.004]$. The interaction between Assessment point and Group was also not significant $[F(1,39) = 0.41, p = 0.528, \eta_p^2 = 0.01]$. Therefore, neither age-related development nor musical training produced improvements in FDS score in relation to baseline performance.

Correlations Between Speech-in-Noise Perception and Phonological Short-Term Memory

The relationship between FDS scores and SRTs at T1 and T2 was assessed using two-tailed Pearson correlations. Results indicated significant covariance in only one of the listening conditions, namely when the SSN was collocated with the target speech at both T1 $(r = -0.35, p = 0.026)$ and T2 $(r = -0.45, p = 0.003)$. Correlations between FDS scores and SRTs in the other three conditions were non-significant (all $p > 0.07$).

DISCUSSION

Effect of Musical Training on Speech-in-Noise Perception

The primary aim of this study was to assess whether additional weekly musical instrument training provided over the course of 1 year improves speech perception under the sorts of challenging acoustic conditions children could realistically expect to experience in a classroom. Namely, environments in which energetic and informational maskers in various spatial relationships with the target speech would tax speech perception. However, there was no significant interaction between Assessment point and Group; that is, musical training was not associated with changes in SiN perception. Interactions that were predicted to show a musicians' advantage for SRTs under various masker and spatial manipulations were also not significant (Group × Assessment point × Masker type; Group × Assessment point × Spatial location). No other study to date has compared effects of musical training on SRTs in children using different masker types and target-masker spatial combinations in 5- to 7-year-old children. Therefore, in what follows, findings from previous cross-sectional and longitudinal studies which show parallels with the present study

but were conducted with children of various ages as well adults will be considered.

In a cross-sectional study by Strait et al. (2012), 7- to 13-year-old children with at least 4 years of musical training or no musical training were tested on different SiN perception tasks. Consistent with the present study's observations, the authors found no evidence for a musicians' advantage for speech perception in collocated babble or SSN. However, there was an advantage for musicians' speech perception when the SSN was spatially separated from the target speech. The masker and spatial conditions used in both studies had the potential to indicate whether musical training improves either peripheral auditory processing, cognition, or both. If the benefits of musical training were for peripheral auditory processing, speech perception under separated and energetic masker conditions would have been predicted because these conditions rely more on peripheral auditory processing than cognition. If benefits of musical training were cognitive, however, speech perception under the more cognitively demanding collocated and informational masker conditions would have been predicted in the musical-training group. In the case that both these processes were improved through musical training, both spatial and masker conditions would have shown improvement. As the cumulative findings of Strait et al. (2012) and the present study indicate no musicians' advantage for collocated conditions accompanied by informational maskers (i.e., babble noise or a single talker, respectively), a cognitive advantage of musical training cannot be concluded. Although Strait et al. (2012) found a musicians' advantage for speech perception under spatially separated energetic masker (i.e., SSN) conditions, the present study failed to demonstrate such trends longitudinally. Therefore, a benefit for musical training for peripheral auditory processing remains to be conclusively established.

A longitudinal musical-training study with children aged 6–9 years conducted by Slater et al. (2015) investigated whether musical training of up to 4 h per week over 2 years improves speech perception in collocated SSN compared to controls who received no musical training. After 1 year, the two groups did not perform significantly differently but a musicians' advantage was found after the second year of training. The discrepancy between this observation and the present study's findings might result from the considerable difference in the amount of the musical training provided in the two studies. However, cross-sectional studies with at least 4 years of musical training (Strait et al., 2012) and adults with over 10 years of musical training (Ruggles et al., 2014; Boebinger et al., 2015) reported no benefits for speech perception in collocated SSN for children either. Further longitudinal investigations are warranted to interpret these conflicting results.

Effect of Musical Training on Phonological Short-Term Memory

A secondary aim of this study was to test if musical training improved phonological short-term memory, which, in turn, could mediate improvements in SiN perception. At baseline, the musical-training group showed significantly higher FDS scores

and this advantage was maintained over time. Although groups were not equally matched at baseline, the ANOVA indicated whether the increase relative to baseline scores over time was greater in the musical training group than controls. The main effect of Assessment point indicated that FDS did not improve significantly over the course of 1 year across groups, and the non-significant interaction between Assessment point and Group meant that the relative increase in FDS was not higher in either group.

These findings contrast with results of Lee et al. (2007) who showed that 12-year-old children with an average of 6 years of musical training had better FDS than non-musicians, and results of Strait et al. (2012) who reported better auditory working memory in musically trained children aged 7 to 13 years. Strait et al. (2012) further reported that the correlation between the number of years of musical training received and auditory working memory ability was "marginally significant" ($r = 0.38$, $p = 0.08$), strongly implying that musical training was causally responsible for the measured between-group difference. Since the studies by Lee et al. (2007) and Strait et al. (2012) were cross-sectional, it cannot be excluded that these findings might be due to pre-existing between-group differences.

However, longitudinal evidence indicates that musically trained children's phonological short-term memory advantage, indicated by cross-sectional studies, are not necessarily due to pre-existing differences masquerading as training effects. A study by Roden et al. (2014) showed that 45 min of weekly musical training over 1 year in 7- to 8-year-old children significantly improved phonological short-term memory capacity. Somewhat surprisingly, the present study, even though methodological very similar (using also a longitudinal design, a comparable cognitive test, similarly aged participants, and a musical-training regimen of similar duration and frequency) failed to find evidence for a musical training-based cognitive improvement.

Correlations Between Speech-in-Noise Perception and Phonological Short-Term Memory

The strength of the relationships between phonological short-term memory and SiN perception was assessed using Pearson correlations between FDS scores and SRTs in the different masker and spatial conditions. Across groups, there was a significant moderate inverse correlation at T1 and T2 when the masker was collocated SSN. Similarly, Strait et al. (2012) found that auditory working memory correlated significantly with SiN perception in spatially separate SSN. Although spatial conditions differed, both studies found that the energetic masker used (i.e., SSN) covaried significantly with memory processes. This suggests that these cognitive skills are most useful when dealing with speech-perception challenges to the auditory periphery. However, it would be more intuitive to expect that cognitive skills should be useful when dealing with the more cognitively demanding maskers (i.e., informational maskers) and spatial conditions (i.e., collocated). Although, less obviously cognitively taxing conditions (e.g., spatially separated SSN maskers) could have a cognitive component for which stronger cognitive abilities could potentially provide benefits.

Limitations

Most prior studies investigating the musicians' advantage used a cross-sectional design, probably due to logistical and practical difficulties associated with the implementation of an actual musical-training intervention. For the present study, a longitudinal design was adopted so as to investigate possible causal relationships between the studied variables. To mimic a realistic context for a training program targeting typically developing young children, and also for logistic reasons, the musical training was delivered as part of the school curriculum. These choices imposed certain limits on the experimental design of the current study. First, the children were not randomly assigned to one of the two groups, limiting the causal claims that could be made by the present study. Their choice to attend the music-focused or sports-focused school determined their group membership. Hence, a bias in terms of participant characteristics (e.g., motivation, cognitive abilities) cannot be ruled out, even though all participants were normally performing pupils and the two groups did not differ in age or maternal socio-economic status. Second, the nature, amount and frequency of musical training was fixed by the curriculum in the music-focused school. It could be argued that other forms of or more musical training could have produced improvements in SiN perception and/or in phonological short-term memory capacity. However, it should be noted that studies using even less musical training have reported significant effects of musical training on cognitive abilities, such as improvements in phonological short-term memory after 45-min-long weekly training over 1 year (Roden et al., 2014) or in reading ability after 30-min-long weekly training for 8 months (Myant et al., 2008). Finally, although the present study considered some potential confounds (i.e., socio-economic status, hearing and language ability) that might have motivated children to take up musical training and might have led to pre-existing between-group inequalities, personality is an additional factor which has shown to be predictive of involvement in musical activities in adults and children (Corrigall et al., 2013; Swaminathan and Schellenberg, 2018). As personality was not measured, it was beyond the scope of this study to evaluate the extent to which this factor contributed to children's motivations to attend the respective schools, and thus represents a potential confound that should be controlled for in future studies.

CONCLUSION

This study assessed the impact of 1 year of musical instrument training on phonological short-term memory and SiN perception in children aged 5–7 years. Musical training improved neither phonological short-term memory, nor SiN perception in any of the listening conditions combining different maskers and spatial target-masker configurations that aimed to simulate realistic classroom conditions. This contrasts with previous studies in similarly aged children reporting evidence of musical-training benefits for SiN perception (Slater et al., 2015) and phonological short-term memory (Roden et al., 2014).

While our study adds to the list of investigations failing to find evidence for a musicians' advantage, more (especially longitudinal) research is warranted to investigate the nature, amount and frequency of musical training required for potential benefits in SiN perception and its underlying cognitive processes.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Faculty of Humanities Research Ethics Committee, University of Pretoria (GW20171130HS). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

AUTHOR CONTRIBUTIONS

DM designed the study, collected and analyzed the data, wrote the manuscript, and prepared the tables. CF assisted with revising the manuscript and responding to reviewers comments. RE collected the data and provided comments. JL and RL supervised the project and provided comments.

FUNDING

REFERENCES

Başkent, D., and Gaudrain, E. (2016). Musician advantage for speech-on-speech perception. *J. Acoust. Soc. Am.* 139, EL51–EL56. doi: 10.1121/1.4942628

Benz, S., Sellaro, R., Hommel, B., and Colzato, L. S. (2016). Music makes the world go round: the impact of musical training on non-musical cognitive functions – a review. *Front. Psychol.* 6:2023. doi: 10.3389/fpsyg.2015.02023

Bergman Nutley, S., Darki, F., and Klingberg, T. (2014). Music practice is associated with development of working memory during childhood and adolescence. *Front. Hum. Neurosci.* 7:926. doi: 10.3389/fnhum.2013.00926

Bidelman, G. M., Weiss, M. W., Moreno, S., and Alain, C. (2014). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *Eur. J. Neurosci.* 40, 2662–2673. doi: 10.1111/ejn.12627

Boebinger, D., Evans, S., Rosen, S., Lima, C. F., Manly, T., and Scott, S. K. (2015). Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.* 137, 378–387. doi: 10.1121/1.4904537

Bradley, J. S., and Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *J. Acoust. Soc. Am.* 123, 2078–2086. doi: 10.1121/1.2839285

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *J. Acoust. Soc. Am.* 109, 1101–1109. doi: 10.1121/1.1345696

Chan, A. S., Ho, Y. C., and Cheung, M. C. (1998). Music training improves verbal memory. *Nature* 396:128. doi: 10.1038/24075

Corbin, N. E., Bonino, A. Y., Buss, E., and Leibold, L. J. (2016). Development of open-set word recognition in children: speech-shaped noise and two-talker speech maskers. *Ear Hear.* 37, 55–63. doi: 10.1097/AUD.0000000000000201

Corrigall, K. A., Schellenberg, E. G., and Misura, N. M. (2013). Music training, cognition, and personality. *Front. Psychol.* 4:222. doi: 10.3389/fpsyg.2013.00222

Cowan, N., Nugent, L. D., Elliott, E. M., Ponomarev, I., and Saults, J. S. (1999). The role of attention in the development of short-term memory: age differences in the verbal span of apprehension. *Child Dev.* 70, 1082–1097. doi: 10.1111/1467-8624.00080

DiSarno, N. J., Schowalter, M., and Grassa, P. (2002). Classroom amplification to enhance student performance. *Teach. Except. Child.* 34, 20–25. doi: 10.1177/004005990203400603

Dockrell, J. E., and Shield, B. (2012). The impact of sound-field systems on learning and attention in elementary school classrooms. *J. Speech Lang. Hear. Res.* 55, 1163–1176. doi: 10.1044/1092-4388(2011/11-0026)

Dole, M., Hoen, M., and Meunier, F. (2012). Speech-in-noise perception deficit in adults with dyslexia: effects of background type and listening configuration. *Neuropsychologia* 50, 1543–1552. doi: 10.1016/j.neuropsychologia.2012.03.007

Fels, J., Buthmann, P., and Vorländer, M. (2004). Head-related transfer functions of children. *Acta Acust. United Acust.* 90, 918–927.

Fleming, D., Belleville, S., Peretz, I., West, G., and Zendel, B. R. (2019). The effects of short-term musical training on the neural processing of speech-in-noise in older adults. *Brain Cogn.* 136:103592. doi: 10.1016/j.bandc.2019.103592

Franklin, M. S., Sledge Moore, K., Yip, C. Y., Jonides, J., Rattray, K., and Moher, J. (2008). The effects of musical training on verbal memory. *Psychol. Music* 36, 353–365. doi: 10.1177/0305735607086044

Freyman, R. L., Helfer, K. S., McCall, D. D., and Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *J. Acoust. Soc. Am.* 106, 3578–3588. doi: 10.1121/1.428211

Fuller, C. D., Galvin, J. J. III, Maat, B., Free, R. H., and Başkent, D. (2014). The musician effect: does it persist under degraded pitch conditions of cochlear implant simulations? *Front. Neurosci.* 8:179. doi: 10.3389/fnins.2014.00179

Hall, J. W. III, Grose, J. H., Buss, E., and Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear Hear.* 23, 159–165. doi: 10.1097/00003446-200204000-00008

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *J. Acoust. Soc. Am.* 115, 833–843. doi: 10.1121/1.1639908

Johnstone, P. M., and Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *J. Acoust. Soc. Am.* 120, 2177–2189. doi: 10.1121/1.2225416

Klatte, M., Bergström, K., and Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Front. Psychol.* 4:578. doi: 10.3389/fpsyg.2013.00578

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., et al. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *J. Neurosci.* 34, 11913–11918. doi: 10.1523/JNEUROSCI.1881-14.2014

Lee, Y. S., Lu, M. J., and Ko, H. P. (2007). Effects of skill training on working memory capacity. *Learn. Instr.* 17, 336–344. doi: 10.1016/j.learninstruc.2007.02.010

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *J. Acoust. Soc. Am.* 117, 3091–3099. doi: 10.1121/1.1873913

Lorenzi, C., Dumont, A., and Füllgrabe, C. (2000). Use of temporal envelope cues by children with developmental dyslexia. *J. Speech Lang. Hear. Res.* 43, 1367–1379. doi: 10.1044/jslhr.4306.1367

MacCutcheon, D., Pausch, F., Fels, J., and Ljung, R. (2018). The effect of language, spatial factors, masker type and memory span on speech-in-noise thresholds in sequential bilingual children. *Scand. J. Psychol.* 59, 567–577. doi: 10.1111/sjop.12466

MacCutcheon, D., Pausch, F., Füllgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., et al. (2019). The contribution of individual differences in memory span and language ability to spatial release from masking in young

children. *J. Speech Lang. Hear. Res.* 62, 3741–3751. doi: 10.1044/2019_JSLHR-S-19-0012

McCreery, R. W., Walker, E., Spratford, M., Lewis, D., and Brennan, M. (2019). Auditory, cognitive, and linguistic factors predict speech recognition in adverse listening conditions for children with hearing loss. *Front. Neurosci.* 13:1093. doi: 10.3389/fnins.2019.01093

Myant, M., Armstrong, W., and Healy, N. (2008). Can music make a difference? A small scale longitudinal study into the effects of music instruction in nursery on later reading ability. *Educ. Child Psychol.* 25, 83–100.

Nelson, P., Kohnert, K., Sabur, S., and Shaw, D. (2005). Classroom noise and children learning through a second language. *Lang. Speech Hear. Serv. Sch.* 36, 219–229. doi: 10.1044/0161-1461(2005/022)

Neuman, A. C., Wroblewski, M., Hajicek, J., and Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear Hear.* 31, 336–344. doi: 10.1097/AUD.0b013e3181d3d514

Parbery-Clark, A., Skoe, E., Lam, C., and Kraus, N. (2009). Musician enhancement for speech-in-noise. *Ear Hear.* 30, 653–661. doi: 10.1097/AUD.0b013e3181b412e9

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* 2:142. doi: 10.3389/fpsyg.2011.00142

Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., and Di Domenico, A. (2019). Noise, age and gender effects on speech intelligibility and sentence comprehension for 11 to 13 years old children in real classrooms. *Front. Psychol.* 10:2166. doi: 10.3389/fpsyg.2019.02166

Purdy, S. C., Smart, J. L., Baily, M., and Sharma, M. (2009). Do children with reading delay benefit from the use of personal FM systems in the classroom? *Int. J. Audiol.* 48, 843–852. doi: 10.3109/14992020903140910

Renfrew, C. E. (1980). *Action Picture Test.* Headington: C. E. Renfrew.

Rhyner, P. M. (ed.) (2009). *Emergent Literacy and Language Development: Promoting Learning in Early Childhood.* New York, NY: Guilford Press.

Roden, I., Grube, D., Bongard, S., and Kreutz, G. (2014). Does music training enhance working memory performance? Findings from a quasi-experimental longitudinal study. *Psychol. Music* 42, 284–298. doi: 10.1177/0305735612471239

Rönnberg, J., Rudner, M., Foo, C., and Lunner, T. (2008). Cognition counts: a working memory system for ease of language understanding (ELU). *Int. J. Audiol.* 47(Suppl. 2), S171–S177. doi: 10.1080/14992020802301167

Ruggles, D. R., Freyman, R. L., and Oxenham, A. J. (2014). Influence of musical training on understanding voiced and whispered speech in noise. *PLoS One* 9:e86980. doi: 10.1371/journal.pone.0086980

Schröder, D. (2011). *Physically Based Real-time Auralization of Interactive Virtual Environments*, Vol. 11. Berlin: Logos Verlag.

Semel, E., Wiig, E. H., and Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals*, 4th Edn, San Antonio, TX: The Psychological Corporation.

Shield, B. M., and Dockrell, J. E. (2003). The effects of noise on children at school: a review. *Build. Acoust.* 10, 97–116. doi: 10.1260/135101003768965960

Shield, B. M., and Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *J. Acoust. Soc. Am.* 123, 133–144. doi: 10.1121/1.2812596

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., and Kraus, N. (2015). Music training improves speech-in-noise perception: longitudinal evidence

from a community-based music program. *Behav. Brain Res.* 291, 244–252. doi: 10.1016/j.bbr.2015.05.026

Stone, M. A., Füllgrabe, C., and Moore, B. C. (2012). Notionally steady background noise acts primarily as a modulation masker of speech. *J. Acoust. Soc. Am.* 132, 317–326. doi: 10.1121/1.4725766

Strait, D. L., O'Connell, S., Parbery-Clark, A., and Kraus, N. (2013). Musicians' enhanced neural differentiation of speech sounds arises early in life: developmental evidence from ages 3 to 30. *Cereb. Cortex* 24, 2512–2521. doi: 10.1093/cercor/bht103

Strait, D. L., Parbery-Clark, A., Hittner, E., and Kraus, N. (2012). Musical training during early childhood enhances the neural encoding of speech in noise. *Brain Lang.* 123, 191–201. doi: 10.1016/j.bandl.2012.09.001

Swaminathan, J., Mason, C. R., Streeter, T. M., Best, V., Kidd, G. Jr., and Patel, A. D. (2015). Musical training, individual differences and the cocktail party problem. *Sci. Rep.* 5:11628. doi: 10.1038/srep11628

Swaminathan, S., and Schellenberg, E. G. (2018). Musical competence is predicted by music training, cognitive abilities, and personality. *Sci. Rep.* 8:9223. doi: 10.1038/s41598-018-27571-2

Swanepoel, D. W., Myburgh, H. C., Howe, D. M., Mahomed, F., and Eikelboom, R. H. (2014). Smartphone hearing screening with integrated quality control and data management. *Int. J. Audiol.* 53, 841–849. doi: 10.3109/14992027.2014.920965

Talamini, F., Altoè, G., Carretti, B., and Grassi, M. (2017). Musicians have better memory than nonmusicians: a meta-analysis. *PLoS One* 12:e0186773. doi: 10.1371/journal.pone.0186773

Tallal, P. (2014). "Experimental studies of language learning impairments: from research to remediation," in *Speech and Language Impairments*, eds D. V. M. Bishop, and L. Leonard, (Hove: Psychology Press).

Vickers, D., Degun, A., Canas, A., Stainsby, T., and Vanpoucke, F. (2016). Deactivating cochlear implant electrodes based on pitch information for users of the ACE strategy. *Adv. Exp. Med. Biol.* 894, 115–123. doi: 10.1007/978-3-319-25474-6_13

Wightman, F., Kistler, D., and Brungart, D. (2006). Informational masking of speech in children: auditory-visual integration. *J. Acoust. Soc. Am.* 119, 3940–3949. doi: 10.1121/1.2195121

Wightman, F. L., and Kistler, D. J. (2005). Informational masking of speech in children: effects of ipsilateral and contralateral distracters. *J. Acoust. Soc. Am.* 118, 3164–3176. doi: 10.1121/1.2082567

Zendel, B. R., West, G., Belleville, S., and Peretz, I. (2019). Music training improves the ability to understand speech-in-noise in older adults. *Neurobiol. Aging* 81, 102–115. doi: 10.1016/j.neurobiolaging.2019.05.015

Check for updates

Executive Functions, Pragmatic Skills, and Mental Health in Children With Congenital Cytomegalovirus (CMV) Infection With Cochlear Implants: A Pilot Study

Ulrika Löfkvist [1,2]*, Lena Anmyr [2,3], Cecilia Henricson [4] and Eva Karltorp [2,3]

[1] Department of Special Needs Education, University of Oslo, Oslo, Norway, [2] Department of Clinical Science, Intervention and Technology, Karolinska Institutet, Solna, Sweden, [3] Karolinska University Hospital, Stockholm, Sweden, [4] Department of Behavioural Sciences and Learning, Linköping University, Linköping, Sweden

Congenital cytomegalovirus (cCMV) infection is the most common cause of progressive hearing impairment. In our previous study around 90% of children with a cCMV infection and CI had severely damaged balance functions (Karltorp et al., 2014). Around 20% had vision impairment, 15% were diagnosed with Autism-Spectrum-Disorder, and 20% with ADHD. One clinical observation was that children with cCMV infection had problems with executive functioning (EF), while controls with a genetic cause of deafness (Connexin 26 mutations; Cx26) did not have similar difficulties. A follow-up study was therefore initiated with the main objective to examine EF and pragmatic skills in relation to mental health in children with a cCMV infection and to draw a comparison with matched controls with Cx26 mutations (age, sex, hearing, non-verbal cognitive ability, vocabulary, and socioeconomic status level). Ten children with a cCMV infection and CI (4.8–12:9 years) and seven children with CI (4:8–12:8 years) participated in the study, which had a multidisciplinary approach. Executive functioning was assessed both with formal tests targeting working memory and attention, parent and teacher questionnaires, and a systematic observation by a blinded psychologist during one test situation. Pragmatics and mental health were investigated with parent and teacher reports. In addition, the early language outcome was considered in non-parametric correlation analyses examining the possible relationships between later EF skills, pragmatics, and mental health. Children with cCMV had a statistically significant worse pragmatic outcome and phonological working memory than controls despite their groups having similar non-verbal cognitive ability and vocabulary. However, there were no statistical differences between the groups regarding their EF skills in everyday settings and mental health. There were associations between early language outcomes and later EF skills and pragmatics in the whole sample.

Conclusion: Children with a cCMV infection are at risk of developing learning difficulties in school due to difficulties with phonological working memory and pragmatic skills in social interactions.

Keywords: executive functions, pragmatcis, mental health, cytomegalo virus infection, cochlear implant

INTRODUCTION

This explorative follow-up study is part of a larger research program with the objective to investigate the effects of different etiological backgrounds in children with pediatric deafness. We have investigated the effects of congenital cytomegalovirus (cCMV) infection in a sample of deaf children with cochlear implants (CI), and results have been related to their executive functioning, pragmatic skills, mental health, and possible interactions with the participants' early language outcome. This has been done in a group of children with CI, deafened due to cCMV infection, and in hearing-matched controls with a genetic cause of deafness: Connexin 26 mutations (Cx26). Congenital CMV infection is known to be related to comorbid conditions, while Cx26 is usually not related to other issues or diagnoses.

Executive functions (EF) are connected to frontal lobe capacity (Kave et al., 2008) and represent underlying, interrelated processing skills, such as working memory, attention, and inhibition/flexibility, which all are important for several functions like communication, social cognition, and learning (Miyake et al., 2000; Diamond, 2013). Children with CI form a heterogeneous population with considerable variation, especially in EF (Figueras et al., 2008; Beer et al., 2014; Kronenberger et al., 2014) but also in the spoken language outcome (Boons et al., 2012; Löfkvist et al., 2012; Walker et al., 2019) and mental health (Hintermair, 2007; Anmyr et al., 2012; Lingås-Haukedal et al., 2018). Poor EF in children with CI may negatively influence pragmatic skills, and especially in subgroups with known comorbid conditions like children with a cCMV infection. Poor attention skills and inferior ability to interpret and use pragmatic cues could affect emotional responses and behavioral actions in social interactions. In turn this might affect personal relationships and mental health. It is therefore valuable to explore the complex relationship between EF, pragmatic skills, and mental health.

Some of the language variation in the population of children with CI may be explained by age at implantation (Dettman et al., 2007; Colletti et al., 2011), non-verbal cognitive ability (Geers et al., 2011), parental sensitivity (Quittner et al., 2013), and socio-economic status (Szagun and Stumper, 2012). Phonological working memory is one EF ability that has previously been associated with language outcome (Gathercole et al., 2008; Wass, 2009), language learning (Willstedt-Svensson et al., 2004), and social interaction (Lyxell et al., 2008). Better language abilities may have positive effects on mental health (Lingås-Haukedal et al., 2018).

The cause of deafness and comorbidity could contribute to explaining some of the still unknown variations in cognitive processing, including poorer EF, which can negatively influence pragmatic skills and/or mental health in preschool and school-aged children with CI. Goberis et al. (2012) investigated pragmatic skills in children aged 3–7 years ($n = 126$) with hearing impairment (HI) and in controls with typical hearing (TH) ($n = 109$). They found that children with HI acquired pragmatic skills at a slower pace than controls with TH, even with targeted intervention strategies (Goberis et al., 2012). Goberis et al. (2012)

did not investigate the possible effects of the cause of deafness in their study cohort.

Half of all sensorineural deafness (50%) is explained by genetic reasons (70% non-syndromic and 30% syndromic) (Alford et al., 2014). The most common non-syndromic genetic causes of deafness are Cx26 mutations (*GJB2*); they are manifested as congenital uni- or bilateral hearing loss/deafness, which can also be progressive. The other 50% of sensorineural deafness is acquired before birth, in infancy, or in early childhood and is explained by non-genetic causes like virus-infections, meningitis, or toxicity during pregnancy (Alford et al., 2014). Congenital cytomegalovirus (cCMV) infection is the most common cause within this group of congenital or early acquired hearing loss/deafness (Grosse et al., 2008). The heterogeneity and incidence of comorbid deficits or diagnoses are high in children with a cCMV infection compared to children with Cx26 deafness who usually do not have other additional diagnoses or deficits related to their cause of deafness (Karltorp et al., 2014).

Congenital CMV infection has a birth rate of 5% per 1,000 births. It has previously been suggested that 80% of all infants who are infected with a CMV infection *in utero* will develop typically without persisting deficits or difficulties within the area of perception, cognition (including language), or motor skills (Boppana et al., 2005). Around 15% of all children with a cCMV infection are diagnosed with a sensorineural HI (Boppana et al., 2005; Grosse et al., 2008). However, some of the children who are born with a cCMV infection will experience a late onset of their hearing loss and will thus not be identified through the universal newborn hearing screening (UNHS) system. Up to 40% of all infected children with cCMV will pass the Oto-Acoustic-Emission (OAE) test at the time when they are born (Fowler et al., 2017). Instead, they will experience later detection and diagnosis of their unilateral or bilateral HI, which may also be progressive (Fowler et al., 2017). So far, it is only the public health care system in the province of Ontario, Canada, that has decided to implement a general cCMV screening, as part of their existing UNHS system, for all newborns [(https://www.newbornscreening.on. ca/en/page/congenital-cytomegalovirus), retrieved 2019-11-24]. Several countries and states in the USA have started to screen for cCMV in all infants who are identified with a hearing loss through the UNHS. Aside from identifying a minority of all children with cCMV infection and HI, the UNHS system will only target infants with cCMV infection who have an HI and not the ones with initially TH but who might have other deficits and clinical symptoms. In the literature, there are reports of children who do not have HI but have vision impairments, motor skills deficits, balance problems, and/or cognitive deficits and, in some cases, neurodevelopmental diagnoses like mental retardation, cerebral palsy (CP), autism spectrum disorder (ASD), or ADHD, and this includes a negative impact on quality of life (Malm and Engman, 2007; Korndewal et al., 2017).

There are several studies that have investigated spoken language in relation to EF abilities like phonological working memory in children with CI (Lyxell et al., 2008; Beer et al., 2014; Kronenberger et al., 2014). Only a few studies have examined more general language abilities like the development of sentence

understanding and speech intelligibility in children with a cCMV infection who use CI (Ramirez Inscoe and Nikolopoulos, 2004; Yoshida et al., 2009). Yoshida et al. (2009) found that language understanding developed at a slower pace in children with cCMV infection ($n = 4$) compared to children with CI who were deafened due to other causes. Ramirez Inscoe and Nikolopoulos (2004) showed in their study that there was a large variability concerning the speech intelligibility level in their cohort of 16 children with cCMV.

We have previously reported that some children with cCMV can catch up and develop adequate speech and language abilities over time, while others may have comorbid conditions (Karltorp et al., 2014). In the study by Karltorp et al. (2014), we found that children with cCMV, including those with typical language test results, had poorer impulse control and attention span during the language and hearing assessment procedure compared to controls with Cx26. On this test occasion we had no formal evaluation of EF, pragmatic skills, or mental health, and there was no psychologist involved in the research team (Karltorp et al., 2014). This unexpected finding was the first indication for us that EF, in particular, could be more difficult for children with cCMV.

Children with profound HI who use CI have been reported to have mental health issues more frequently than peers with TH (Hintermair, 2007). Nevertheless, recent findings in a Norwegian study displayed that the mental health of children with CI, aged 5; 0–12; 11 years, was similar to age-matched children with TH (Lingås-Haukedal et al., 2018). Lingås-Haukedal et al. (2018) examined health-related quality of life in 186 children with CI, as reported by parents, and they found that about 50% of the children with CI had levels comparable to peers with TH ($n = 80$). The possible influence on mental health in relation to the cause of deafness was not examined in the study by Lingås-Haukedal et al. (2018).

Mental health can be assessed with the Strengths and Difficulties Questionnaire (SDQ), which was originally developed in nearly identical versions for parents and teachers of children aged 4–16 (Goodman, 1997; Goodman et al., 1998). The SDQ can be used as part of a clinical assessment, as a treatment–outcome measure, and as a research tool (Goodman et al., 2000). The SDQ has been found to be a reliable and valid questionnaire for use in samples of deaf/hard of hearing children (Cornes, 2007; Hintermair, 2007). In a study by Hintermair (2013) EF and mental health were evaluated in children with HI, and in relation to their social communication skills by using two questionnaires—Behavior Rating Inventory Executive Function (BRIEF) (EF abilities; Gioia et al., 2000) and SDQ (mental health)—together with a communicative competence scale (Hintermair, 2013). The questionnaires were rated by teachers of 214 children with HI, who had a mean age of 12;4 years, and results were compared to normative data of 720 children. There was a statistically significant higher rate of EF difficulties in all children with HI compared to the norm data. Children who attended mainstream schools were rated to have better communicative competence than children who attended special schools for deaf children. A regression analysis revealed that better executive functioning and communicative competence in children with HI was associated with a lower incidence of behavioral problems (Hintermair, 2013). Seemingly,

difficulties in verbal language abilities were not only related to EF outcome but also to social behavior in children with HI (Hintermair, 2013). Worse EF may have an influence on literacy, prosody, and language abilities (Lyxell et al., 2009) and may also negatively affect pragmatic skills in children with HI (Goberis et al., 2012; Hintermair, 2013), especially for children with initially atypical brain patterns in early childhood (Kave et al., 2008; Korndewal et al., 2017). Poor phonological working memory and short attention span are, for instance, known to affect children's ability to understand instructions and to retrieve words from their long-term memory (Lyxell et al., 2009). These difficulties can be negatively associated with linguistic and social skills in verbal interactions, in particular with regard to interactions in noisy environments, such as classrooms or playgrounds. To our knowledge, there are no previous studies in the literature that have explored cognitive abilities, like EF, pragmatic skills, as well as mental ill health in children with cCMV who use CI.

The objective of the present study was to explore EF, pragmatic skills, and mental ill health in children with an acquired deafness (cCMV infection) using CI and who have no known additional diagnoses like ADHD, Developmental Language Disorder (DLD), or Autism-Spectrum-Disorder (ASD) and compare this to well-matched controls who were deafened due to a genetic non-syndromic deafness (Cx26 mutations). The groups were matched on the basis of age, hearing, vocabulary, parents' education level, and non-verbal cognitive ability.

Several research questions were addressed:

1. Do children with CI have *worse EF* results in relation to norm data, regardless of the cause of deafness, and do children with cCMV infection have even poorer executive functioning compared to children with genetic non-syndromic deafness (Cx26)?
   The hypothesis was that all children with CI would have a worse EF outcome than children with TH (norm data) (Kronenberger et al., 2014), and that children with a cCMV infection would have even poorer EF results than children with Cx26 mutations. The reason for this hypothesis was that a congenital virus infection may be related to additional diagnoses, atypical brain patterns, and virus-related deficits (Karltorp et al., 2014; Korndewal et al., 2017).

2. Do children with a cCMV infection who use CI have *worse pragmatic skills and mental health* than well-matched children with Cx26 in comparison to norm data?
   The hypothesis was that children with a cCMV infection would have worse pragmatic skills than controls. We hypothesized that worse pragmatic skills in children with cCMV infection may be explained not only by their deafness but also by the consequences of their congenital virus infection (Korndewal et al., 2017).

3. Is there a relationship between EF, pragmatic skills, mental health, and early language abilities in children with CI, regardless of the cause of deafness?
   The hypothesis was that there would be a relationship among EF, pragmatic skills, and mental health in all children with CI, regardless of their cause of deafness and that better speech and language understanding in early childhood could be related to an improved later outcome (Goberis et al., 2012; Hintermair, 2013).

METHODS

The current follow-up study had a long-term approach, which included data collection and retrospective reviews of medical journals, and it is part of a larger research study program at the Auditory Implant Center, Karolinska Institutet, aiming to explore the effects of etiological factors in children with CI, who have different causes of deafness, in relation to their listening skills, cognitive abilities, mental health, and linguistic outcome. This study was carried out in accordance with the recommendations of the Regional Ethical Review Board in Stockholm, Sweden. All participants were first provided with written information about the study. Written informed consent was then obtained from the parents of all participants, in accordance with the Declaration of Helsinki. The protocol was approved by the Regional Ethical Review Board in Stockholm, Sweden; DN 2012:/2.

Participants

Inclusion criteria: children with cCMV or Cx26 who were older than 4 years and younger than 13 years at the time of the study, who used their CI during all waking hours, who did not have a confirmed and known additional diagnose(s) related to deficits in the domain of executive functioning (ASD, ADHD) or pragmatic skills (DLD, ASD), and who had at least one parent who spoke Swedish at home. Families with a child who fulfilled the criteria and who had been implanted at the Auditory Implant Center, Karolinska University Hospital, which covers half of the Swedish population (i.e., five million people), were invited to take part in the follow-up study. Parents were first provided with written information about the study and then, if interested, they were asked to sign an informed consent of participation form. Children who could read (older than 8 years) also signed a consent of participation. Seven children with cCMV were excluded because they were too young or too old, and two children with cCMV were excluded because they had several additional diagnoses aside from their deafness. There was one participant with Cx26 who fulfilled the criteria and who initially agreed to participate but later decided not to participate in the study.

The final study sample consisted of 17 children ($N = 17$) aged 4.8–12.9 years (mean 8.2; Md 7; 8 years)—eight girls and nine boys with a confirmed cCMV infection or Cx26 mutations with CI who met the inclusion criteria. There were no statistically significant differences between the groups (cCMV vs. Cx26) regarding age (Z = −0.05, $p = 0.96$, $r = 0.01$), sex [χ^2 (1, $n =$ 17) = 1.63, $p = 0.34$] or parent educational level for mothers (Z = −0.40, $p = 0.69$, $r = 0.10$) or fathers (Z = −1.53, $p = 0.13$, $r =$ 0.37). All parents had at least a high school or a university degree, which is common to most parents in the Swedish population. The participants came from different parts of Sweden, and a majority came from the Stockholm area. All but one family had been offered some kind of Family-Centered Early-Intervention (FCEI) option. Nine families had received regular services (once a week or every second week) from a speech–language pathologist or a teacher of the deaf at their local habilitation team using an auditory verbal approach for at least 1 year after the first CI surgery. Seven families had received similar intervention options but less frequently. One child, who was identified late with severe-to-profound hearing loss, had not received FCEI services with focus on parent engagement and spoken language skills before or after the first CI surgery (**Table 1**).

Children With cCMV

All children ($n = 10$), six girls and four boys, had been screened at birth with OAE, and five children passed the first hearing screening without remarks. All children with cCMV were tested with an MR investigation before the CI intervention, and 100% ($n = 10$) had results that indicated slightly atypical patterns (white substance), mainly in the frontal regions of the brain (level 1 of three levels, where a higher level indicates more injuries) (Karltorp et al., 2014). The parents reported that there were no close family members with ADHD, ASD, or DLD. Some of the children with cCMV had been introduced and exposed to sign language or supported signs in daycare settings and in their home environment in early childhood. At the time of the follow-up study, however, only a few of them used signs themselves, and the majority of children went to mainstream preschools or schools. One child went to a special school for those with hearing impairment that had an adjusted listening environment, smaller class size, and spoken Swedish as the educational language. No child with a cCMV infection attended a deaf school. The nine children who attended mainstream schools had a certain degree of an adjusted listening environment in their mainstream classrooms. They were included in typical classes, with more pupils than in special schools, and there was a large variety in the type and level of support available for the individual child and their family.

Children With Cx26 Mutations

All children ($n = 7$), two girls and five boys, had been screened at birth with OAE. According to the parents, none of the families had close family members with ADHD, DLP, or ASD. All children communicated primarily with spoken language at home and in preschool/school. A few children knew and used sign-supported language or sign language. All but one had been going to mainstream daycare since they were toddlers, and they continuously went to mainstream preschools/schools, close to their homes, at the time of the study. One child with Cx26 attended a special school for hearing-impaired children. The rest of the group of children with Cx26 had a similar situation compared to children with cCMV who were mainstreamed (typical class sizes, some adjustment of the listening environment, a large variation in the type and level of individual child support in their preschool/school).

Procedure

All participating families had visited the same Auditory Implant Center at Karolinska University Hospital since their child received their first CI. Families were scheduled for a duration of around 4 h at the follow-up occasion (see **Table 2**). The team at the Auditory Implant Center had previously assessed the child both pre- and post-implantation with a fixed test protocol and with the same test procedures. The participating children were randomly scheduled to meet a multidisciplinary team containing

TABLE 1 | Participant demographics concerning ages (months) when individual children were identified with a hearing impairment (HI), ages at identification of cause of deafness (cCMV or Cx26), ages when the children received their 1st and 2nd CI; type of Family-Centered Early Intervention (FCEI) actions after identification of HI, and the chronological ages of the children at the follow-up study.

Child	Age at HI id.	Age at id. of etiology	Age at 1st CI	Age at 2nd CI	FCEI	Age at follow-up
CMV-1	18	24	22	22	1	128 (10.7 year)
CMV-2	13	24	17	72	1	82 (6.8 year)
CMV-3	36	44	44	44	2	130 (11.0 year)
CMV-4	0	10	12	12	1	106 (8.2 year)
CMV-5	0	0	10	16	1	63 (5.3 year)
CMV-6	6	16	17	39	2	67 (5.6 year)
CMV-7	12	20	21	24	1	118 (9.8 year)
CMV-8	0	18	18	30	2	155 (12.9 year)
CMV-9	0	9	9	9	2	57 (4.8 year)
CMV-10	30	64	67	*	3	81 (6.8 year)
Md (min-max)	9 (0–36)	19 (0–64)	18 (9–67)	24 (9–72)	2 (1–3)	99 (57–155)
Cx26-11	2	10	8	49	1	108 (9.0 year)
Cx26-12	2	52	48	*	2	70 (6.0 year)
Cx26-13	10	19	19	23	2	153 (12.8 year)
Cx26-14	1	9	9	9	1	57 (4.8 year)
Cx26-15	19	23	22	27	2	140 (11.8 year)
Cx26-16	0	0	14	*	2	64 (5.3 year)
Cx26-17	2	22	34	95	2	93 (7.8 year)
Md (min-max)	2 (0–19)	19 (0–52)	19 (8–48)	27 (9–95)	2 (1–2)	99 (57–155)

*Age at HI id., age when the child was identified with hearing impairment; Age at id. of etiology, age when the child's cause of deafness was identified; *Three children (CMV-10, Cx26-12, Cx26-16) had bimodal hearing (CI+HA); type of FCEI, family-centered intervention actions during the 1st year after 1st hearing aid fitting, with focus on individual parent guidance and with an auditory-verbal approach; 1 = Yes, on a regular basis; 2 = Yes, but not on a regular basis; 3 = No FCEI offered.*

experienced clinicians/researchers: a medical doctor, speech-language pathologist, audiologist, social worker. In addition, a blinded psychologist who had no previous knowledge about the individual children and who did not know which group each participant belonged to (cCMV or Cx26) was included as a team member to perform the EF tests and behavior observations (Karltorp et al., 2014). Before the visit, parents and teachers had already filled out questionnaires that measured executive functions, pragmatics, and the mental health of the child. After the test occasion, the blinded psychologist, for validity reasons, observed the recorded video-based material from the test occasion. X-ray data (MR) and other child-related information regarding early clinical findings were retrieved from the individual children's medical records and were then reviewed in the data collection process by a medical doctor who was part of the multidisciplinary research team. The medical doctor met all families at the follow-up occasion for a short interview with the parents about their early FCEI services, family background (hereditary for ADHD and ASD etc.) and the child's medical health.

Measures
Executive Functions—Tests, Questionnaire, and Qualitative Analyses of Behavior
Everyday attention level
Everyday attention level was assessed with the Test of Everyday Attention for Children (TEA-Ch) in children older than 6 years (Heaton et al., 2001; Manly et al., 2001). The TEA-Ch test

has previously been translated to Swedish and used for other clinical groups, such as 7-years-old children with low birth weight (Starnberg et al., 2018), but there are so far no Swedish norms on the test. The TEA-Ch is a test that assesses everyday attention capacity and is presented both in an auditory or visual modality. The TEA-Ch consists of nine subtests; Sky Search, Score!, Creature Count, Sky Search DT, Map Mission, Walk-don't walk, Opposite Worlds, and Code transmission. These subtests assessed the participant's ability to sustain, select, and shift their attention (Manly et al., 2001). In the present study, the test procedure was conducted as suggested in the manual. The subtest Score Dual Task (to discriminate between two sound tracks only by listening) was excluded in the present study, for reliability reasons because it was too difficult to perform for participants with CI.

Phonological working memory
Phonological working memory (a non-word repetition task that is a relatively pure measure of the phonological loop capacity, Baddeley, 2012) and General working memory (the capacity to simultaneously store and process information, Wass, 2009) was assessed by using two subtests—Serial Recall on non-words and Sentence, Completion and recall, respectively—from the SIPS test battery (Wass, 2009) in children older than 5 years. In the Serial Recall on non-words subtest, children listened to standardized and recorded non-word material that was presented from loudspeakers and with gradually increasing numbers of non-words in a row. The children decided on the comfortable

TABLE 2 | Description of assessment tools used at different test occasions in the study.

Assessment tools study	Preop	1 year with CI	3 years with CI	Follow-up
TESTS				
Reynell-III (Language understanding)	X	X	X	
Expressive grammar scale	X	X	X	
SIR-2 (Speech intelligibility)	X	X	X	
BNT (expressive vocabulary)				X
Lexical-semantic error analysis (BNT)				X
FAS and Animal (word fluency ability)				X
Ravens (non-verbal cognitive ability)				X
TEA-Ch (attention level)				X
EBA-R (observational analysis scale)				X
SIPS; phonological working memory				X
SIPS; general working memory				X
Speech recognition (silence)				X
Speech recognition (noise)				X
QUESTIONNAIRES				
BRIEF (parents) (EF skills)				X
BRIEF (teachers) (EF skills)				X
CCC-2 (parents) (pragmatics)				X
SDQ (parent) (mental health)				X
SDQ (teachers) (mental health)				X

hearing level before the assessment. Then, participants were asked to repeat the non-word utterances as accurately as they could. The percentage of correctly reproduced consonants in the whole test was calculated. In the Sentence, Completion and recall subtest, the number of correctly recalled real words were counted. Examples of sentences were, "The sky is blue and the grass is… (green) (participant fills in)" and "You sit on a chair, and you sleep in a… (bed) (participant fills in)." Then, the test administer asks the participant, "Which words did you say?" These two cognitive tests have been used in children with TH and typical development and in clinical groups from around 6 years of age (Wass et al., 2008; Lyxell et al., 2009; Henricson et al., 2012).

Executive functioning in the home and a preschool/school environment

Executive functioning in the home and a preschool/school environment was rated in a questionnaire by parents and the child's primary teacher, respectively, who filled in the Behavior Rating Inventory of Executive Function (BRIEF) to evaluate possible behavioral problems concerning EF in everyday settings (at home and in preschool/school, respectively) (Gioia et al., 2000; Isquith et al., 2004). BRIEF functional scales were used to screen for possible behavioral problems in executive functioning in everyday life situations. The individual subscale results of BRIEF can be summarized in three different functional scales: Behavior Rating Inventory (BRI), Metacognition Index (MI), and the Global Executive Composite (GEC). Caregivers of all children filled out the questionnaire, as did the child's teacher (preschool or school). The BRIEF questionnaire has been translated to

Swedish, but there is yet no validation of the test or Swedish norms available. We therefore compared the study results with the American norms. T-scores ≥65 were considered clinically significant, and any scores ≥70 were extremely high (Gioia et al., 2000; Isquith et al., 2004).

Emotional, behavioral, and attention rating (EBA-R)

Emotional, Behavioral, and Attention Rating (EBA-R), an in-house developed observational and qualitative analysis scale (Henricson and Löfkvist, **Appendix 1**), was used to evaluate the child's behavior during the test session with the psychologist (TEA-Ch). It was conducted by the blinded psychologist who also reviewed videotapes afterwards to confirm or adjust the initial observational rating results. Several categories were rated: Expression of positive emotions; Frustration level; Restlessness level; Focus level; Problem solving (structured ability, logical behavior); and Problem solving (unstructured ability, chaotic behavior) (see **Appendix 1**).

Pragmatic Skills

The second edition of the Swedish version of the parent report questionnaire Child Communication Checklist (CCC-2) was used to examine the children's pragmatic skills (Bishop, 2003). This assessment tool includes Swedish norms for children between 4–16 years (https://www.pearsonassessment.se/ccc-2). The checklist, which had 70 different statements, was filled in by parents and then analyzed afterwards with computerized scoring. The CCC-2 consists of 10 subscales; A–Speech; B–Syntax; C–Semantics; D–Coherence; E–Inappropriate initiations; F– Stereotypic language; G–Use of context; H–Non-verbal communication; I–Social relations; and J–Interests.

Mental Health

The SDQ is a 25-item screening questionnaire. Each item is rated 0 = not true, 1 = somewhat true, or 2 = certainly true (Goodman, 1997; Malmberg et al., 2003), in which 10 items reflect strengths, 14 reflect difficulties, and 1 is neutral but is scored as a difficulty item on the peer problems subscale (Goodman, 1997). A small number of negatively worded items are reverse scored. The items are grouped in five subscales containing five items each. The subscales are emotional symptoms, conduct problems, hyperactivity-inattention, peer problems, and prosocial behavior. Each subscale score ranges from 0 to 10. Higher scores on the prosocial behavior subscale reflect strengths, whereas higher scores on the other four subscales reflect difficulties. A total difficulty score is calculated by adding the sum of scores on the emotional, conduct, hyperactivity, and peer problems subscales, with a possible range of 0 to 40 (Goodman, 1997). The construction of cut-off values is based on normative SDQ scoring, as proposed by Goodman (1997). A total of 10% of a norm sample with the highest scores were classified as abnormal, the next 10% as borderline, and the remaining 80% as normal. These cut-offs varied between informant versions as well as across subscales and the total difficulties scale (Goodman, 1997, 2005). The psychometric properties of the Swedish parent-rated version of the SDQ have been evaluated by Smedje et al. (1999).

Speech Recognition

Sound field hearing thresholds were assessed by presenting frequency-modulated tones at octave frequencies from 0.125–6 kHz. The hearing tests were conducted using best-aided conditions (bilateral CI or in bimodal fashion; CI and HA) for speech in silence and in multi-source noise (Asp et al., 2012). The speech recognition in quiet was conducted with a 25-item list of monosyllabic words presented at 65 dB SPL level. The noisy conditions consisted of a presentation of stationary speech-shaped noise from ±45° to ±135° azimuth (uncorrelated signals), which resulted in a signal-to-noise ratio of 0 dB.

Screening of Non-verbal Cognitive Ability

All children were assessed with the Raven colored progressive matrices (Raven et al., 2003). This test evaluates an individual's ability to discover and interpret visual patterns and can be viewed as a screening tool for IQ. There are, so far, no Swedish norms on Ravens, and we therefore used the validated and standardized English norms for comparisons between participants and children with TH (Raven et al., 2003).

Language Abilities

Children were assessed by way of expressive vocabulary/picture naming by using a validated Swedish version of the Boston Naming Test (BNT) (Kaplan et al., 1983; Tallberg, 2005). The BNT has been normed for Swedish children aged 6–15 years (N = 152) (Brusewitz and Tallberg, 2010). The Boston Naming Test is an open-set test that consists of 60 pictures that the child is asked to name. In the current study, we did not allow phonological or semantic prompting. Synonyms and subordinated words were counted as correct words. A lexical-semantic error analysis was performed with the purpose of exploring more in-depth semantic knowledge of incorrect responses besides form scoring the number of correct responses on the BNT (Löfkvist et al., 2014). Word fluency tasks included Animal word fluency (semantically based) and FAS letter word fluency, a phonemically based task that not only measures word retrieval from the long-term memory but also targets EF indirectly, considering the individual's use of strategies in the process of retrieving words from their long-term memory. Both these two tests have been normed in Swedish children aged 6–15 years (N = 130) (Tallberg et al., 2011).

Early Language Abilities

The Reynell-III test evaluates expressive and receptive language abilities and was originally developed for children aged 0–7 years with TH (Edwards et al., 1997). A validated Swedish norm study of the *receptive test* part was conducted in a group of Swedish children with TH and typical development (Eriksson and Grundström, 2000). The results showed that children aged 2:6–3:5 years (N = 122) had comparable results with age-matched English children (Edwards et al., 1997). The Swedish norm data has a narrow age range. As the English and Swedish norm data showed similar results, it was therefore decided to use the English validated norms as comparisons with clinical data in the present study. Reynell-III was used to measure language understanding pre-op as well as after one year and three years

post-op as part of the regular follow-up procedure for all children who have been implanted with CI at the Auditory Implant Center, Karolinska University Hospital, including the current study sample (Edwards et al., 1997).

Furthermore, experienced speech–language pathologists who were the same clinicians who performed the Reynell-III assessment pre-op, and after 1 and 3 years after the first CI, also rated the level of *expressive grammar* (level 1–8) and the child's level of *speech intelligibility* (see **Appendix 2**) (Allen et al., 2001; Löfkvist, 2014). The expressive grammar-rating scales (level 1: "no use of voice with intent" to level 8: "typical or correct expressive grammar and sentence level") were developed within a Swedish context, primarily for use in children with HI, but may be used in other groups, including children with TH (see **Appendix 2**, Löfkvist, 2014).

The Speech Intelligibility Rating Scales (SIR-2) was specifically developed for use in children with HI and consists of a 5-level rating scale from "recognizable words in speech" to "connected speech is intelligible to all listeners" (Allen et al., 2001). The reliability of the SIR was originally tested and validated in 54 English children with CI, aged 1; 2–10 years. Experienced speech–language pathologists at the Auditory Implant Center at Karolinska University Hospital rated the SIR-2 before the first cochlear implantation and, thereafter, every 12 months until the child reached level 5. The SIR has been translated into Swedish and implemented at the Auditory Implant Center, but has not yet been validated in the Swedish context.

Statistical Analyses

Potential group differences (cCMV infection vs. Cx26) were examined with Mann Whitney U-tests that included effect size indicators; $r = Z/\sqrt{N}$ and a Chi-square test, and Spearman's correlations were used to examine the possible relationships between executive function, pragmatics, and mental health in the whole study sample (N = 17). As the sample size was small, and as it had a wide age range, only non-parametric statistical analyses were performed in the calculations. Individual data on BRIEF, phonological working memory, CCC-2, SDQ, and early language and speech intelligibility results after 3 years with the first CI are presented in **Appendix 3**.

RESULTS

We addressed three research questions in the current follow-up study that were related to possible similarities and differences in EF outcome, pragmatics, and mental health in a sample of deaf children with CI and with different etiological backgrounds. The groups (cCMV and Cx26) were initially matched based on age, hearing (CI), vocabulary (BNT; raw scores), and non-verbal cognitive ability (Ravens matrices). There were no statistically significant differences between groups (cCMV and Cx26) regarding the speech recognition outcome (**Table 4**), parent education level, early language abilities pre-op and after 1 year with the first CI (**Table 3**), or expressive grammar levels after 3 years with the first CI (p > 0.05) (**Table 3**). Nevertheless, there were two statistically significant group median differences for language understanding (Reynell-III) (Edwards et al., 1997)

TABLE 3 | Early speech, language, and hearing outcome (pre-op, post-op after 1 and 3 years with 1st CI), and age at walking (months), on group level (cCMV infection and Cx26 mutations), including statistical values for group comparisons (Mann Whitney *U*-test, and calculated effect sizes).

Abilities/tests (Md, min-max)	cCMV infection (*n* = 10)	Cx26 mutations (*n* = 7)	Z	*p*-value	r
Language understanding					
(Reynell-III, raw scores)					
Pre-op	0 (0–1), (*n* = 9)	0 (0–42)	−1.01	*p* = 0.31	0.25
1-year post-op	13 (3–43), (*n* = 9)	25 (17–53), (*n* = 6)	−1.89	*p* = 0.06	0.49
3- years post-op	47 (37–52), (*n* = 9)	51 (51–54), (*n* = 5)	−2.09	*p* = 0.04	0.56
Speech intelligibility rating					
(SIR-2, clinical judgement)					
Pre-op	1 (1–5)	1 (1–4)	−0.69	*p* = 0.49	0.17
1-year post-op	2 (2–3), (*n* = 9)	3 (2–4), (*n* = 6)	−1.86	*p* = 0.06	0.45
3-years post-op	4 (2–5), (*n* = 9)	5 (4–5)	−2.08	*p* = 0.04	0.50
Expressive grammar level					
(EGL, clinical judgement)					
Pre-op	2 (1–8)	2 (1–7)	−0.41	*p* = 0.68	0.10
1-year post-op	6 (4–7), (*n* = 9)	6 (5–8), (*n* = 6)	−1.05	*p* = 0.30	0.27
3-year post-op	7 (7–8), (*n* = 9)	8 (7–8)	−1.39	*p* = 0.17	0.34
Age at walking (months)	18 (12–23), (*n* = 9)	12 (11–13)	−3.05	*p* = 0.002	0.76

Language understanding; Reynell-III (Edwards et al., 1997); Speech Intelligibility Rating (SIR-2) (Allen et al., 2001), rating scale 1–5; Expressive Grammar Level (EGL), rating scale 1–8 (Löfkvist, 2014); Pure-Tone Average (PTA) performed with best aided situation; Age at walking, information from parents.

and speech intelligibility (SIR-2) (Allen et al., 2001) after 3 years, with better results for children with Cx26 mutations (**Table 3**). Individual test results on BRIEF, CCC-2, SDQ, and the two working-memory tasks are presented in **Appendix 3**.

Question 1: *Do children with CI have worse EF results in relation to norm data, regardless of cause of deafness, and do children with a cCMV infection have even poorer executive functioning compared to children with genetic non-syndromic deafness (Cx26)?*

Executive Functioning on Tests (Working Memory, Attention)
Working Memory
There was one statistically significant difference between children with cCMV and children with Cx26 on the phonological working memory test ($Z = −2.30$, $p = 0.02$, $r = 0.56$), with worse results for children with cCMV, while there were no statistically significant differences between groups on general working memory ($Z = −0.95$, $p = 0.34$, $r = 0.23$).

Attention Level
Attention level was assessed with the TEA-Ch test in all children older than 6 years. Although, there were only six children with cCMV and four children with Cx26, one statistically significant difference was found on one subscale; "walk don't walk" targets impulse control under time pressure ($Z = −2.0$, $p = 0.04$, $r = 0.63$). Due to missing data on some subtests for a few individuals and in combination of the small numbers in the sample, it was not possible to further evaluate whether the results were comparable or worse than for peers with TH in the same ages (norm data).

Emotional, Behavioral, and Attention Rating (EBA-R)
There were no statistically significant group differences on any of the scales: Expression of positive emotions; Frustration level ($Z = −1.61$, $p = 0.11$, $r = 0.39$); Restlessness level ($Z = −1.49$, $p = 0.14$, $r = 0.36$); Focus level ($Z = −1.30$, $p = 0.20$, $r = 0.32$); Problem solving (structured ability, logical behavior) ($Z = −1.93$, $p = 0.05$, $r = 0.49$); or Problem solving (unstructured ability, chaotic behavior) ($Z = −1.69$, $p = 0.09$, $r = 0.41$).

Executive Functions in Everyday Settings (Home and Preschool/School)
The group median results of the BRIEF rating indicated slightly worse results than expected in relation to norm data for children with TH, but there was a large variation within the cCMV group. The majority of children with cCMV were within limits of typical levels compared to American norm data. We found no statistically significant group differences (cCMV and Cx26) ($p > 0.05$). Nevertheless, there were three individuals with cCMV who had poorer EF results than controls and in relation to norm data, which should be examined further in more in-depth investigations by a clinical psychologist (see **Figures 1**, **2**).

Although there was some variation in outcome between individuals within the Cx26-group, there was no child with genetic deafness who reached a t-score over 65 on either the BRI, MI, or GEC, indicating that children with Cx26 were within typical levels for children with TH in the same ages (norms). This suggests that children with Cx26 deafness did not have specific EF problems at home or in preschool/school. One child with Cx26 had results that scored higher than average on working memory and shifting (two subscales in
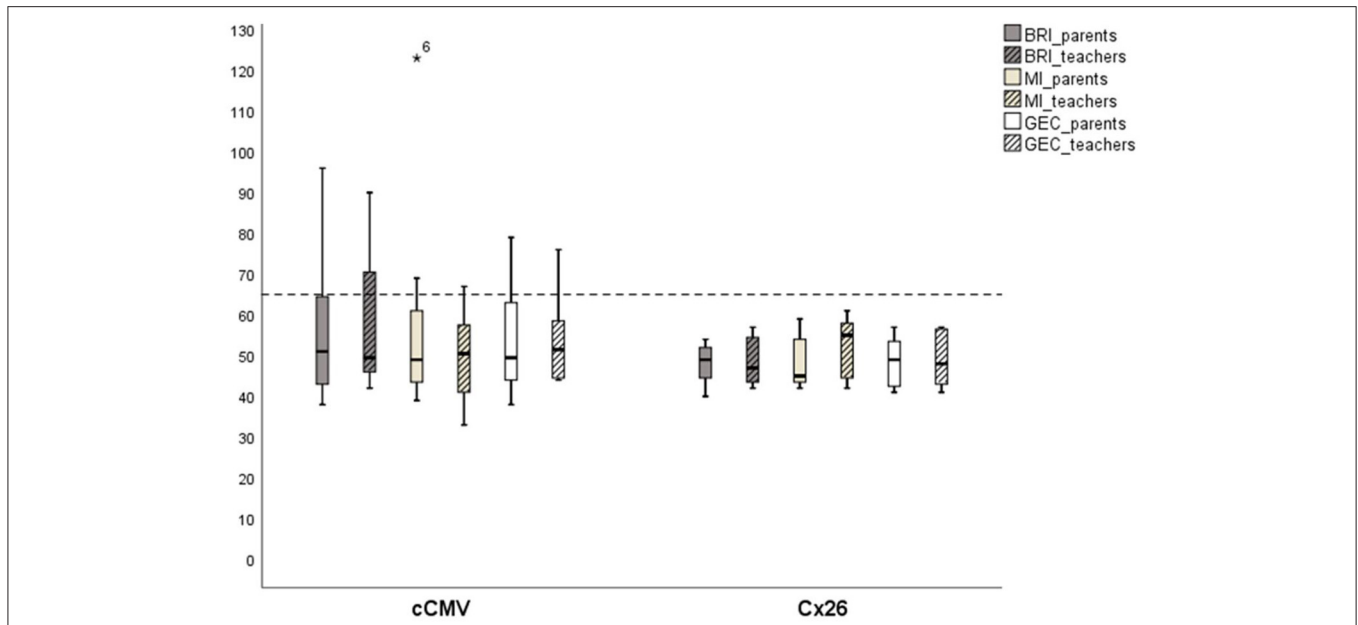
FIGURE 1 | *Executive functions in home and pre-school/school environment (BRIEF-results)*. Children deafened due to congenital cytomegalovirus (cCMV) infection (*n* = 8) and connexin 26 (Cx26) (*n* = 7), showing *t*-score results on group level; Behavior Rating Inventory (BRI), Metacognition Index (MI), and Global Executive Composite (GEC), rated by parents and teachers. Scores over 65 is considered to be clinically atypical. *Participant 6.
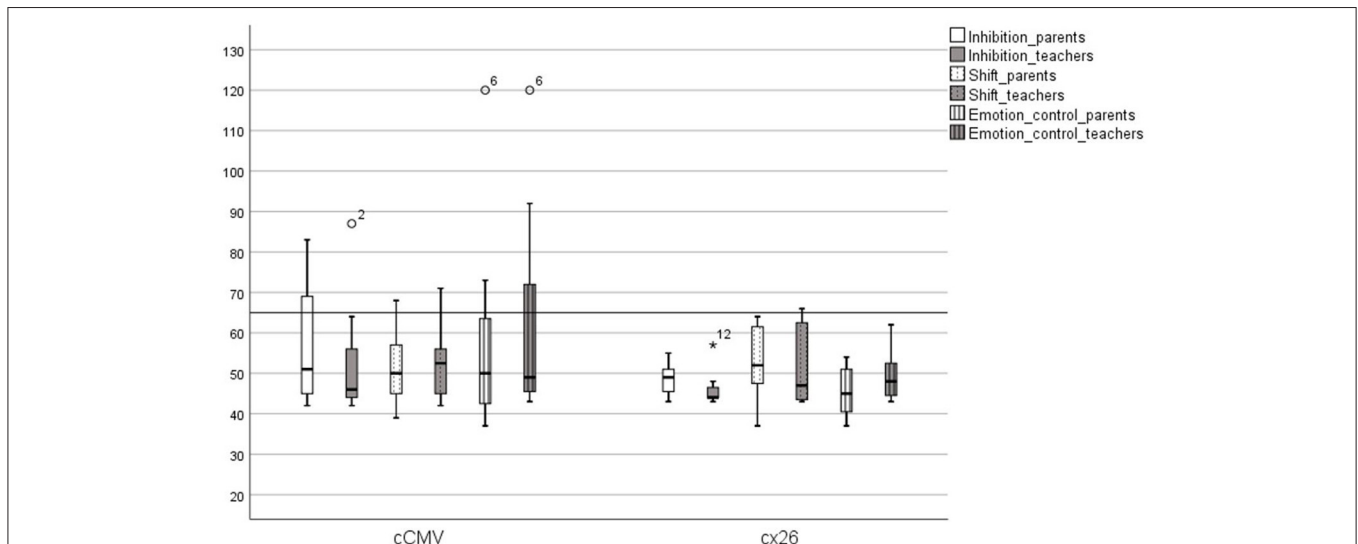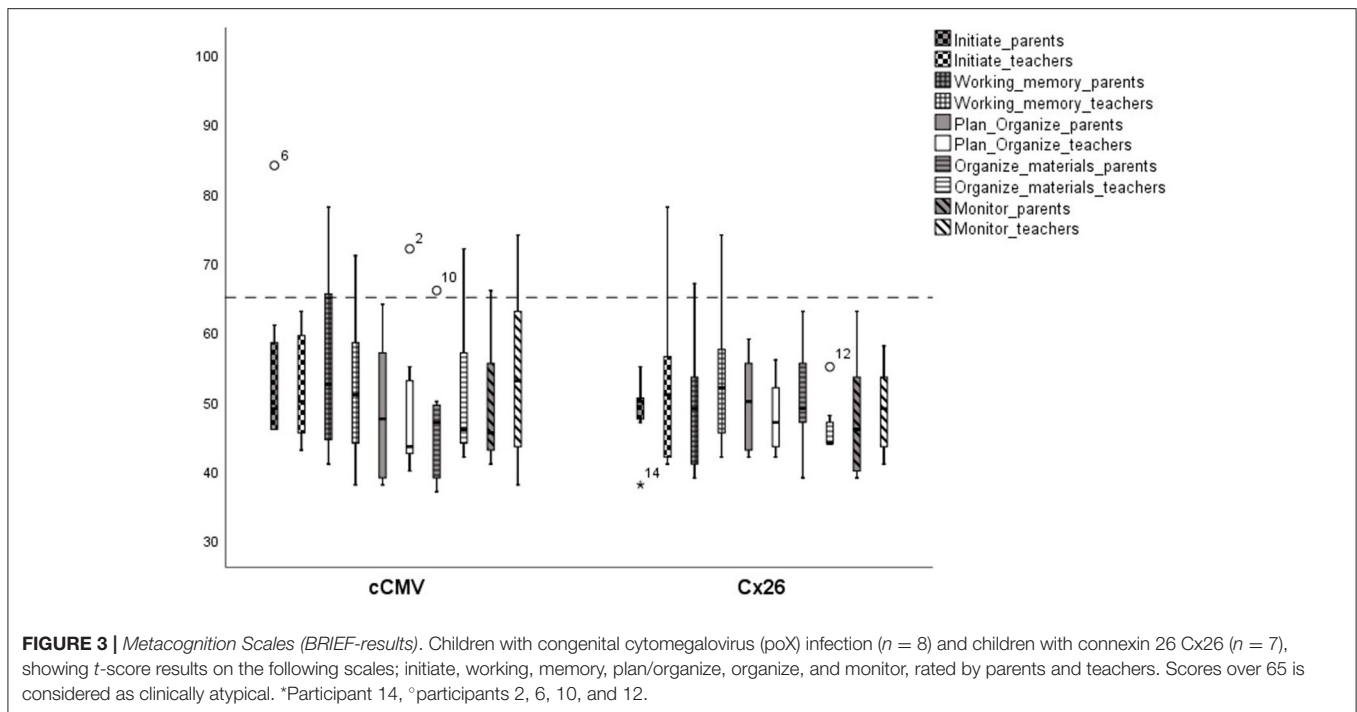

FIGURE 2 | *Behavioral Regulation Scales (BRIEF-results)*. Children with congenital cytomegalovirus (cCMV) infection and children with connexin 26 (Cx26), showing *t*-score results on the following scales; inhibition, shift, and emotion control, rated by parents and teachers. *Participant 12, °participants 2 and 6.

BRIEF), which means that this child could have slightly worse results than expected, but not clinically atypical (see **Figures 1**–**3**).

To summarize, our first hypothesis that children with CI in both groups (cCMV and Cx26) had worse EF outcomes than children with TH was only partly confirmed by these pilot results. Children with cCMV had statistically significant worse phonological memory abilities than children with Cx26.

Due to the small sample size and missing data from the TEA-Ch test we could not conclude that children with cCMV had substantially poorer attention and impulse control than children with Cx26 mutations. Three individuals with cCMV had BRIEF results that indicated they should be referred to a clinical psychologist for a more thorough investigation of their EF, while there were none in the control group with similar indications.

FIGURE 3 | *Metacognition Scales (BRIEF-results)*. Children with congenital cytomegalovirus (poX) infection (*n* = 8) and children with connexin 26 Cx26 (*n* = 7), showing *t*-score results on the following scales; initiate, working, memory, plan/organize, organize, and monitor, rated by parents and teachers. Scores over 65 is considered as clinically atypical. *Participant 14, °participants 2, 6, 10, and 12.

Question 2: *Do children with a cCMV infection who use CI have worse pragmatic skills and mental health status than well-matched children with Cx26 in comparison to norm data?*

Pragmatic Skills

Results on the parent questionnaire CCC-2, measuring the child's pragmatic skills, showed significant differences between groups (cCMV infection and Cx26) on the IGK/total raw score ($Z = -2.28$, $p = 0.02$, $r = 0.57$), and on two subscales; Initiatives ($Z = -2.40$, $p = 0.02$, $r = 0.60$), and Use of context ($Z = -2.87$, $p = 0.002$, $r = 0.72$) (**Figures 4**, **5**).

Mental Ill Health

All children in the sample, with a few exceptions, had typical results on mental health (SDQ) compared to norm data (**Table 4**). There were only three statistical differences between groups (cCMV and Cx26) on the SDQ results for individual subscales for results reported by fathers, which was related to more conduct problems and peer problems in the group of children with cCMV (see **Table 5**).

The second hypothesis we had before the study was that children with cCMV would have worse pragmatic skills than hearing-matched controls due to (presumed) worse executive functioning. The results showed statistically significant differences between groups (cCMV and Cx26), both on total raw score and on subscales that are related to conversational skills (initiatives and use of context); both are important for social cognition and could be related to *attention skills* and *flexibility* (EF). We hypothesized that worse pragmatic skills in children with cCMV could be explained by not only their auditory deprivation and HI but also by other consequences related to

atypical MR findings and a congenital virus infection, which is known to be associated with other deficits (Karltorp et al., 2014). The results indicated that the statistically significant group differences and effect sizes on CCC-2 were not only explained by the HI only but by other reasons too. However, the sample size was small, which made it difficult to generalize the findings on population level.

Question 3: *Is there a relationship between EF, pragmatic skills, mental health, and early language abilities in children with CI, regardless of cause of deafness?*

Correlation Analyses

There were some correlations among EF, pragmatic skills, mental health, and level of language understanding and speech intelligibility rating after 3 years with the first CI, and these are presented in **Table 6**. These results represent the whole sample (cCMV and Cx26). The results showed statistically significant correlations both between the higher level of pragmatic skills and early language abilities as well as for pragmatic skills, mental health levels rated by parents, and some weaker correlations with phonological working memory (**Table 6**).

One initial hypothesis was that there would be a relationship between EF, pragmatic skills, and mental health in all children with CI, regardless of their cause of deafness, and that better speech and language understanding in early childhood would be related to better outcomes in pragmatics, EF, and social behavior in later childhood (Goberis et al., 2012; Hintermair, 2013). Apparently, children with cCMV showed more of that expected interaction pattern, but, because of the small sample
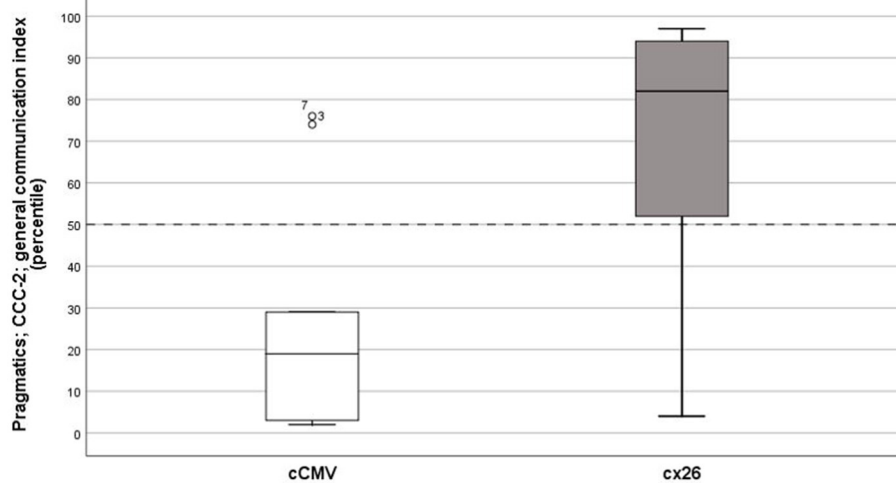
FIGURE 4 | *Pragmatic skills*. General communication index, median percentile score on the Children's Communication Checklist (CCC), second edition, on group level congenital cytomegalovirus (cCMV) infection; *n* = 9 and connexin 26 (Cx26); *n* = 7.
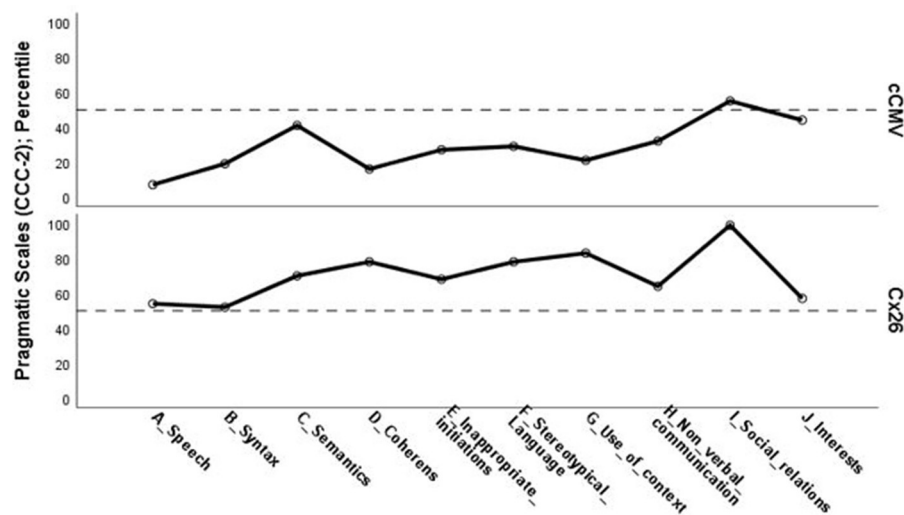

FIGURE 5 | *Pragmatic skills*. Percentile results (Md) on individual scales on the Children's Communication Checklist (CCC), second edition, on group level; congenital cytomegalovirus (cCMV) infection: *n* = 9, and connexin 26 (Cx26): *n* = 7.

size, the statistical correlation results had to be interpreted with caution, and the relation did not showing a casual effect.

DISCUSSION

In this follow-up study of children with cCMV compared to well-matched controls, we started with three research questions that arose after previous findings indicated that children with cCMV might have specific EF difficulties, which could affect their social or pragmatic development/behavior (Karltorp et al., 2014). As a group, most participants had age-adequate EF results compared to American norms for TH children

concerning EF in everyday settings, rated by parents and teachers, which was somewhat surprising considering previous findings in the literature (Figueras et al., 2008; Kronenberger et al., 2014; Korndewal et al., 2017). When looking more closely at the subgroup patterns (cCMV vs. Cx26), and in relation to individual results, there were three children with cCMV who did not perform like typically developed children with TH and who should therefore be referred to a clinical psychologist to conduct more in-depth investigations of their EF.

On a group level, children with cCMV did have statistically significant worse phonological working memory than matched controls, but there was no group difference on general

TABLE 4 | Language, non-verbal cognition and hearing outcome measures (median, range, and statistical values), compared on group level (cCMV vs. Cx26), including the effect sizes.

	cCMV (n = 10)	Cx26 (n = 6)	Z	p-value	r
Vocabulary (BNT)					
Raw scores	31 (0–46)∞	34 (23–53)	−0.82	p = 0.43	0.05
Stanine	4 (1–9)∞	3 (1–9)	−0.48	p = 0.63	0.02
Lexical-Semantic error analysis (BNT)					
Semantic relevant errors	11 (6–24)∞	11 (6–22)	−0.42	p = 0.68	0.01
Semantic irrelevant errors	11 (3–22)∞	10 (0–21)	−0.53	p = 0.60	0.07
No responds	0 (0–13)∞	0 (0–8)	−0.14	p = 0.89	<0.01
Phonemic word fluency (FAS letter WF)					
Total numbers	18 (0–34)	13 (1–37)*	−0.20	p = 0.84	<0.01
Semantic word fluency (Animal WF)					
Total numbers	13 (2–19)	10 (7–17)	−0.33	p = 0.74	<0.01
Non-verbal cognitive ability					
Raw scores	25 (14–34)	32 (14–35)	−1.36	p = 0.17	0.12
Speech recognition					
Quiet (%)	84 (64–100)§	68 (48–100)	−0.83	p = 0.41	0.05
Noise (%)	50 (32–68)×	56 (32–68)¤	−0.52	p = 0.60	0.02

Missing data in the Cx26-group: one child (Cx16) only participated with parent questionnaires, *(n = 5), ¤(n = 4). Missing data in the cCMV-group: §(n = 9); ×(n = 8); ∞(n = 9); one child (CMV2) did not want to participate in the BNT for unknown reason, and therefore there are also missing data on the lexical-semantic error analysis for one individual.

TABLE 5 | Mental health.

		cCMV (n = 9)	Cx26 (n = 7)	Z	p-value	r
SDQ—total score	F	6 (2–12)	3 (0–7)	−1.75	p = 0.08	0.44
	M	6 (0–12)	3 (0–9)	−1.01	p = 0.31	0.25
	T	4 (0–16)	3 (0–10)	0.00	p = 1.00	–
A. Emotional symptoms	F	0 (0–2)	1 (0–2)	−1.01	p = 0.31	0.25
	M	0 (0–3)	1 (0–4)	−1.74	p = 0.08	0.44
	T	4 (0–16)	0 (0–3)	−0.86	p = 0.39	0.22
B. Conduct problems	F	2 (0–3)	0 (0–1)	−2.23	p = 0.03	0.56
	M	1 (0–3)	0 (0–3)	−0.82	p = 0.41	0.21
	T	0 (0–3)	0 (0–2)	−0.15	p = 0.88	0.04
C. Hyperactivity	F	3 (1–6)	0 (0–1)	−1.51	p = 0.13	0.38
	M	3 (0–8)	0 (0–5)	−1.39	p = 0.16	0.35
	T	2 (0–10)	2 (0–7)	−0.89	p = 0.39	0.22
D. Peer problems	F	1 (0–3)	0 (0–1)	−2.12	p = 0.03	0.53
	M	0 (0–5)	0 (0–1)	−0.42	p = 0.68	0.11
	T	0 (0–2)	0 (0–1)	−0.43	p = 0.67	0.11
Prosocial behavior	F	8 (6–10)	10 (7–10)	−1.94	p = 0.05	0.49
	M	9 (7–10)	10 (8–10)	−0.88	p = 0.38	0.22
	T	7 (5–10)	9 (2–10)	−0.07	p = 0.95	0.02

Total score on SDQ and scores on subscales (Md, min-max) on group level (cCMV and Cx26), including Md group comparisons (Mann Whitney). F, fathers; M, mothers; T, teachers. Missing data in cCMV-group: reports from one mother (n = 8), two fathers (n = 7) and three teachers (n = 6). Missing data in Cx26-group: report from one father (n = 6). Higher scores on total score and subscales A–D indicate more difficulties, while higher scores on the prosocial behavior subscale reflect strengths.

TABLE 6 | Correlation coefficients for EF skills, pragmatics and mental health rated by mothers and fathers, and early language abilities after 3 years with 1st CI; language understanding (Reynell-III) and speech intelligibility (SIR-2), for children with cCMV and Cx26.

	1	2	3	4	5	6	7
1. Phonological WM	–	−0.56*	0.61*	0.55	0.39	−0.48	−0.65*
2. EF skills, GEC (BRIEF)		–	−0.58*	−0.42	−0.34	−0.70**	0.38
3. Pragmatics (CCC-2)			–	0.74**	0.63*	−0.73**	−0.63*
4. Language understanding (Reynell)				–	0.65*	−0.48	−0.48
5. Speech intelligibility (SIR-2)					–	-32	−0.43
6. SDQ, total score, M						–	0.65*
7. SDQ, total score, F							–

*EF skills, GEC, executive function skills, Global Executive Composite in BRIEF. M, mothers; F, fathers. Missing data from individual tests are reported in **Tables 3–5**. *p ≤ 0.05; **p ≤ 0.01.*

working memory. This indicates group-specific differences in how linguistic information is processed. Children with cCMV appeared to find it especially more difficult to process

phonologically based information without semantic clues than children with Cx26. The children in the whole sample showed variation in the vocabulary outcome, but there were no statistically significant subgroup differences regarding the vocabulary size (total score on BNT) or the lexical-semantic error response analysis. Children with cCMV performed well on the FAS letter-fluency task, which means that children in the sample (on a group level) had sufficient and effective strategies to learn words and retrieve lexical-semantic information from their long-term memory despite the fact that they also had a worse ability to process non-words (Löfkvist et al., 2012; Löfkvist, 2014).

Some individuals with a cCMV infection did not complete all the tests, due either to fatigue or for unknown reasons,

while children with Cx26 did not complain about fatigue in the same way, indicating worse *attention abilities* in children with a cCMV infection. Still, the EBR observation performed by the blinded psychologist did not show statistically significant group differences in performance during the test situation while performing cognitive tasks. One limitation was that the EBR observation was only performed in one test situation. It would have been useful to perform the same kind of observation also during the language and hearing assessment to further explore possible group differences related to the fatigue and attention level of participants during other assessments at the same test occasion. The attention measures in the TEA-Ch test were especially difficult to interpret, mainly due to few completed test results. The discovery of fatigue and attention difficulties in participants who did not complete the TEA-Ch-test suggests that a more sensitive measure of attention may be more informative to use in future studies. The result, however, also indicated that the TEA-Ch-test was challenging for all participants because it was assessing attention skills (Figueras et al., 2008; Beer et al., 2014; Kronenberger et al., 2014).

Children with CI, regardless of their cause of deafness, had more or less typical mental health results that were comparable to norm data of age-matched children with TH, which is a positive result. Only two aspects related to mental health differed in the two subgroups. Fathers reported conduct problems and poor peer functioning in children with cCMV compared to the hearing-matched controls with Cx26, which could be related to worse EF (Lyxell et al., 2009) and/or poorer pragmatic skills (Goberis et al., 2012).

There were statistically significant differences between groups on the total score of the pragmatic skills questionnaire (CCC-2) as well as for the subscales *initiatives* and *use of context* (in dialogues). Children with cCMV appeared to have worse ability to make use of the context in social interactions, and, according to parent reports, they used more initiatives that were irrelevant in verbal interactions despite having a similar language level and non-verbal cognitive ability as children with Cx26. The children with cCMV could be at risk of having more affected pragmatic skills than controls due to a later HI diagnosis age, resulting from their progressive hearing loss, and deviant pragmatic skills that were more related to their congenital CMV infection and atypical brain patterns (Karltorp et al., 2014).

Most et al. (2010) investigated pragmatic skills in a sample of 24 children with HI aged 6; 3–9; 4 years who had CI (*n* = 11) or used hearing aids (*n* = 13) and with 13 controls with TH. The pragmatic skills were similar for all participants with HI regardless or type of hearing technology. On a group level, children with HI had statistically significant poorer outcomes than children with TH. The authors concluded that their less effective pragmatic skills could be explained by impaired auditory perception of spoken language, less flexible use of language in combination with deficits in theory of mind, less exposure to different pragmatic situations, and poor use of repairing strategies. An additional explanation for their delayed pragmatic skills was their late diagnosis of HI (1; 8 years), which was influenced negatively by prolonging the length of auditory deprivation, especially for the deaf children who had a mean

age of 2; 6 years when they received their first CI. Cause of deafness was not investigated in the study (Most et al., 2010). The study findings by Most et al. (2010) showed that children with late identification and management of HI had a more delayed acquisition of pragmatic competence, which could lead to consequences, not only in social interaction with friends, but also in learning situations.

The fathers in the current study reported statistically significant less-well conduct levels and peer problems in the cCMV infection group compared to controls. These two functions are interrelated and associated with social behavior. Poor behavior could lead to affected peer relations. Conduct level might also be related to pragmatics (Goberis et al., 2012), EF abilities like attention and phonological working memory (Lyxell et al., 2009), as well as social behavior (Hintermair, 2013). We found some correlations in the whole sample (cCMV and Cx26) between early language skills after 3 years with the first CI and later outcomes at the follow-up study, not only with pragmatic skills but also with phonological working memory and mental health. Better early language skills were associated with better pragmatic skills and phonological working memory at later ages. Better pragmatic skills were also related to better mental health. These findings should be investigated further in larger groups (cCMV and Cx26) to find out if there are more specific subgroup differences and if this has any relation to the children's own perceived mental health. In the current study, only parents and teachers responded on behalf of their child. An analysis of the child's own perceived mental health could give another result with worse self-perceived mental health in comparison to the view of the child's parents, which has been reported in previous studies of children with CI (Anmyr et al., 2012).

The participating children in the two groups were initially matched based on age, hearing-level, vocabulary knowledge, non-verbal cognitive ability, home language situation (at least one parent who speak Swedish), and no other known additional diagnoses besides the deafness. Furthermore, at the time of the follow-up study we found no differences between groups based on socio-economic status (parental education level). Nonetheless, there were some significant statistical group differences between children that were related to their early childhood. Children with cCMV on average started to walk later than children with Cx26, which is suggestive of a balance problem that has been reported on before (Karltorp et al., 2014). They also had statistically significant worse language understanding and speech intelligibility after 3 years with their first CI compared to the results of controls with Cx26 (Ramirez Inscoe and Nikolopoulos, 2004; Yoshida et al., 2009), while there were no group differences after 1 year with their first CI. Apparently, children with a cCMV infection developed their spoken language at a slower pace than children with Cx26 despite there being a possible better hearing situation as infants in some cases of cCMV infection due to the late onset of HI, which is a result that has been reported on before (Ramirez Inscoe and Nikolopoulos, 2004; Yoshida et al., 2009; Karltorp et al., 2014). In a follow-up study by Yoshida et al. (2017), in 16 children with a cCMV infection and with a mean follow-up time of 7.8 years after the first CI, the authors found that some children who had initial delayed

language had caught up in speech and language understanding. Yet, there were some children who instead showed increased difficulties related to the incidence of additional diagnose(s) and more brain abnormalities in infancy (Yoshida et al., 2017), which is similar to the results of the current study, with a large variation of outcome, especially in the cCMV group.

Another difference between groups (cCMV and Cx26) in the current study was their early daycare environment. All but one child with Cx26 went to mainstream daycare from the start and continued to be mainstreamed onwards, while some children with cCMV infection had initially attended special units for deaf and hearing-impaired children, with more exposure to sign language and total communication, before they changed to mainstream preschools/schools. All participants with cCMV had parents who were TH; all children with cCMV therefore had access to spoken language throughout their early childhood in their home environment. We therefore have no reason to believe that the initial different daycare settings would explain later group differences in EF abilities or language outcome, including their pragmatic skills. The worse spoken language understanding level after 3 years with CI in children with cCMV infection could potentially be related to limited exposure of spoken language in daycare, but is more likely explained by previous findings that there is a slower pace in speech and language development in children with cCMV compared to other subgroups of children with CI (Ramirez Inscoe and Nikolopoulos, 2004; Yoshida et al., 2009).

Study Limitations and Future Studies

Although the present study was limited in the number of individuals, the pilot study contributed with new knowledge about executive functioning, pragmatic skills, and mental health in deaf children with cCMV who use CI as well as for matched controls with Cx26 mutations. Future studies should look more closely into individual results in children with a cCMV infection. It would be beneficial to conduct a study with a longitudinal study design to further examine the developmental aspects of executive functions and pragmatic skills and include theory-of mind as an aspect in relation to the children's mental health, including their own self-perceived opinion and perspectives. Comparative cross-sectional studies should include more participants with cCMV and controls with TH who are matched based on age and socio-economic status and preferably also including a control group of typically hearing children with ADHD.

To conclude, children with a cCMV infection who used CI, and who did not have previous known diagnoses like ADHD, DLD, or ASD, had worse pragmatic skills and phonological

working memory compared to well-matched controls with Cx26 and CI. Both groups with CI had typical mental health according to parent and teacher reports; some fathers' reports, however, showed more conduct problems and poor peer functioning in the group of children with cCMV infection. Parents and teachers did not report severe EF difficulties in everyday settings on group level. Better early language skills after 3 years of CI use was correlated to better pragmatic skills and mental health at later ages. The results indicate that it is important to identify children with cCMV as early as possible and support them and their families with preventive language stimulation actions, including specific training of social and pragmatic skills. Besides listening and language abilities, social cognition and EF should be assessed on a regular basis. This might limit the risk that subgroups like children with cCMV are left behind in social interaction and learning situations.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Regional Ethical Review Board in Stockholm, Sweden. All participants were first provided with written information about the study. Then a written informed consent was obtained from the parents of all participants, in accordance with the Declaration of Helsinki. The protocol was approved by the Regional Ethical Review Board in Stockholm, Sweden; DN 2012:/2.

AUTHOR CONTRIBUTIONS

All authors contributed with planning of the study, conducted the data collection, and performed data analysis/interpretation as well as discussed results and commented on the manuscript at all stages. UL conducted most of the statistical analysis as well as prepared the draft.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02808/full#supplementary-material

REFERENCES

Alford, R. L., Arnos, K. S., Fox, M., Lin, J. W., Palmer, C. G., Pandya, A., et al. (2014). American College of Medical Genetics and Genomics guideline for the clinical evaluation and etiological diagnosis of hearing loss. *Genet. Med.* 16, 347–355. doi: 10.1038/gim.2014.2

Allen, C., Nikopoulos, T. P., Dyar, D., and O'Donoghue, G. M. (2001). Reliability of a rating scale for measuring speech intelligibility

after pediatric cochlear implantation. *Otol. Neurotol.* 22, 19–31. doi: 10.1097/00129492-200109000-00012

Anmyr, L., Larsson, K., Olsson, M., and Freijd, A. (2012). Strengths and difficulties in children with cochlear implants – comparing self-reports with reports from parents and teachers. *Int. J. Pediatr. Otorhi.* 76, 1107–1112. doi: 10.1016/j.ijporl.2012.04.009

Asp, F., Mäki-Torkko, E., Karltorp, E., Harder, H., Hergils, L., Eskilsson, G., et al. (2012). Bilateral versus unilateral cochlear implants in children: speech

recognition, sound localization, and parental reports. *Int. J. Audiol.* 51, 817–832. doi: 10.3109/14992027.2012.705898

Baddeley, A. (2012). Working memory and language: an overview. *Annu. Rev. Psychol.* 63, 1–29. doi: 10.1146/annurev-psych-120710-100422

Beer, J., Kronenberger, W. G., Castellanos, I., Colson, B. G., Henning, S. C., and Pisoni, D. B. (2014). Executive functioning skills in preschool-age children with cochlear implants. *J. Speech Lang. Hear. Res.* 57, 1521–1534. doi: 10.1044/2014_JSLHR-H-13-0054

Bishop, D. V. M. (2003). *The Children's Communication Checklist Version 2 (CCC-2)*. London: Psychological Corporation.

Boons, T., Brokx, J. P., Dhooge, I., Frijns, J. H., Peeraer, L., Verneulen, A., et al. (2012). Predictors of spoken language development following pediatric cochlear implantation. *Ear Hear.* 33, 617–639. doi: 10.1097/AUD.0b013e3182503e47

Boppana, S. E., Fowler, K. B., Pass, R. F., Rivera, L. B., Bradford, R. D., Lakeman, F. D., et al. (2005). Congenital cytomegalovirus infection: association between virus in infancy and hearing loss. *J. Pediatr.* 146, 817–823. doi: 10.1016/j.jpeds.2005.01.059

Brusewitz, K., and Tallberg, I. M. (2010). The boston naming test and swedish children: normative data and response analysis. *Eur. J. Dev. Psychol.* 7, 265–280. doi: 10.1080/17405620802234500

Colletti, L., Mandàla, M., Zoccante, L., Shannon, R. V., and Colletti, V. (2011). Infants versus older children fitted with cochlear implants: performance over 10 years. *Int. J. Pediatr. Otorhi.* 75, 504–509. doi: 10.1016/j.ijporl.2011.01.005

Cornes, A. J. (2007). *Development of Computer-Administered Sign Language (Auslan) Versions of Commonly Used Child Mental Health Instruments and Examination of Their Reliability, Validity and Acceptability by Australian Deaf Adolescents*. Sydney, NSW: Department of Psychological Medicine; University of Sydney.

Dettman, S., Pinder, D., Briggs, R. J., Dowell, R. C., and Leigh, J. R. (2007). Communication development in children who receive the cochlear implant younger than 12 months: risks versus benefits. *Ear Hear.* 28, 11S−18S. doi: 10.1097/AUD.0b013e31803153f8

Diamond, A. (2013). Executive functions. *Annu. Rev. Psychol.* 64, 135–168. doi: 10.1146/annurev-psych-113011-143750

Edwards, S., Garman, M., Hughes, A., Letts, C., and Sinka, I. (1997). *The Reynell Developmental Langue Scales (RDLS III)*. Windsor: NFER-Nelson Publishing Company Ltd.

Eriksson, L., and Grundström, P. (2000). "Reynell developmental language scales III, språkförståelsedelen," in *Översättning Och Normering, Samt Studie Över Sambandet Mellan Testresultat Och Föräldrars Utbildningsnivå [Degree project]*. Umeå: Umeå University; Speech and Language Department.

Figueras, B., Edwards, L., and Langdon, D. J. (2008). Executive function and language in deaf children. *Deaf Stud. Deaf Educ.* 13, 362–377. doi: 10.1093/deafed/enm067

Fowler, K. B., McCollister, F. P., Sabo, D. L., Shoup, A. G., Owen, K. E., Woodruff, J. L., et al. (2017). A targeted approach for congenital cytomegalovirus screening within newborn hearing screening. *Pediatrics* 139:e20162128. doi: 10.1542/peds.2016-2128

Gathercole, S. E., Briscoe, J., Thorn, A., Tiffany, C., and ALSPAC Study Team (2008). Deficits in verbal long-term memory and learning in children with poor phonological short-term memory skills. *Q. J. Exp. Psychol.* 61, 474–490. doi: 10.1080/17470210701273443

Geers, A. E., Strube, M. J., Tobey, E. A., and Moog, J. S. (2011). Epilogue: factors contributing to long-term outcomes of cochlear implantation in early childhood. *Ear Hear.* 32, 84S−92S. doi: 10.1097/AUD.0b013e3181ffd5b5

Gioia, G. A., Isquith, P. K., Guy, S. C., and Kenworthy, L. (2000). Behavior rating inventory of executive function. *Child Neuropsychol.* 6, 235–238. doi: 10.1076/chin.6.3.235.3152

Goberis, D., Beams, D., Dalpes, M., Abrisch, A., Baca, R., and Yoshinaga-Itano, C. (2012). The missing link in language development of deaf and hard of hearing children: pragmatic language development. *Semin. Speech Lang.* 33, 297–309. doi: 10.1055/s-0032-1326916

Goodman, R. (1997). The strengths and difficulties questionnaire: a research note. *j. Child Psychol. Psychiatry* 38, 581–586. doi: 10.1111/j.1469-7610.1997.tb01545.x

Goodman, R. (2005). *Youth in Mind; SDQ-Information for Researchers and Professionals About the Strengths and Difficulties Questionnaire*. Aailable online at: http://www.sdqinfo.org

Goodman, R., Ford, T., Simmons, H., Garward, R., and Metzler, H. (2000). Using the strengths and difficulties questionnaire (SDQ) to screen for child psychiatric disorders in a community sample. *Br. J. Psychiatr.* 177, 534–539. doi: 10.1192/bjp.177.6.534

Goodman, R., Meltzer, H., and Bailey, V. (1998). The strengths and difficulties questionnaire: a pilot study on the validity of the self-report version. *Eur. Child Adolesc. Psychiatry* 7, 125–130. doi: 10.1007/s007870050057

Grosse, S. D., Ross, D. S., and Dollard, S. C. (2008). Congenital cytomegalovirus (CMV) infection as cause of permanent bilateral hearing loss: a quantitative assessment. *J. Clin. Virol.* 41, 57–62. doi: 10.1016/j.jcv.2007.09.004

Heaton, S. C., Reader, S. K., Preston, A. S., Fennell, E. B., Puyana, O. A., Gill, N., et al. (2001). The test of everyday attention for children (TEA-Ch): patterns of performance in children with ADHD and clinical controls. *Child Neuropsychol.* 7, 251–264. doi: 10.1076/chin.7.4.251.8736

Henricson, C., Wass, M., Lidestam, B., Möller, C., and Lyxell, B. (2012). Cognitive skills in children with Usher syndrome type 1 and cochlear implants. *Int. J. Pediatr. Otorhi.* 76, 1449–1457. doi: 10.1016/j.ijporl.2012.06.020

Hintermair, M. (2007). Prevalence of socioemotional problems in deaf and hard of hearing children in Germany. *Am. Ann. Deaf* 152, 320–330. doi: 10.1353/aad.2007.0028

Hintermair, M. (2013). Executive functions and behavioral problems in deaf and hard-of-hearing students at general and special schools. *J. Deaf Stud. Deaf Educ.* 18, 244–259. doi: 10.1093/deafed/ent003

Isquith, P. K., Gioia, G. A., and Espy, K. A. (2004). Executive function in preschool children: examination through everyday behavior. *Dev. Neuropsychol.* 26, 403–422. doi: 10.1207/s15326942dn2601_3

Kaplan, E., Goodglass, H., and Weintraub, S. (1983). *The Boston Naming Test, 2nd Edn*. Philadelphia, PA: Lea and Febiger.

Karltorp, E., Löfkvist, U., Lewensohn-Fuchs, I., Lindström, K., Westblad, M. E., Fahnehjelm, K. T., et al. (2014). Impaired balance and neurodevelopmental disabilities among children with congenital cytomegalovirus infection. *Acta Paediatr.* 103, 1165–1173. doi: 10.1111/apa.12745

Kave, G., Kigel, S., and Kochva, R. (2008). Switching and clustering in verbal fluency tasks throughout childhood. *J. Clin. Exp. Neuropsychol.* 30, 349–359. doi: 10.1080/13803390701416197

Korndewal, M. J., Oudesluys-Murphy, A. M., Kroes, A. C. M., Vossen, A. C. T. M., and de Melker, H. E. (2017). Congenital cytomegalovirus infection: child development, quality of life and impact on daily life. *Pediatr. Infect. Dis. J.* 36, 1141–1147. doi: 10.1097/INF.0000000000001663

Kronenberger, W. G., Colson, B. G., Henning, S. C., and Pisoni, D. B. (2014). Executive functioning and speech-language skills following long-term use of cochlear implants. *J. Deaf Stud. Deaf Educ.* 19, 456–470. doi: 10.1093/deafed/enu011

Lingås-Haukedal, C., von Koss Torkildsen, J., Lyxell, B., and and Wie, O. B. (2018). Parents' perception of health-related quality of life in children with cochlear implants: the impact of language skills and hearing. *J. Speech Lang. Hear. Res.* 61, 2084–2098. doi: 10.1044/2018_JSLHR-H-17-0278

Löfkvist, U. (2014). *Lexical and semantic development in children with cochlear implants* (Thesis). Karolinska Institutet, Stockholm, Sweden.

Löfkvist, U., Almkvist, O., Lyxell, B., and Tallberg, I. M. (2014). Lexical and semantic ability in groups of children with cochlear implants, language impairment and autism spectrum disorder. *Int. J. Pediatr. Otorhinolaryngol.* 78, 253–263. doi: 10.1016/j.ijporl.2013.11.017

Löfkvist, U., Lyxell, B., Almkvist, O., and Tallberg, I. M. (2012). Word fluency performance and strategies in children with cochlear implants: age-dependant effects? *Scand. J. Psychol.* 53, 467–474. doi: 10.1111/j.1467-9450.2012.00975.x

Lyxell, B., Sahlén, B., Wass, M., Ibertsson, T., Larsby, B., Hällgren, M., et al. (2008). Cognitive development in children with cochlear implants: relations to reading and communication. *Int. J. Audiol.* 47, S47–52. doi: 10.1080/14992020802307370

Lyxell, B., Wass, M., Sahlén, B., Samuelsson, C., Asker-Arnason, L., Ibertsson, T., et al. (2009). Cognitive development, reading and prosodic skills in children with cochlear implants. *Scand. J. Psychol.* 50, 463–474. doi: 10.1111/j.1467-9450.2009.00754.x

Malm, G., and Engman, M. L. (2007). Congenital cytomegalovirus infections seminars. *Fetal Neonatal Med.* 12, 154–159. doi: 10.1016/j.siny.2007.01.012

Malmberg, M., Rydell, A. M., and Smedje, H. (2003).Validity of the swedish version of strengths and difficulties questionnaire (SDQ-SWE). *Nord. J. Psychiatry* 57, 357–363. doi: 10.1080/08039480310002697

Manly, T., Anderson, V., Nimmo-Smith, I., Turner, A., Watson, P., and Robertson, I. H. (2001). The differential assessment of children's attention: the test of everyday attention for children (TEA-Ch), normative sample and ADHD performance. *J. Child Psychol. Psyc.* 42, 1065–1081. doi: 10.1111/1469-7610.00806

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., and Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: a latent variable analysis *Cogn. Psychol.* 41, 49–100. doi: 10.1006/cogp.1999.0734

Most, T., Shina-August, E., and Meilijison, S. (2010). Pragmatic abilities of children with hearing loss using cochlear implants or hearing aids compared to hearing children. *J. Deaf Stud. Deaf Educ.* 15, 422–437. doi: 10.1093/deafed/enq032

Quittner, A. L., Cruz, I., Barker, D. H., Tobey, E., Eisenberg, L. S., Niparko, J. K., et al. (2013). Effects of maternal sensitivity and cognitive and linguistic stimulation on cochlear implant users' language development over four years. *J. Pediatr.* 162, 343–348. doi: 10.1016/j.jpeds.2012.08.003

Ramirez Inscoe, J. M., and Nikolopoulos, T. P. (2004). Cochlear implantation in children deafened by cytomegalovirus: speech perception and speech intelligibility outcomes. *Otol. Neurotol.* 25, 479–482. doi: 10.1097/00129492-200407000-00014

Raven, J., Raven, J. C., and Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales.* San Antonio, TX: Pearson Assessment.

Smedje, H., Broman, J. E., Hetta, J., and von Knorring, A. L. (1999). Psychometric properties of a Swedish version of the "strengths and difficulties questionnaire." *Eur. Child Adolesc. Psychiatry* 8, 63–70. doi: 10.1007/s007870050086

Starnberg, J., Norman, M., Westrup, B., Domellöf, M., and Berglund, S. K. (2018). Lower cognitive test scores at age 7 in children born with marginally low birth weight. *Pediatr. Res.* 83, 1129–1135. doi: 10.1038/pr.2018.35

Szagun, G., and Stumper, B. (2012). Age or experience? The influence of age at implantation and social and linguistic environment on language development in children with cochlear implants. *J. Speech Lang. Hear. Res.* 55, 1640–1654. doi: 10.1044/1092-4388(2012/11-0119)

Tallberg, I. M. (2005). The boston naming test in swedish: normative data. *Brain Lang.* 94, 19–31. doi: 10.1016/j.bandl.2004.11.004

Tallberg, I. M., Carlsson, S., and Lieberman, M. (2011). Children's word fluency strategies. *Scand. J. Psychol.* 52, 35–42. doi: 10.1111/j.1467-9450.2010.00842.x

Walker, E. A., Redfern, A., and Oleson, J. J. (2019). Linear mixed-model analysis to examine longitudinal trajectories in vocabulary depth and breadth in children who are hard of hearing. *J. Speech Lang. Hear. Res.* 62, 525–542. doi: 10.1044/2018_JSLHR-L-ASTM-18-0250

Wass, M. (2009). *Children with cochlear implants - cognition and reading ability* (Thesis). Linköping University. Available online at: http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-51735

Wass, M., Ibertsson, T., Lyxell, B., Sahlén, B., Hällgren, M., Larsby, B., et al. (2008). Cognitive and linguistic skills in Swedish children with cochlear implants - measures of accuracy and latency as indicators of development. *Scand. J. Psychol.* 49, 559–576. doi: 10.1111/j.1467-9450.2008.00680.x

Willstedt-Svensson, U., Löfqvist, A., Almqvist, B., and Sahlén, B. (2004). Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *Int. J. Audiol.* 43, 506–515. doi: 10.1080/14992020400050065

Yoshida, H., Kanda, Y., Takahashi, H., Miyamoto, I., Yamamoto, T., and Kumagami, H. (2009). Cochlear implantation in children with congenital cytomegalovirus infection. *Otol. Neurotol.* 30, 725–730. doi: 10.1097/MAO.0b013e3181b1212e

Yoshida, H., Takahash, H., Kanda, Y., Kitaoka, K., and Hara, M. (2017). Long-term outcomes of cochlear implantation in children with congenital cytomegalovirus infection. *Otol. Neurotol.* 38, e190–e194. doi: 10.1097/MAO.0000000000001483

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor declared a shared affiliation, though no other collaboration, with one of the authors CH at the time of the review.

Listening Difficulties in Children: Behavior and Brain Activation Produced by Dichotic Listening of CV Syllables

David R. Moore[1,2,3*], **Kenneth Hugdahl**[4,5,6], **Hannah J. Stewart**[1,7], **Jennifer Vannest**[1,8,9], **Audrey J. Perdew**[1], **Nicholette T. Sloat**[1], **Erin K. Cash**[1] and **Lisa L. Hunter**[1,2]

[1] Communication Sciences Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States, [2] Department of Otolaryngology, University of Cincinnati College of Medicine, Cincinnati, OH, United States, [3] Manchester Centre for Audiology and Deafness, University of Manchester, Manchester, United Kingdom, [4] Department of Biological and Medical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway, [5] Department of Psychiatry, Haukeland University Hospital, Bergen, Norway, [6] Department of Radiology, Haukeland University Hospital, Bergen, Norway, [7] Division of Psychology and Language Sciences, University College London, London, United Kingdom, [8] Division of Neurology and Pediatric Neuroimaging Research Consortium, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, United States, [9] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States

Listening difficulties (LiD) are common in children with and without hearing loss. Impaired interactions between the two ears have been proposed as an important component of LiD when there is no hearing loss, also known as auditory processing disorder (APD). We examined the ability of 6–13 year old (y.o.) children with normal audiometric thresholds to identify and selectively attend to dichotically presented CV syllables using the Bergen Dichotic Listening Test (BDLT; www.dichoticlistening.com). Children were recruited as typically developing (TD; $n = 39$) or having LiD ($n = 35$) based primarily on composite score of the ECLiPS caregiver report. Different single syllables (ba, da, ga, pa, ta, ka) were presented simultaneously to each ear (6×36 trials). Children reported the syllable heard most clearly (non-forced, NF) or the syllable presented to the right [forced right (FR)] or left [forced left (FL)] ear. Interaural level differences (ILDs) manipulated bottom-up perceptual salience. Dichotic listening (DL) data [correct responses, laterality index (LI)] were analyzed initially by group (LiD, TD), age, report method (NF, FR, FL), and ILD (0, ± 15 dB) and compared with speech-in-noise thresholds (LiSN-S) and cognitive performance (NIH Toolbox). fMRI measured brain activation produced by a receptive speech task that segregated speech, phonetic, and intelligibility components. Some activated areas [planum temporale (PT), inferior frontal gyrus (IFG), and orbitofrontal cortex (OFC)] were correlated with dichotic results in TD children only. Neither group, age, nor report method affected the LI of right/left recall. However, a significant interaction was found between ear, group, and ILD. Laterality indices were small and tended to increase with age, as previously reported. Children with LiD had significantly

larger mean LIs than TD children for stimuli with ILDs, especially those favoring the left ear. Neural activity associated with Speech, Phonetic, and Intelligibility sentence cues did not differ significantly between groups. Significant correlations between brain activity level and BDLT were found in several frontal and temporal locations for the TD but not for the LiD group. Overall, the children with LiD had only subtle differences from TD children in the BDLT, and correspondingly minor changes in brain activation.

INTRODUCTION

Listening is often considered to be the active counterpart of passive hearing; "paying thoughtful attention to sound" (Keith et al., 2019; after Merriam-Webster). By definition, therefore, children with LiD may have problems with thought, attention, or hearing. In practice, a considerable number of children seen at audiology clinics who have LiD are, on further testing, found to have normal audiograms, the pure-tone detection, gold-standard measure of hearing (Hind et al., 2011). For these children, a wide variety of symptoms are reported by caregivers (American Academy of Audiology, 2010; Moore and Hunter, 2013) that may be summarized as difficulty responding to meaningful sounds while ignoring irrelevant sounds. For at least 40 years, children with these symptoms have, following further testing, been diagnosed by some audiologists as having an auditory processing disorder (APD), but that diagnosis has not gained universal acceptance (Moore, 2018), so we will generally refer to the symptoms here by the more generic and non-diagnostic term LiD.

Impaired interactions between the two ears have been proposed as an important component of LiD, based mainly on studies of DL, the simultaneous presentation of different acoustic signals to the two ears (Broadbent, 1956; Kimura, 1961; Keith, 2009). However, other aspects of binaural interaction, including binaural (Moore et al., 1991; Pillsbury et al., 1991) and spatial (Cameron and Dillon, 2007) release from masking have received substantial attention as contributors to LiD in adults and in children. Many other aspects of hearing and listening have also been studied in children with LiD (Moore et al., 2010; Weihing et al., 2015; de Wit et al., 2016; Wilson, 2018) leading, overall, to the emergence of two dominant hypotheses concerning the nature of the problem experienced by these children. The first, and more traditional hypothesis is that a disorder, (C)APD, is primarily a result of impaired processing of auditory neural signals in the central auditory system, defined as the brain

pathway from the auditory nerve to the auditory cortex (Rees and Palmer, 2010). The second, disruptive hypothesis is that LiD are due primarily to impaired speech/language synthesis, inattention, or other executive function impairment in cortical processing of auditory information beyond the central auditory system. The study reported here was motivated by an attempt to distinguish between these hypotheses.

There is a rich history of studies of DL in children going back at least to Kimura (1963). Many of the early studies included children with a variety of learning problems, of which reading disability was perhaps the most common. Interestingly, several of these studies appeared to equate language and other abilities now considered to be cognitive with central auditory processing. However, Roeser et al. (1983), studying dichotic CV syllables in both TD 6–10 y.o. children and children with language impairment, concluded that the "dichotic CV syllables test has limited prognostic value in identifying auditory processing dysfunction in children classified in having a learning disability."

More recently, some clinical DL tests have focused on listener reports of words (Keith, 2009), especially the spoken digits 1–10 (Musiek, 1983), that carry a substantial memory and executive control load in addition to their linguistic and acoustic demands. Nevertheless, dichotic digits, often described confusingly as a test of binaural integration (American Academy of Audiology, 2010; Brenneman et al., 2017), has become one of the most common clinical tests of APD (Emanuel et al., 2011). Various dichotic digit-based training programs have been proposed as interventions for the remediation of APD (Moncrieff et al., 2017). However, recent research (Cameron et al., 2016; Cameron and Dillon, 2020) has questioned whether dichotic digits testing involves any binaural interaction. These researchers found that performance on a *diotic* version of the test (presenting the same digits simultaneously to the two ears) correlates highly ($r = 0.8$) with performance on the dichotic version. The results suggest that, while binaural hearing may be disrupted during listening to dichotic digits, multiple, diverse abilities (acoustic discrimination, semantic identification, attentive listening, separation of two simultaneously presented sounds, accurate recall of heard digits) determine performance on these tasks.

As part of a larger Cincinnati Children's Hospital (CCH) program to investigate the nature of LiD in 6–13 y.o. children with normal audiometric thresholds (SICLID), we examined those children's ability to identify dichotically presented CV syllables using the BDLT (see www.dichoticlistening.com). In the BDLT (Hugdahl et al., 2009), two different CV syllables (from ba, da, ga, ka, pa, ta) are presented simultaneously, one to each ear,

and the listener is asked either to report the first or most clearly heard syllable (NF condition), or selectively to report only that syllable presented to the left (FL) or to the right (FR) ear. In the NF condition, also known as the "Listen" mode, the proportion of syllables presented to the right ear that is correctly reported consistently exceeds the proportion presented to the left ear that is correctly reported, a REA.

The REA is a long established, robust, bottom-up, stimulus-driven, perceptual effect. Historically, it has been found to decrease with age (Kimura, 1963), and also to decrease in children with learning problems (Harris et al., 1983). Based on observations of adult patients with large temporal lobe lesions, and volunteers with sodium amytal silencing of a whole hemisphere, the REA was proposed to reflect left hemisphere dominance for processing of speech (Kimura, 1961). More recently, the REA has been reflected in left-hemifield dominant activation of the auditory cortex in studies using fMRI (Hugdahl et al., 1999; Rimol et al., 2005; Westerhausen and Hugdahl, 2010). It is modulated by top-down, cognitive influences (attention, executive function, working memory, training), reflected in FR and, particularly, FL performance (Kompus et al., 2012; Hugdahl and Westerhausen, 2016). An acoustic ILD between the syllables can offset the REA and thus serve as a physical measure (in dB) of a cognitive construct (Hugdahl et al., 2008; Westerhausen et al., 2009). For these reasons, as well as its simplicity and the extensive literature on it, the BDLT is well suited to investigate neural processes of listening in children. This study represents the first, to our knowledge, where BDLT has been used to examine children with LiD/APD.

Data from two other SICLID test suites, the LiSN-S listening of sentences in spatialized noise (Cameron and Dillon, 2007), and the NIH Cognition Toolbox (Weintraub et al., 2013), are briefly presented here to examine possible correlations with functions revealed by BDLT testing. In particular, we were interested to know how BDLT data related to specific measures of space- and talker-based grouping of sounds, and of presumed underpinning cognitive function.

Previous studies have shown that BDLT "Concentrate" modes (FL, FR) activate different brain regions in adults when contrasted with the Listen mode (Westerhausen and Hugdahl, 2010). Thus, FR activates a "dorsal attention network," consisting of the rDLPFC and, weakly, lDLPFC and the bilateral occipital cortex. FL activates a "cognitive control network," consisting of the bilateral angular gyrus, DLPFC, and anterior cingulate cortex (Westerhausen et al., 2010). We have taken another approach to examine the neural mechanisms underlying performance on the BDLT. Specifically, we used a sentence listening and speaker identification test to produce BOLD activation inside a 3T MRI scanner. We contrasted aspects of the sentence listening task to isolate components of receptive speech (speech, phonetics, intelligibility). We examined the relationship between BOLD activation (Scott et al., 2000; Halai et al., 2015) for each group (TD, LiD) and children's performance by BDLT listening mode (NF, FL, FR) and interaural acoustic bias (ILD).

Testing the hypothesis that children with LiD have problems with cortical language, attention, and executive function beyond the central auditory system, the predictions of this study were that (i) children with LiD will perform normally on BDLT in Listen mode but will have difficulty in Concentrate mode, based on their overall tendency to perform poorly on cognitive tasks despite normal hearing; and (ii) children with LiD will have atypical top-down brain activation contrasts (Intelligibility, Speech), but typical Phonetic contrasts associated with BDLT performance. To investigate these predictions, we examined BDLT performance in normally hearing 6–13 y.o. children and correlated that performance with other tests of speech perception, cognition, and speech-evoked fMRI.
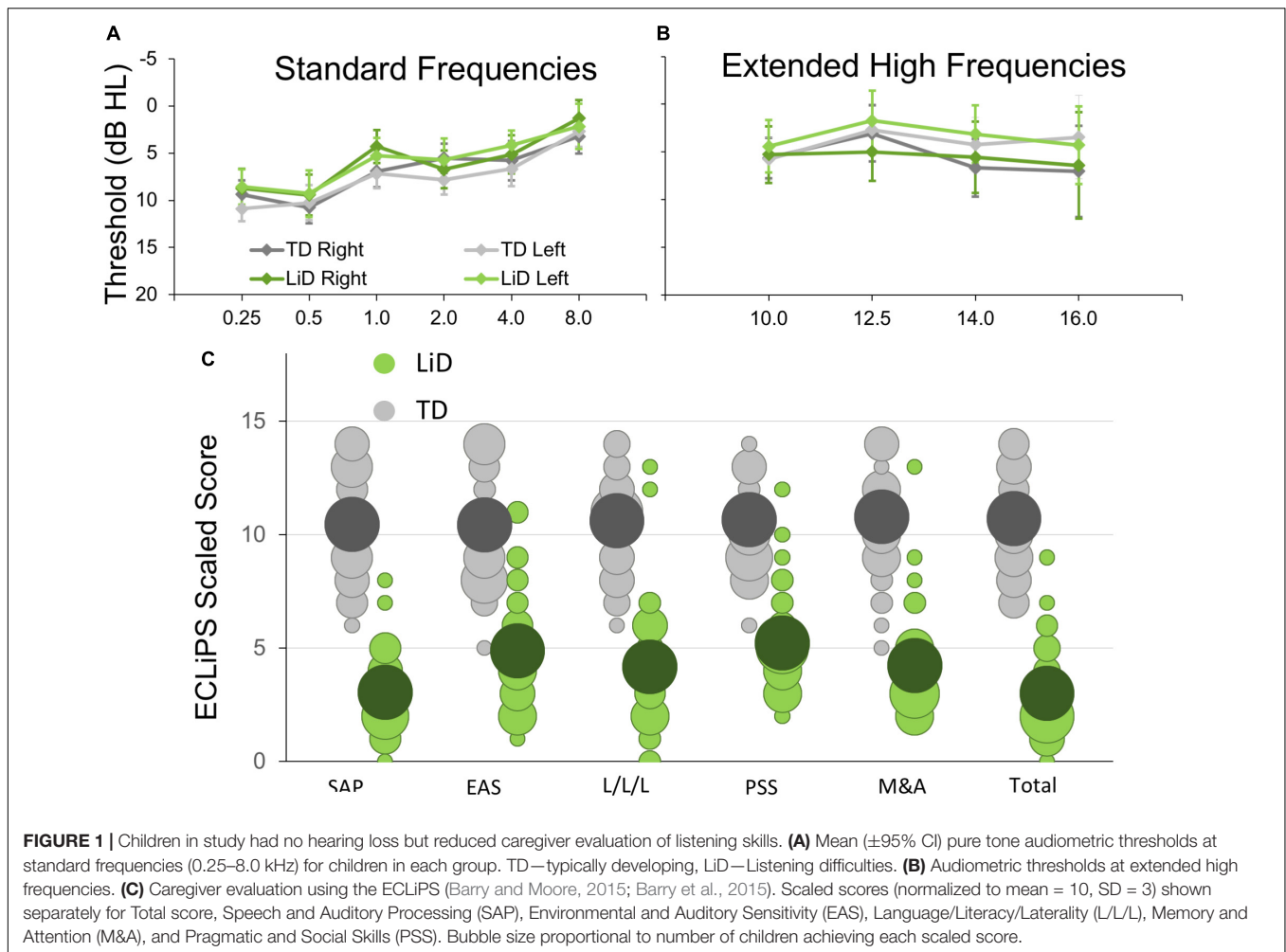
MATERIALS AND METHODS

Participants

Children with LiD were recruited initially from a medical record review study of over 1100 children assessed for APD at CCH (Moore et al., 2018). Caregivers of children diagnosed with APD (including those with a "Disorder" or a "Weakness") who responded to invitation to participate were sent questionnaires including the ECLiPS, below, and a background questionnaire on relevant demographic, medical (otology and neurology), and educational (learning disorders) issues. Those who completed and returned the questionnaires were invited to bring their child into the lab for a study visit. Over time, recruitment expanded to include the use of CCH IRB approved materials, advertising, and messages via print, electronic, social, and digital media at hospital locations and in the local and regional area for participation of families with children who had a "LiD," or were "without any known or diagnosed learning problem." Following a positive response and a brief telephone interview to screen for listening status, families were sent the same questionnaire pack and were invited for a study visit as described below.

Seventy four children aged 6–13 y.o. completed BDLT testing and most of the secondary behavioral testing. All of these children had normal hearing, bilaterally, defined as clear ears, A-type tympanometry, and pure tone thresholds $\leq$ 20 dB at octave frequencies between 0.25 and 8 kHz (**Figure 1A**) using standard audiometric procedures. Additional, extended high frequency audiometry (10–16 kHz; **Figure 1B**) was also obtained, but inclusion did not require any criterion level of performance at those frequencies. Seventy children received MRI scanning (95%).

The ECLiPS questionnaire (Barry and Moore, 2015; Roebuck and Barry, 2018) is a 38-item inventory asking users to agree or disagree (five-point Likert scale) with simple statements about their child's listening and related skills. Total standardized ECLiPS scores $\geq$ 5 designated TD, and scores < 5, or a previous diagnosis of APD, designated LiD, resulting in 39 TD children (mean age 9.84 years, SD = 2.19) and 35 children with LiD (mean age 10.16, years; SD = 2.14; **Figure 1C**). Of the children in the LiD group previously diagnosed with APD (n = 9; see below), two scored 5 or more on the ECLiPS. These children were nevertheless included in the LiD group.

Demographics, audiological status, secondary testing of auditory and speech perception, and cognitive performance of

FIGURE 1 | Children in study had no hearing loss but reduced caregiver evaluation of listening skills. **(A)** Mean (±95% CI) pure tone audiometric thresholds at standard frequencies (0.25–8.0 kHz) for children in each group. TD—typically developing, LiD—Listening difficulties. **(B)** Audiometric thresholds at extended high frequencies. **(C)** Caregiver evaluation using the ECLiPS (Barry and Moore, 2015; Barry et al., 2015). Scaled scores (normalized to mean = 10, SD = 3) shown separately for Total score, Speech and Auditory Processing (SAP), Environmental and Auditory Sensitivity (EAS), Language/Literacy/Laterality (L/L/L), Memory and Attention (M&A), and Pragmatic and Social Skills (PSS). Bubble size proportional to number of children achieving each scaled score.

the larger SICLID sample (*n* = 146) will be reported in greater detail elsewhere.

Behavioral Tests

Bergen Dichotic Listening Test (BDLT)

Digitally recorded test materials were provided by the Department of Biological and Medical Psychology, University of Bergen, Norway. Test materials and general procedures are described in detail elsewhere (Hugdahl et al., 2009). Listeners were seated in a sound treated booth and instructed to attend to and verbally repeat speech sounds presented via Sennheiser HD 25-1 headphones connected to a laptop PC. Control software was Direct RT. Two different CV syllables from a list of six (/ba/,/da/,/ga/,/ka/,/pa/,/ta/) were presented simultaneously, one to each ear at an initial level of 65 dB SPL. Each trial was started manually by the tester when the participant was ready. In an NF condition, the listener was asked to report the syllable they "heard best." Alternately, the listener was asked to report only that syllable presented to the left (FL) or to the right (FR) ear.

A test session started with 12 practice trials (NF). For the first six of these (ILD = 0 dB), the listener had to repeat one of two identical syllables correctly in five/six trials to proceed. For the

second six trials, different syllables were presented to each ear, ILD varied between + 15 (right louder) through -15 to 0 dB each two trials, and the listener had again to get five/six trials correct to proceed. The practice trials were repeated if a listener did not achieve the prescribed correct response rate. Five children (three LiD, two TD, in addition to the 74) were excused from the experiment when they failed to achieve the prescribed correct response rate. Data collection sessions (×6) each consisted of 36 trials containing every possible pair combination. The first two, NF sessions had 12 trials each of ILD = 0, + 15, −15. In randomized order, there followed two FR and two FL sessions, with 12 trials each of ILD = 0, −15, + 15 dB (first session), and ILD = 0, + 15, −15 dB (second session). A short break was provided between each session. Data were downloaded to REDCap (Harris et al., 2009, 2019) for storage and analysis.

LiSN-S

The LiSN-S task[1] (Cameron and Dillon, 2007) measures ability to attend, hear, and recall sentences in the presence of distracting sentences. LiSN-S was administered using a laptop, a task-specific soundcard, and Sennheiser HD 215 headphones. Participants

---

[1] www.LiSN-S.com

were asked to repeat a series of target sentences ("T"), presented directly in front (0°), while ignoring two distracting talkers. There were four listening conditions, in which the distractors change voice (different or same as target) and/or (virtual) position (0° and 90° relative to the listener). The test was adaptive; the level of the target speaker decreased or increased in SNR relative to the distracting talkers as the listener responded correctly or incorrectly. Testing continued for a minimum of 22 trials per condition (including five practice items that did not contribute to the score). Testing stopped when SEM < 1 or after 30 trials. The 50% correct SNR was either the "Low cue SRT" (same voice, 0° relative to the listener) or the "High cue SRT" (different voice, 90° relative to the listener). Three "derived scores" were the Talker Advantage, Spatial Advantage, and Total Advantage, so-called because each is the difference between SRTs from two conditions.

NIH Cognition Toolbox

Cognition was assessed using the NIH Toolbox – Cognition Domain battery of tests (Weintraub et al., 2013). Participants completed testing online or via iPad app in accordance with the current NIH recommendations in a private sound attenuated booth or quiet room. The battery contains up to seven different standardized cognitive instruments measuring different aspects of vocabulary, memory, attention, executive functioning, etc. The precise composition of the testing battery is dependent on participant age. Sixty five participants in this study completed the picture vocabulary test (TPVT), flanker inhibitory control and attention task (Flanker), DCCS test, and PSMT. Each test produced an age-corrected standardized score and the scores of all four tests were combined to calculate a single, Early Childhood Composite. Additional tests, contributing to the Crystallized, Fluid and Total Composite scores, were the LSWM, the PCPS, and the RR.

Toolbox picture vocabulary test is an adaptive test in which the participant is presented with an audio recording of a word and selects which of four pictures most closely matches the meaning of the word. In the Flanker, testing inhibition/attention, the participant reports over 40 trials the direction of a central visual stimulus (left or right, fish or arrow) in a string of five similar, flanking stimuli that may be congruent (same direction as target) or incongruent (opposite direction). The DCCS tests cognitive flexibility (switching attention). Target and test "card" stimuli vary along two dimensions, shape and color. Participants are asked to match test cards to the target card according to a specified dimension that varies for each trial. Both the Flanker and DCCS score accuracy and reaction time. PSMT assesses episodic memory by presenting an increasing number of illustrated objects and activities, each with a corresponding audio-recorded descriptive phrase. Picture sequences vary in length from 6 to 18 pictures depending on age, and participants are scored on the cumulative number of adjacent pairs remembered correctly over two learning trials.

Magnetic Resonance Imaging
Stimuli and Task

fMRI scanning included an active speech categorization task. Sixteen BKB sentences (Bench et al., 1979) recorded by a single male North American speaker under studio recording conditions were presented using sparse scanning procedures ("HUSH"; Schmithorst and Holland, 2004; Deshpande et al., 2016). Specifically, sentences were presented during a 6 s silent interval followed by 6 s of fMRI scanning (details below). Following methods described by Scott et al. (2000), recordings limited to < 3.8 kHz but otherwise unprocessed were delivered as "Clear" speech sentences. "Rotated" speech stimuli were created by rotating each sentence spectrally around 2 kHz using the (Blesser, 1972) technique. Rotated speech was not intelligible, though some phonetic features and some of the original intonation were preserved. "Rotated and Vocoded" speech stimuli were created by applying six-band noise-vocoding (Shannon et al., 1995) to the rotated speech stimuli. While the rotated noise-vocoded speech was completely unintelligible, the character of the envelope and some spectral detail was preserved. The listener's task was to make a button press after each sentence presentation, indicating whether a cartoon image ("human" or "alien") matched the speaker of the sentence. In familiarization trials, before scanning, the clear speech was introduced as "human" and the rotated/vocoded speech as "alien." Each participant completed three practice trials with verbal feedback from the tester. If a trial was completed incorrectly, the stimuli and instructions were reintroduced until the listener showed understanding.

Procedure

All listeners wore foam ear plugs to attenuate the scanner noise, but they were still able to hear clearly the stimuli delivered binaurally (diotically) via MR-compatible circumaural headphones. Listeners completed 48 matching trials, 16 of each sentence type, with no feedback. To maintain scanner timings, the sentence task continued regardless of whether a response was made. However, if a response was not made on three trials in a row, the tester provided reminders/encouragement over the scanner intercom between stimuli presentations.

Imaging

MRI was performed using a 3T Phillips Ingenia scanner with a 64-channel head coil and Avotec audiovisual system. The scanning protocol included a T1-weighted anatomical scan (1 mm isotropic resolution) and the fMRI task described above using a sparse acquisition approach ("HUSH"; TE = 30 ms, TR = 2000 ms, voxel size = 2.5 × 2.5 × 3.5 mm, 39 slices ascending).

Analysis
Behavioral Analysis

ECLiPS, LiSN-S, and NIH Toolbox data were separately analyzed in two-way mixed effects ANOVA, with the Group variable (TD/LiD) and within-subject variables for subtests. Separate t-tests were used to examine composite scores.

Dichotic listening data were first analyzed in a four-way mixed effects ANOVA, with the variables 2 Groups (TD/LiD) × 2 Ear (Right, Left) × 3 Attention (NF, FR, FL) × 3 ILD (0, + 15, −15), and number of correct reports as the dependent variable. The Group variable was treated as a between-group variable, while

the ear, intensity, and attention variables were treated as within-subject variables. In a second three-way ANOVA, we reduced the design to the variables Group × Attention × Intensity, and with the LI score as dependent variable. The LI score controlled for differences in overall performance between the participants, and was calculated according to the formula [(REar – LEar)/(REar + LEar) × 100].

To elucidate differences between groups in sensitivity to manipulating the physical acoustic environment of the stimuli, a third, two-way ANOVA further reduced the variables to Group × ILD, again based on the LI scores. Follow-up *post hoc* tests of main- and interaction-effects were done with Fisher's LSD test. Significance threshold was set at *p* = 0.05 for all tests.

Correlations between DL, care-giver report, spatialized listening, and cognitive function were conducted using Pearson's coefficient between age-corrected DL-LI across ILD, and ECLiPS Total Score, LiSN-S Low Cue and Talker Advantage, and NIH Toolbox Total Composite.

Imaging Analysis

First-level fMRI data were processed using FSL (FMRIB Software Library[2]). Anatomical T1 data and functional data were first reoriented using FSL's *fslreorient2std*. Next, the T1 data were brain extracted using FSL's BET. The brain extracted T1 image was then normalized and resampled to the 2 mm isotropic MNI ICBM 152 non-linear sixth generation template using FSL's FLIRT. For the functional data, the initial three time points were discarded to allow protons to reach T1 relaxation equilibrium. Slice timing correction was carried out using FSL's "*slicetimer*" and BET, respectively. Outlying functional volumes were detected using FSL's "*fsl_motion_outliers*" with the default RMS intensity difference. Cardiac and respiration signals were regressed out using AFNI's "*3dretroicor.*" Motion correction of the BOLD time-series was carried out using MCFLIRT. Motion-related artifacts were regressed from the data by setting up a general linear model (GLM) using six motion parameters. The amount of motion during the scans did not differ between groups.

Second level analysis was also conducted using FSL. A GLM approach was used to create group activation maps based on contrasts between conditions for all participants (i.e., regardless of LiD/TD status). Group composite images were thresholded using a family-wise error correction (*p* < 0.001) and clustering threshold of *k* = 4 voxels. Three BOLD activation contrasts were used as localizers responding to different aspects of speech perception (Halai et al., 2015 modified from Scott et al., 2000). First, the "Speech" activation map contrasts a signal having intelligibility, intonation, phonetics, and sound with one lacking all these attributes except sound (clear > rotated/vocoded). Second, the "Intelligibility" activation map contrasts a signal having all attributes with one retaining intonation, phonetics, and sound (clear > rotated). Third, the "Phonetics" activation map contrasts a signal having intonation, phonetics, and sound with one having only sound (rotated > rotated + vocoded).

---

[2]https://www.fsl.fmrib.ox.ac.uk/fsl/

Regions of Interest (ROI) Analysis

These three activation maps were used to identify brain regions showing significantly increased activation for speech, phonetic features, and intelligibility. These active regions were used as ROIs for correlation analysis with the DL behavioral data within which significant group differences between TD and LiD were hypothesized. Statistical analysis used JASP (v. 0.10.2) to plot data regressions and calculate correlations. Differences between correlation coefficients of each group were tested using Fisher's r-to-z transformation.
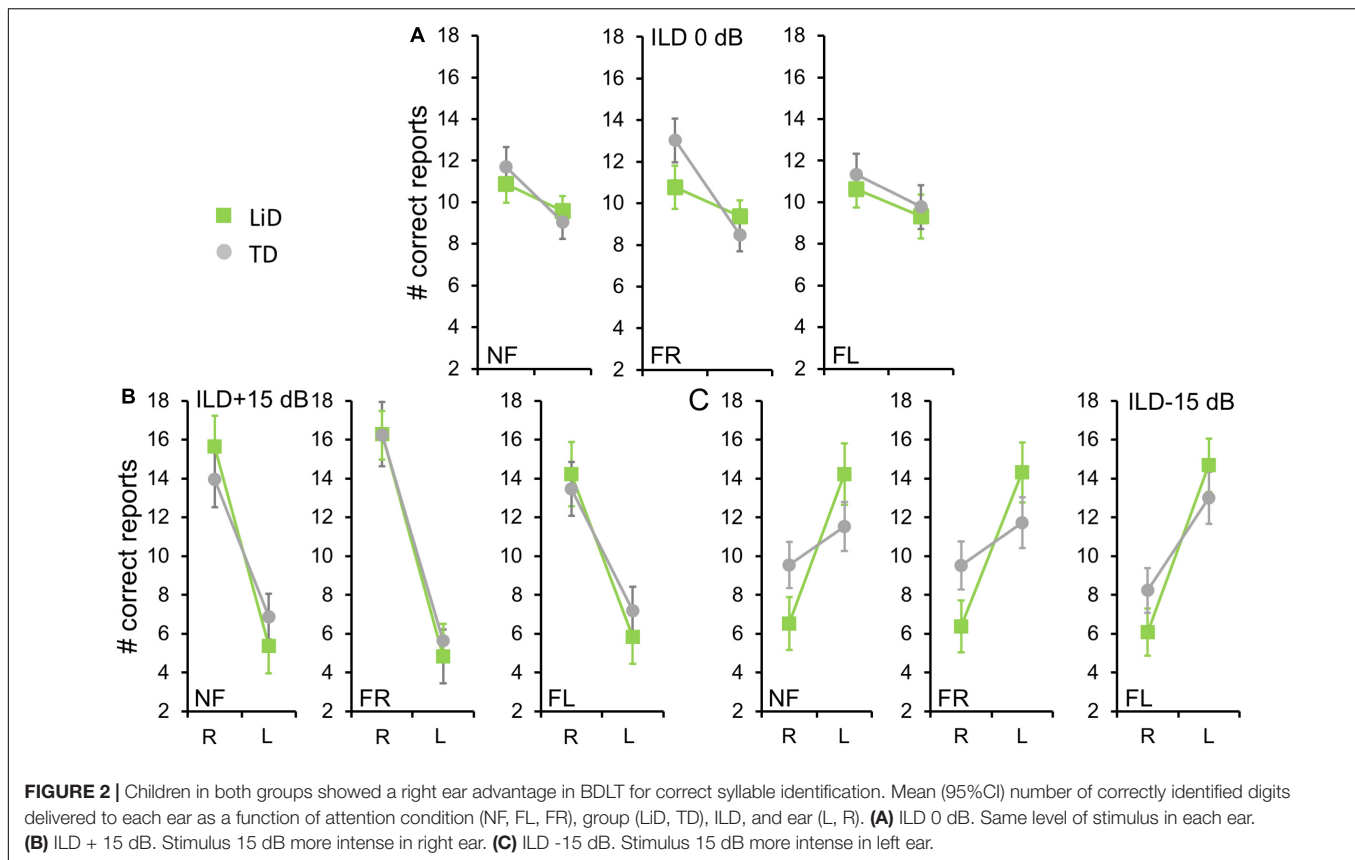
RESULTS

Audiometry and Caregiver Report

No significant difference in pure tone auditory threshold detection was found between children who were TD and those who had LiD across either the standard (**Figure 1A**) or extended (**Figure 1B**) frequency range. Children formed a continuum of listening abilities, as assessed by caregivers, but two groups, TD and LiD were segregated, primarily on their total score on the ECLiPS (**Figure 1C**). Two children in the LiD group who overlapped with the TD range of scores, and an additional 11 children with LiD had a clinical diagnosis of APD.

Dichotic Listening

Children of all ages were generally able to complete the full 216 trials of BDLT testing in about 30 min, although there was a significant attrition rate as testing continued since the task is not the most engaging and fatigue was commonplace in both groups. Participants with LiD were more likely to become frustrated or upset by the task. Frequent check-ins with the participant were needed and short breaks (a few minutes) were not uncommon. However, neither fatigue nor inattention was a basis for exclusion. Forced conditions were counterbalanced.

We first examined the BDLT results of all children in terms of number of syllables correctly identified, with a maximum possible score under each condition of 24 (12 trials × 2 blocks. 3 ILDs × 3 Attention conditions; **Figure 2**). For no ILD (ILD 0 dB), all three attention conditions (NF, FL, FR) showed a significant REA in both groups (**Figure 2A**) but there was no significant difference between attention condition. That REA became larger for ILD + 15 dB (**Figure 2B**) and reversed for ILD -15 dB (**Figure 2C**), all as expected from the DL literature, except that a REA was obtained even in the FL condition at ILD 0 dB. For the ILD -15 dB condition, it appeared that the ear differences were smaller for the TD than for the LiD group. An overall four-way ANOVA was first run with the factor "Age" as covariate to control for the small, non-significant age difference between the groups (see below). This analysis showed a significant three-way interaction of ILD by ear by group: $F(2,142) = 5.70$, $p = 0.004$, partial eta^2 = 0.07. The interaction was followed-up with Tukey's HSD test which showed that, while both groups were able to shift to a significant left ear advantage during the ILD -15 dB condition,

FIGURE 2 | Children in both groups showed a right ear advantage in BDLT for correct syllable identification. Mean (95%CI) number of correctly identified digits delivered to each ear as a function of attention condition (NF, FL, FR), group (LiD, TD), ILD, and ear (L, R). **(A)** ILD 0 dB. Same level of stimulus in each ear. **(B)** ILD + 15 dB. Stimulus 15 dB more intense in right ear. **(C)** ILD -15 dB. Stimulus 15 dB more intense in left ear.

this ability was exaggerated in the LiD group, controlling for multiple comparisons.

To investigate group differences further, we next examined the LI (**Figure 3A**). Three-way ANOVA showed a significant effect of ILD: $F(2,142) = 4.45$, $p = 0.013$, partial eta^2 = 0.06. There was a significant two-way interaction of ILD by group (**Figure 3B**): $F(2,142) = 6.87$, $p = 0.001$, partial eta^2 = 0.08. Tukey's HSD test showed significantly higher Laterality in the LiD group in the ILD −15 dB condition, controlling for multiple comparisons. Also shown on **Figure 3B** are typical young adult data (NF condition) from the study of Westerhausen et al. (2009). At ILD 0 dB, the LI (REA) was smallest for the LiD group (6%), larger for the TD group (13%), and largest for the Adults (28%). Note that Westerhausen's adult data were near parallel with the LiD data, but that the LiD data showed a stronger left ear influence at each ILD. Asymmetry of LI between the ILD ± 15 dB was more marked for the TD than the LiD group, with TD children, like adults, showing a much larger LI for ILDs favoring the right ear. By contrast, children with LiD had larger but near symmetric LIs for ILD ± 15 dB. Both groups of children showed different immature response patterns. Of the 35 children with LiD who completed DL testing, 22 had been evaluated for, and nine had a diagnosis of APD. None of the means of their DL scores (reports and LI) differed significantly from that of the 26 children in the LiD group without an APD diagnosis (27 independent samples t-tests; $p = 0.27$–0.99). LIs for three age groups, with both TD and LiD children together and divided approximately to equalize

the number of children in each group, are shown in **Figure 4**. As above, small differences were seen between the age groups, with older children overall having slightly larger unsigned LIs than younger children, although not significant, as indicated from three-way ANOVA with age × attention × ILD as variables. Note the positive LIs in the ILD 0 dB FL condition.

In summary, neither group, age, nor attention condition affected the LI of right/left recall. However, a significant interaction was found between group (LiD, TD) and ILD. Children with LiD were more influenced by large ILDs, especially favoring the left ear, than were TD children and were thus less able to modulate performance through attention, and more driven by the physical properties of the acoustic stimuli.

Auditory Perceptual and Cognitive Function

Listening to sentences in "spatialized" noise (LiSN-S) was significantly ($p < 0.01$) impaired in children with LiD on the Low Cue and High Cue conditions, and the derived Talker Advantage measure (**Figure 5A**). This pattern of results suggested that the children with LiD had problems with both the procedural demands and the specifically auditory demands of the task.

Related to the disability of children with LiD to perform the listening task (LiSN-S), we found they also had impaired performance on all subtests of the NIH Toolbox, summarized
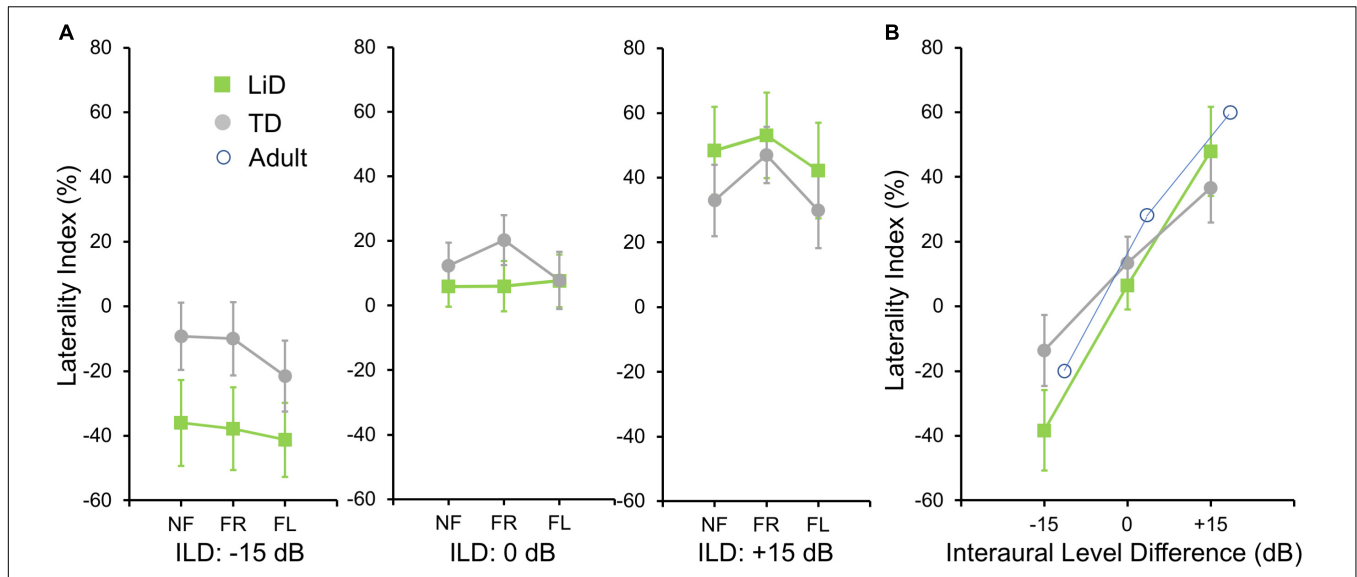
FIGURE 3 | BDLT laterality index varied more in LiD group than in TD group as a function of ILD. **(A)** Same comparisons as in **Figure 2** expressed as mean (95% CI) percentage correct responses right ear relative to left ear (see text). **(B)** Mean (± 95% CI) LI as a function of ILD averaged across attention conditions in each group. Adult data from Westerhausen et al. (2009).
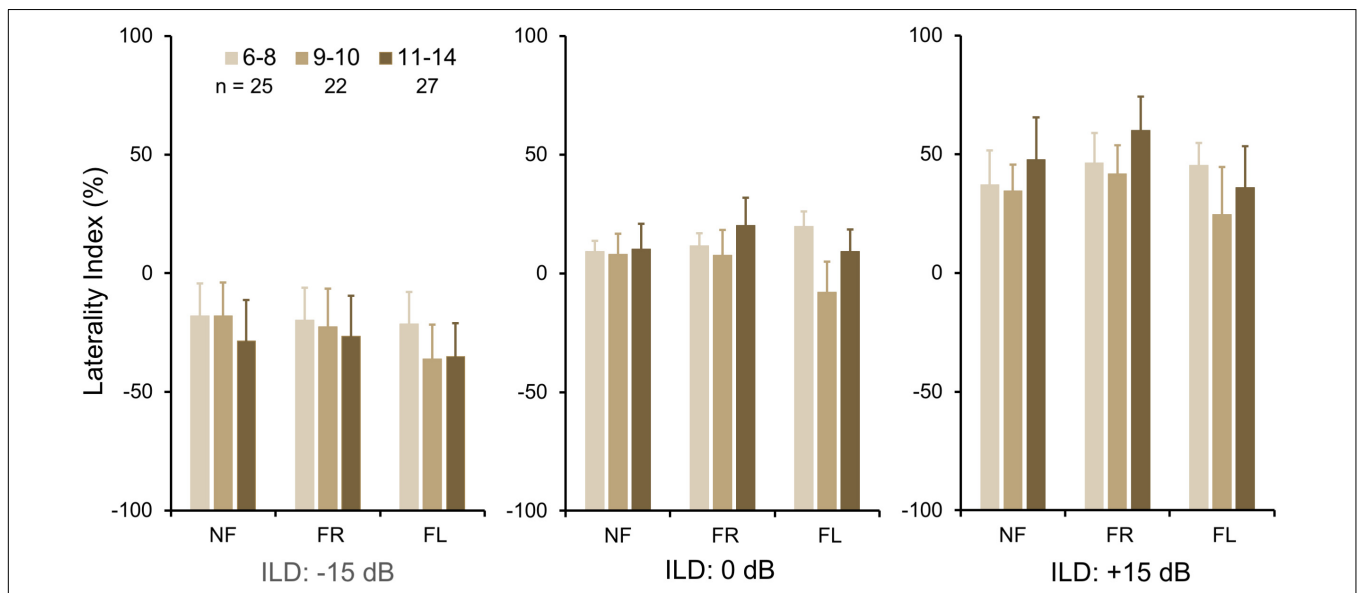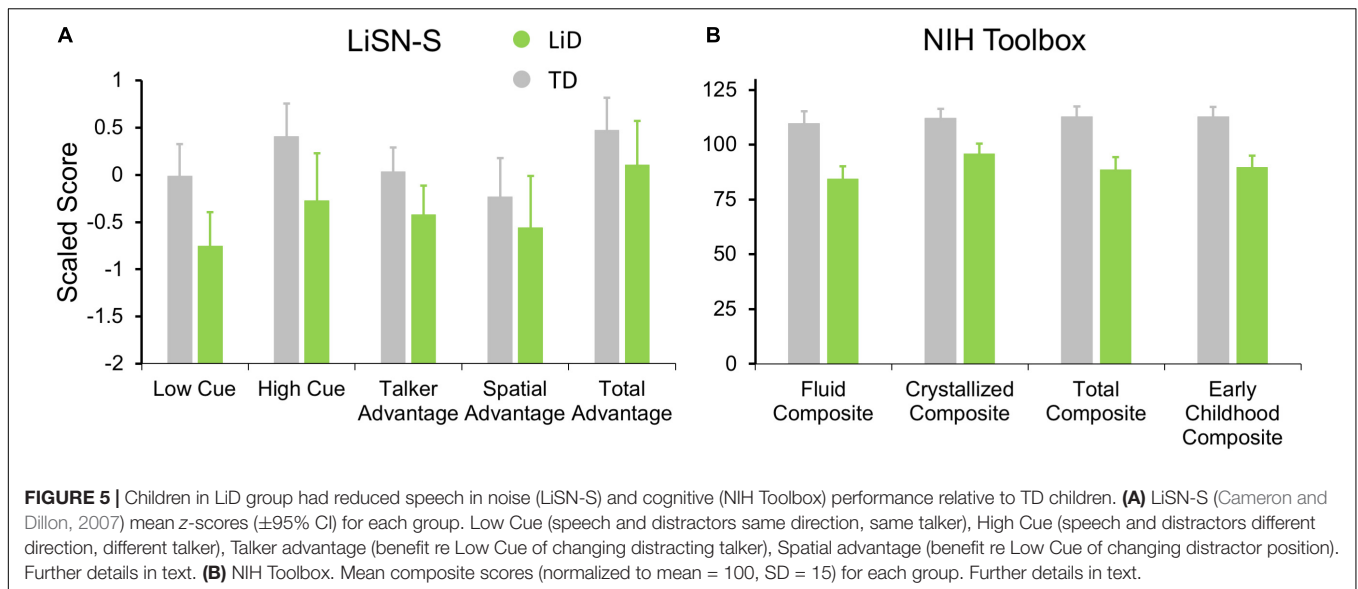

FIGURE 4 | BDLT laterality index increased slightly, but non-significantly with age. Same comparisons as in **Figures 2**, **3**; mean (95% CI) LI averaged across groups. Note that REA did not increase significantly across this age range.

in **Figure 5B** (all $p < 0.001$). The mean standard score of the children with LiD was poorest on the Fluid Composite, composed of the visually based NIH Toolbox DCCS, Flanker, PSM, LSWM, and PCPS subtests. Performance was also significantly impaired on the Crystallized Composite, consisting of the PV and RR subtests. The PV was the only subtest on which success was partly dependent on auditory perception and receptive language function. However, results for the PV subtest (mean difference between LiD and TD groups = 15.9 points) were similar to those of the RR subtest (16.0 points). It therefore appears that the LiD

group had a generalized, multi-modal mild cognitive impairment relative to the TD group.

Correlations Between Behavioral Measures

Few significant correlations were observed between the ECLiPS, the LiSN-S or the NIH Toolbox data, and DL-LI measures. From a total of 99 comparisons, only nine LiSN-S and Toolbox measures were significant at $p < 0.01$, uncorrected for multiple

FIGURE 5 | Children in LiD group had reduced speech in noise (LiSN-S) and cognitive (NIH Toolbox) performance relative to TD children. **(A)** LiSN-S (Cameron and Dillon, 2007) mean z-scores (±95% CI) for each group. Low Cue (speech and distractors same direction, same talker), High Cue (speech and distractors different direction, different talker), Talker advantage (benefit re Low Cue of changing distracting talker), Spatial advantage (benefit re Low Cue of changing distractor position). Further details in text. **(B)** NIH Toolbox. Mean composite scores (normalized to mean = 100, SD = 15) for each group. Further details in text.

comparisons. For the ECLiPS (not shown), only 3/9 comparisons were significant at $p < 0.05$, and all three comparisons were for ILD −15 dB, at which LIs and differences between groups were largest (**Figure 3B**). Similar patterns were seen for the LiSN-S and the Toolbox (**Figure 6**). Correlations between LiSN-S Spatial and Talker Advantage with LI during FR −15 and FL + 15 were just significant, after Bonferroni adjustment (two of 18 comparisons, $p \leq 0.006$). Toolbox Composite data showed some strong and consistent correlations (**Figure 6B**). For example, the Fluid Composite was significantly associated with LI ($p < 0.001$) for FR −15; all four Toolbox measures showed low cognitive performance associated with strongly negative LI.

fMRI

All children performed well in the scanner on the active speech categorization task, although the TD children performed more accurately, and with shorter reaction times, than those with LiD (**Figure 7**).

Neural activity associated with listening to the Speech, Phonetic, and Intelligibility of the sentences did not significantly differ between the two groups. BOLD activation from across all children (regardless of group) is shown in **Figure 8**.

Activation patterns for Speech included bilateral auditory cortices (middle temporal gyrus, superior temporal gyrus, and Heschl's gyrus), PT, left temporal fusiform cortex, inferior temporal gyrus, OFC, and right parahippocampal gyrus. In the left OFC, a significant correlation (r) was observed among the TD children between BOLD activation and the BDLT FR0 attention condition (**Figure 9A**). In contrast, children with LiD lacked such a correlation.

Phonetic activation was seen in a more restricted region of posterior AC in the STG and PT bilaterally and right Heschl's gyrus (**Figure 8B**). Left PT activation was correlated with the BDLT LI in TD children for the NF0 attention condition (**Figure 9B**). For children with LiD, the relationship flipped, but not to the extent of a significant correlation.
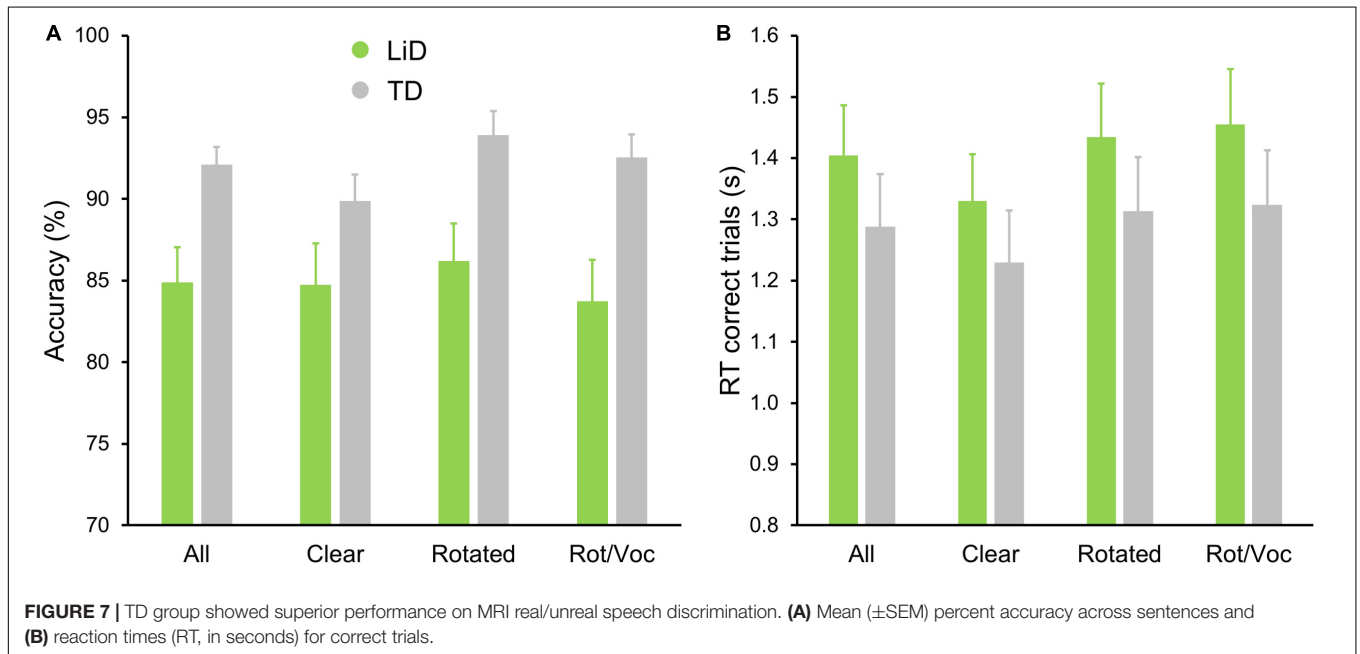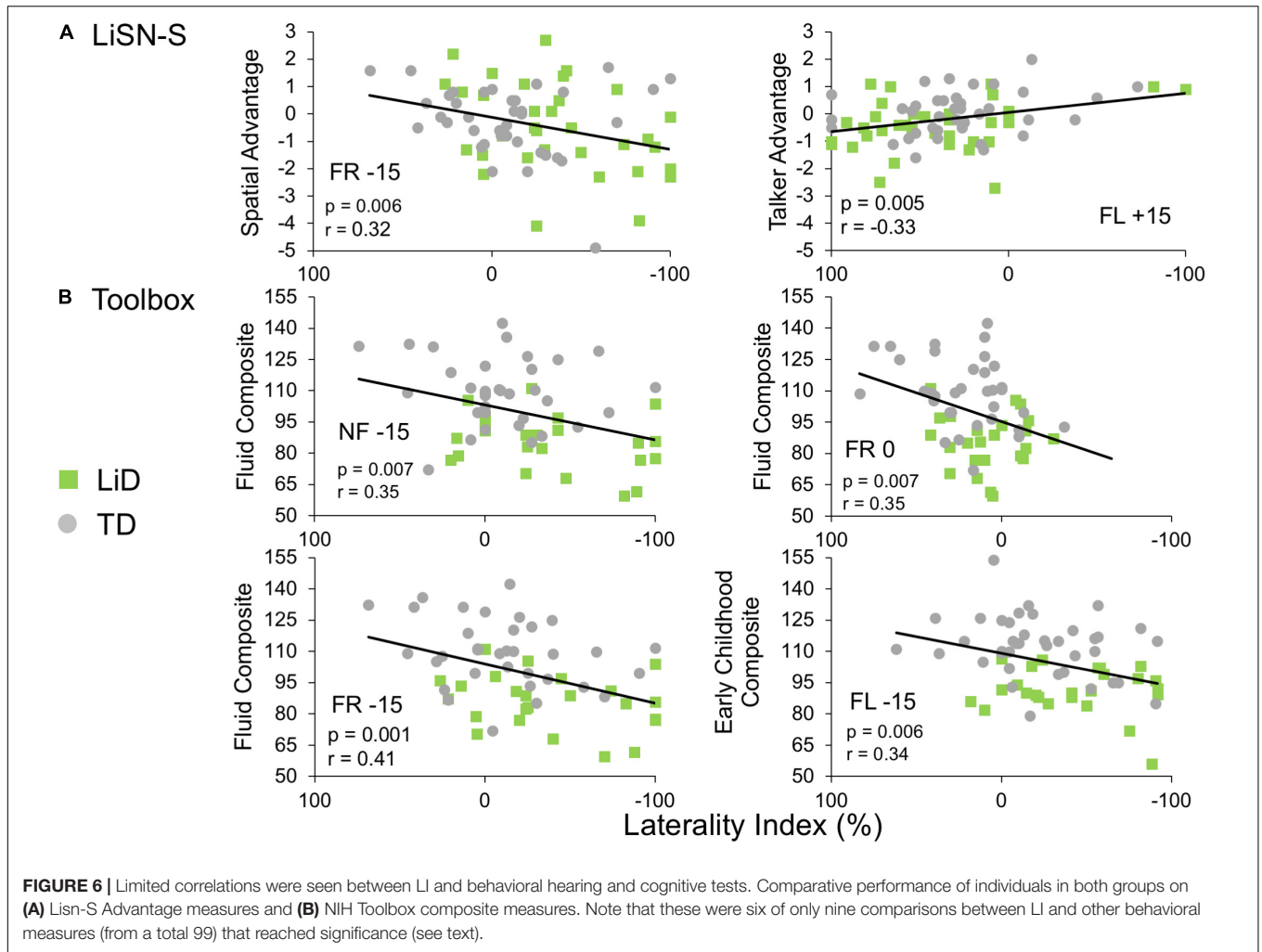
The Intelligibility contrast revealed increased activation in the superior temporal gyrus, with a long anterior to posterior profile from the left temporal pole and along the STG and a more anterior temporal pole locus on the right (**Figure 8C**). Significant correlations with BDLT LI were seen with the left IFG under both NF0 and NF15 conditions in the TD children. Again, no such correlations were observed in the children with LiD (**Figures 9C,D**).
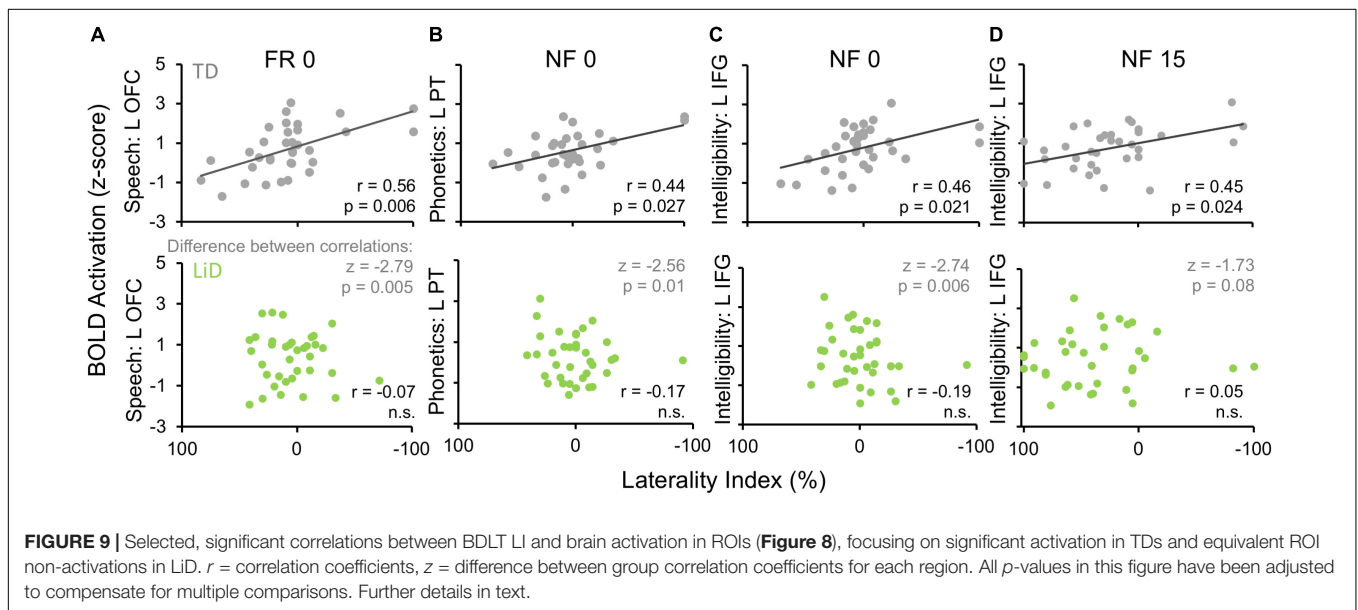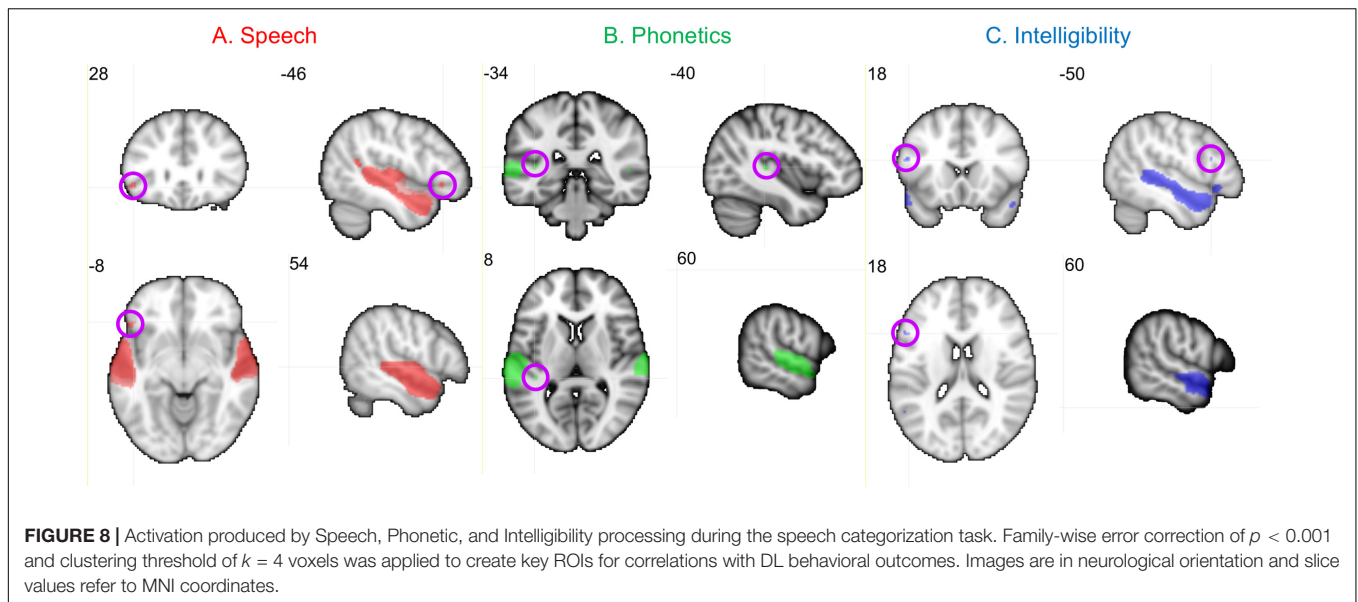
The difference between the two groups in correlations of brain activation with attention conditions was significant (z) for speech-FR0 ($p = 0.005$), phonetics-NF0 ($p = 0.01$), and intelligibility-NF0 ($p = 0.006$) but not intelligibility-NF15 ($p = 0.006$).

DISCUSSION AND CONCLUSION

Listening Difficulties

Children with LiD performed normally in the BDLT Listen (NF) mode, as hypothesized, but they also performed normally in the BDLT Concentrate (FL, FR) modes, despite significantly impaired performance on speech-in-noise and cognitive tests. This is a new finding since most previous research on auditory processing differences between TD and non-TD children has focused on specific impairments in processing capacity of the left hemisphere, reflected in differential scores in the FL mode (Westerhausen and Hugdahl, 2010). The normal performance of children with LiD in the ILD 0 dB condition suggests that their cognitive insufficiency did not prevent them performing the DL task. Moreover, no significant differences were found between the groups on the right ear or left ear scores, suggesting no systematic hemispheric processing differences. Rather, the children with LiD were found to have a generalized disability to benefit from ILDs between the dichotic stimuli. The small REA found in both groups is consistent with weak REAs reported in a previous study of young children (Passow et al., 2013).

FIGURE 6 | Limited correlations were seen between LI and behavioral hearing and cognitive tests. Comparative performance of individuals in both groups on **(A)** Lisn-S Advantage measures and **(B)** NIH Toolbox composite measures. Note that these were six of only nine comparisons between LI and other behavioral measures (from a total 99) that reached significance (see text).

FIGURE 7 | TD group showed superior performance on MRI real/unreal speech discrimination. **(A)** Mean (±SEM) percent accuracy across sentences and **(B)** reaction times (RT, in seconds) for correct trials.

FIGURE 8 | Activation produced by Speech, Phonetic, and Intelligibility processing during the speech categorization task. Family-wise error correction of $p < 0.001$ and clustering threshold of $k = 4$ voxels was applied to create key ROIs for correlations with DL behavioral outcomes. Images are in neurological orientation and slice values refer to MNI coordinates.

FIGURE 9 | Selected, significant correlations between BDLT LI and brain activation in ROIs (**Figure 8**), focusing on significant activation in TDs and equivalent ROI non-activations in LiD. r = correlation coefficients, z = difference between group correlation coefficients for each region. All p-values in this figure have been adjusted to compensate for multiple comparisons. Further details in text.

Performance on BDLT of children with LiD was more affected by varying ILD than was that of TD children. This could be because the children with LiD had a primary auditory problem, or that they were less able to offset greater sound level at either ear through attention modulation (Westerhausen and Hugdahl, 2010). Poor LiSN-S performance, particularly on the spatial advantage measure, may indicate a binaural interaction problem (Cameron and Dillon, 2008; Glyde et al., 2013; Cameron et al., 2014). Correlations between LI and LiSN-S advantage measures at high ILD support this interpretation, but LiD and TD groups did not differ in this respect. Inattention is a primary symptom of LiD, although its relationship to APD is controversial (Moore et al., 2010; Moore, 2018). However, there seems to be general agreement that many if not most children undergoing APD evaluation have attention difficulties

that, at least, need to be taken into account by the examining audiologist (American Academy of Audiology, 2010; Sharma et al., 2014; Moore et al., 2018).

Age

There have been few studies of children using the BDLT. Age effects in this study generally mirrored those previously reported (Hugdahl et al., 2001; Passow et al., 2013), although the REA tended to get larger with increasing age, in contrast to a recent, NF-only study (Kelley and Littenberg, 2019). In fact, comparison between current LiD data and adult data of Westerhausen et al. (2009) suggested a robust, consistent increase in right ear influence with age, across ILD, supporting a "right ear weakness" hypothesis for LiD. This contrasted with TD children who had more of a "REA amplification" pattern of development where

changes in LI with increasing age were asymmetric between left- and right-leading ILDs.

fMRI

Both groups of children (LiD, TD) used the same brain areas to perform the sentence-listening task but relations between brain area activation and BDLT LI suggested the areas are used differently by the two groups. Left OFC was related to LI forced attention for TD but not for LiD. Other areas (left PT and IFG) also related to LI activation in TD, but not LiD. These results were all somewhat independent of ILD or task type (NF or forced). As outlined in the Introduction, it was predicted that, if the LiD children have an auditory processing deficit, we would find similar relationships between cortical activity and behavioral results for the BDLT attention conditions (i.e., FL and FR) in both groups, and different relationships between groups for the intensity manipulations in the BDLT (i.e., −15, 0, and 15). The reverse would suggest the LiD children have language processing deficits.

A lack of group differences in BOLD activations for the sentence-listening task in the MRI scanner suggests that the groups of children do not differ in the brain areas used to process auditorily presented sentences. However, the relationships between these brain areas and BDLT laterality suggest that these brain areas are used differently by each group. In the TD group, activation in a specific cortical area used for top-down processing of speech (left OFC) was related to degree of laterality on a BDLT forced attention condition. However, this relationship was not found in the LiD group. Similarly, activation in areas used for bottom-up integration (sound, phonetics, and intonation in the left PT and IFG) was related to laterality of NF attention at different intensity levels in the TD group, but not in the LiD group. Lesser relationships were observed in other areas, with an overall pattern of some limited correlations with laterality among the TD group and lack of correlation in the LiD group.

Different group relationships were found between cortical activity and BDLT behavioral results for the attention conditions and also in the intensity manipulations. This suggests that the LiD children do not have a clear pattern of cortical reorganization associated with auditory processing. These results do not indicate a redistribution of cortical listening areas in children with LiD but, instead, a reorganization as to how these areas are engaged during language listening. Specifically, TD children showed a pattern of higher engagement of specific cortical listening areas used to support better listening task performance. This pattern was not observed in children with LiD.

Implications for Listening Difficulties in Children

Clinical use of dichotic assessment for APD in children has mostly used dichotic digits (Musiek, 1983; Emanuel et al., 2011). As discussed in Section "Introduction," the results of that assessment do not distinguish between an auditory and a cognitive explanation of those children's LiD, although the test may correctly identify children without hearing loss as having an auditory perceptual or speech coding problem. In other

studies, using the less cognitively but more auditorily demanding BDLT, older children and adults with a wide variety of learning, neurological, and mental health diagnoses had a generally weak left ear performance in the FL condition (Westerhausen and Hugdahl, 2010). This was interpreted as a means for testing top-down executive function that we found here, in the Toolbox data, to be impaired in children with LiD. However, we did not find a consistent poor performance on the FL task in that group.

A number of observations have been made about dichotic ear advantages in children with APD (Moncrieff et al., 2017). Some studies have focused on the prevalence of a REA or LEA, suggesting balance between the two is initially more even but unstable, but that a consistent REA emerges with age (Moncrieff, 2011). However, the current results, and others (Hugdahl et al., 2001) show that the absolute level of LI increases (i.e., larger LEA and REA) with age, contrary to the report of Moncrieff (2011), and that use of a binary LEA/REA distinction can be misleading. Other results in adults have shown that larger LIs, either positive or negative, are associated with better accuracy on the BLDT (Hirnstein et al., 2014).

A new term, "amblyaudia" was introduced by Moncrieff et al. (2016) to designate "an abnormally large asymmetry between the two ears during DL tasks with either normal or below normal performance in the dominant ear." The results of the study reported here only partially supported amblyaudia in the children with LiD. Their performance was statistically indistinguishable from that of the TD children at ILD 0 dB, the usual condition for testing. But, consistent with the definition, there was a larger than normal asymmetry in the ILD 15 dB conditions, with normal or below normal number of correctly reported digits in the right ear and a greater than normal number of correct reports in the left ear. In a review, Whitton and Polley (2011) discussed amblyaudia in the context of long-term effects of conductive hearing loss on auditory system plasticity, induced in children predominantly by otitis media. While building a convincing case from the literature for such plasticity, the relevance of such findings to the children in this study is unclear; a similar proportion of children in each group had a history of PE tubes (Hunter et al., 2020).

Several forms of dichotic training have been proposed, and we know of at least two that are in current evaluation or practice for the treatment of amblyaudia (DIID—Musiek, 2004; ARIA—Moncrieff et al., 2017) and other abnormalities detected through DL evaluation (Emanuel et al., 2011). Unfortunately, it remains unclear what sort of benefits might be obtained from such training or whether any of the proposed methods generalize to improved listening in real-world challenging environments. Hugdahl et al. (2009) present several arguments *against* training using the BDLT to treat impaired performance on the FL instruction task. These arguments are, briefly, that the DL task is very simple and therefore unchallenging, that it shows little or no learning effect, and that executive functions are not amenable to training. It is unclear what form of training might be useful for normalizing specific DL behavior patterns of the children with LiD in the current study. However, interventions that improve auditory attention should be generally useful for these children.

In summary, we found little evidence for impaired DL on the BDLT or brain activation differences for children with LiD

compared with TD children. A significant reduction of the LI was found in the LiD group when the left ear stimulus was presented at a reduced level compared with the right ear stimulus. Brain activation was correlated with LI in some frontal and temporal regions for children in the TD group, but not for those in the LiD group.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Cincinnati Children's Institutional Review Board (IRB). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

REFERENCES

American Academy of Audiology (2010). *Clinical Practice Guidelines: Diagnosis, Treatment, and Management of Children and Adults With Central Auditory Processing Disorder*. Available at: https://audiology-web.s3.amazonaws.com/migrated/CAPD%20Guidelines%208-2010.pdf_539952af956c79.73897613.pdf (accessed December16, 2019).

Barry, J. G., and Moore, D. R. (2015). *Evaluation of Children's Listening and Processing Skills (ECLIPS)*. London: Mrc-T Publishing.

Barry, J. G., Tomlin, D., Moore, D. R., and Dillon, H. (2015). Use of questionnaire-based measures in the assessment of listening difficulties in school-aged children. *Ear Hear.* 36:e00300-13.

Bench, J., Kowal, A., and Bamford, J. (1979). The Bkb (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *Br. J. Audiol.* 13, 108–112. doi: 10.3109/03005367909078884

Blesser, B. (1972). Speech perception under conditions of spectral transformation. I. Phonetic characteristics. *J. Speech Hear. Res.* 15, 5–41. doi: 10.1044/jshr.1501.05

Brenneman, L., Cash, E., Chermak, G. D., Guenette, L., Masters, G., Musiek, F. E., et al. (2017). The relationship between central auditory processing, language, and cognition in children being evaluated for central auditory processing disorder. *J. Am. Acad. Audiol.* 28, 758–769. doi: 10.3766/jaaa.16119

Broadbent, D. E. (1956). Successive responses to simultaneous stimuli. *Q. J. Exp. Psychol.* 8, 145–152. doi: 10.1080/17470215608416814

Cameron, S., and Dillon, H. (2007). Development of the listening in spatialized noise-sentences test (Lisn-S). *Ear Hear.* 28, 196–211. doi: 10.1097/aud.0b013e318031267f

Cameron, S., and Dillon, H. (2008). The listening in spatialized noise-sentences test (LISN-S): comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *J. Am. Acad. Audiol.* 19, 377–391. doi: 10.3766/jaaa.19.5.2

Cameron, S., and Dillon, H. (2020). Are "Dichotic" deficits uniquely dichotic? investigating dichotic performance with the dichotic digits difference test (DddT) in a large clinical population of children referred for an auditory processing assessment. *J. Am. Acad. Audiol.* doi: 10.376/jaaa.19037 [Epub ahead of print].

Cameron, S., Dillon, H., Glyde, H., Kanthan, S., and Kania, A. (2014). Prevalence and remediation of spatial processing disorder (SPD) in Indigenous children in regional Australia. *Int. J. Audiol.* 53, 326–335. doi: 10.3109/14992027.2013.871388

Cameron, S., Glyde, H., Dillon, H., Whitfield, J., and Seymour, J. (2016). The dichotic digits difference test (DddT): development, normative data, and test-retest reliability studies part 1. *J. Am. Acad. Audiol.* 27, 458–469. doi: 10.3766/jaaa.15084

de Wit, E., Visser-Bochane, M. I., Steenbergen, B., Van Dijk, P., Van Der Schans, C. P., and Luinge, M. R. (2016). Characteristics of auditory processing disorders: a systematic review. *J. Speech Lang. Hear Res.* 59, 384–413.

Deshpande, A. K., Tan, L., Lu, L. J., Altaye, M., and Holland, S. K. (2016). fmri as a preimplant objective tool to predict postimplant oral language outcomes in children with cochlear implants. *Ear Hear.* 37, e263–e272. doi: 10.1097/aud.0000000000000259

Emanuel, D. C., Ficca, K. N., and Korczak, P. (2011). Survey of the diagnosis and management of auditory processing disorder. *Am. J. Audiol.* 20, 48–60. doi: 10.1044/1059-0889(2011/10-0019)

Glyde, H., Buchholz, J. M., Dillon, H., Cameron, S., and Hickson, L. (2013). The importance of interaural time differences and level differences in spatial release from masking. *J. Acoust. Soc. Am.* 134, EL147–EL152. doi: 10.1121/1.4812441

Halai, A. D., Parkes, L. M., and Welbourne, S. R. (2015). Dual-echo fMRI can detect activations in inferior temporal lobe during intelligible speech comprehension. *Neuroimage* 122, 214–221. doi: 10.1016/j.neuroimage.2015.05.067

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., et al. (2019). The Redcap consortium: Building an international community of software platform partners. *J. Biomed. Inform.* 95:103208. doi: 10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J. G. (2009). Research electronic data capture (Redcap)–a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* 42, 377–381. doi: 10.1016/j.jbi.2008.08.010

Harris, V. L., Keith, R. W., and Novak, K. K. (1983). Relationship between two dichotic listening tests and the Token test for children. *Ear Hear.* 4, 278–282. doi: 10.1097/00003446-198311000-00003

Hind, S. E., Haines-Bazrafshan, R., Benton, C. L., Brassington, W., Towle, B., and Moore, D. R. (2011). Prevalence of clinical referrals having hearing thresholds within normal limits. *Int. J. Audiol.* 50, 708–716. doi: 10.3109/14992027.2011.582049

Hirnstein, M., Hugdahl, K., and Hausmann, M. (2014). How brain asymmetry relates to performance - a large-scale dichotic listening study. *Front. Psychol.* 4:997. doi: 10.3389/fpsyg.2013.00997

Hugdahl, K., Bronnick, K., Kyllingsbaek, S., Law, I., Gade, A., and Paulson, O. B. (1999). Brain activation during dichotic presentations of consonant-vowel and musical instrument stimuli: a 15O-Pet study. *Neuropsychologia* 37, 431–440. doi: 10.1016/s0028-3932(98)00101-8

Hugdahl, K., Carlsson, G., and Eichele, T. (2001). Age effects in dichotic listening to consonant-vowel syllables: interactions with attention. *Dev. Neuropsychol.* 20, 445–457. doi: 10.1207/s15326942dn2001_8

Hugdahl, K., and Westerhausen, R. (2016). Speech processing asymmetry revealed by dichotic listening and functional brain imaging. *Neuropsychologia* 93, 466–481. doi: 10.1016/j.neuropsychologia.2015.12.011

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., and Hamalainen, H. (2008). The effect of stimulus intensity on the right ear advantage in dichotic listening. *Neurosci. Lett.* 431, 90–94. doi: 10.1016/j.neulet.2007.11.046

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., Laine, M., and Hamalainen, H. (2009). Attention and cognitive control: unfolding the dichotic listening story. *Scand. J. Psychol.* 50, 11–22. doi: 10.1111/j.1467-9450.2008.00676.x

Hunter, L. L., Blankenship, C. M., Sloat, N. T., Perdew, A., Stewart, H. J., and Moore, D. R. (2020). Peripheral auditory involvement in childhood listening difficulty. *Ear Hear.* (in press).

Keith, R. W. (2009). *Scan−3:C Tests for Auditory Processing Disorders for Children.* London: Pearson.

Keith, R. W., Tektas, M., Ramsay, K., and Delaney, S. (2019). Development and standardization of the university of cincinnati auditory processing inventory (UCAPI). *Int. J. Audiol.* 58, 373–378. doi: 10.1080/14992027.2019.1585973

Kelley, K. S., and Littenberg, B. (2019). Dichotic listening test-retest reliability in Children. *J. Speech Lang. Hear. Res.* 62, 169–176. doi: 10.1044/2018_jslhr-h-17-0158

Kimura, D. (1961). Cerebral dominance and the perception of verbal stimuli. *Can. J. Psychol.* 15, 166–171. doi: 10.1037/h0083219

Kimura, D. (1963). Speech lateralization in young children as determined by an auditory test. *J. Comp. Physiol. Psychol.* 56, 899–902. doi: 10.1037/h0047762

Kompus, K., Specht, K., Ersland, L., Juvodden, H. T., Van Wageningen, H., Hugdahl, K., et al. (2012). A forced-attention dichotic listening fmri study on 113 subjects. *Brain Lang.* 121, 240–247. doi: 10.1016/j.bandl.2012.03.004

Moncrieff, D., Keith, W., Abramson, M., and Swann, A. (2016). Diagnosis of amblyaudia in children referred for auditory processing assessment. *Int. J. Audiol.* 55, 333–345. doi: 10.3109/14992027.2015.1128003

Moncrieff, D., Keith, W., Abramson, M., and Swann, A. (2017). Evidence of binaural integration benefits following Aria training for children and adolescents diagnosed with amblyaudia. *Int. J. Audiol.* 56, 580–588. doi: 10.1080/14992027.2017.1303199

Moncrieff, D. W. (2011). Dichotic listening in children: age-related changes in direction and magnitude of ear advantage. *Brain Cogn.* 76, 316–322. doi: 10.1016/j.bandc.2011.03.013

Moore, D. R. (2018). Editorial: auditory processing disorder. *Ear Hear.* 39, 617–620. doi: 10.1097/aud.0000000000000582

Moore, D. R., Ferguson, M. A., Edmondson-Jones, A. M., Ratib, S., and Riley, A. (2010). Nature of auditory processing disorder in children. *Pediatrics* 126, e382–e390.

Moore, D. R., and Hunter, L. L. (2013). Auditory processing disorder (APD) in children: a marker of neurodevelopmental syndrome. *Hear. Balance Commun.* 11, 160–167. doi: 10.3109/21695717.2013.821756

Moore, D. R., Hutchings, M. E., and Meyer, S. E. (1991). Binaural masking level differences in children with a history of otitis media. *Audiology* 30, 91–101. doi: 10.3109/00206099109072874

Moore, D. R., Sieswerda, S. L., Grainger, M. M., Bowling, A., Smith, N., Perdew, A., et al. (2018). Referral and diagnosis of developmental auditory processing disorder in a large, United States hospital-based audiology service. *J. Am. Acad. Audiol.* 29, 364–377. doi: 10.3766/jaaa.16130

Musiek, F. E. (1983). Assessment of central auditory dysfunction: the dichotic digit test revisited. *Ear Hear.* 4, 79–83. doi: 10.1097/00003446-198303000-00002

Musiek, F. E. (2004). The Diid: a new treatment for Apd. *Hear. J.* 57:50. doi: 10.1097/01.hj.0000293049.80297.cd

Passow, S., Muller, M., Westerhausen, R., Hugdahl, K., Wartenburger, I., Heekeren, H. R., et al. (2013). Development of attentional control of verbal auditory perception from middle to late childhood: comparisons to healthy aging. *Dev. Psychol.* 49, 1982–1993. doi: 10.1037/a0031207

Pillsbury, H. C., Grose, J. H., and Hall, J. W. III (1991). Otitis media with effusion in children. Binaural hearing before and after corrective surgery.

Arch. Otolaryngol. Head Neck Surg. 117, 718–723. doi: 10.1001/archotol.1991.01870190030008

Rees, A., and Palmer, A. R. (2010). *The Auditory Brain.* Oxford: Oxford University Press.

Rimol, L. M., Specht, K., Weis, S., Savoy, R., and Hugdahl, K. (2005). Processing of sub-syllabic speech units in the posterior temporal lobe: an fmri study. *Neuroimage* 26, 1059–1067. doi: 10.1016/j.neuroimage.2005.03.028

Roebuck, H., and Barry, J. G. (2018). Parental perception of listening difficulties: an interaction between weaknesses in language processing and ability to sustain attention. *Sci. Rep.* 8:6985.

Roeser, R. J., Millay, K. K., and Morrow, J. M. (1983). Dichotic consonant-vowel (Cv) perception in normal and learning-impaired children. *Ear Hear.* 4, 293–299. doi: 10.1097/00003446-198311000-00006

Schmithorst, V. J., and Holland, S. K. (2004). Event-related fmri technique for auditory processing with hemodynamics unrelated to acoustic gradient noise. *Magn. Reson. Med.* 51, 399–402. doi: 10.1002/mrm.10706

Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2000). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123(Pt 12), 2400–2406. doi: 10.1093/brain/123.12.2400

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303

Sharma, M., Dhamani, I., Leung, J., and Carlile, S. (2014). Attention, memory, and auditory processing in 10- to 15-year-old children with listening difficulties. *J. Speech Lang. Hear. Res.* 57, 2308–2321. doi: 10.1044/2014_jslhr-h-13-0226

Weihing, J., Guenette, L., Chermak, G., Brown, M., Ceruti, J., Fitzgerald, K., et al. (2015). Characteristics of pediatric performance on a test battery commonly used in the diagnosis of central auditory processing disorder. *J. Am. Acad. Audiol.* 26, 652–669. doi: 10.3766/jaaa.14108

Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., et al. (2013). Cognition assessment using the Nih Toolbox. *Neurology* 80, S54–S64.

Westerhausen, R., and Hugdahl, K. (2010). "Cognitive control of auditory laterality," in *The Two Halves of the Brain: Information Processing in the Cerebral Hemispheres*, eds K. Hugdahl and R. Westerhausen (Cambridge: Mit Press).

Westerhausen, R., Moosmann, M., Alho, K., Belsby, S. O., Hamalainen, H., Medvedev, S., et al. (2010). Identification of attention and cognitive control networks in a parametric auditory fmri study. *Neuropsychologia* 48, 2075–2081. doi: 10.1016/j.neuropsychologia.2010.03.028

Westerhausen, R., Moosmann, M., Alho, K., Medvedev, S., Hamalainen, H., and Hugdahl, K. (2009). Top-down and bottom-up interaction: manipulating the dichotic listening ear advantage. *Brain Res.* 1250, 183–189. doi: 10.1016/j.brainres.2008.10.070

Whitton, J. P., and Polley, D. B. (2011). Evaluating the perceptual and pathophysiological consequences of auditory deprivation in early postnatal life: a comparison of basic and clinical studies. *J. Assoc. Res. Otolaryngol.* 12, 535–547. doi: 10.1007/s10162-011-0271-6

Wilson, W. J. (2018). Evolving the concept of Apd. *Intern. J. Audiol.* 57, 240–248.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read
for greatest visibility
and readership

FAST PUBLICATION
Around 90 days
from submission
to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative,
and constructive
peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers
acknowledged by name
on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: info@frontiersin.org | +41 21 510 17 00

REPRODUCIBILITY OF RESEARCH
Support open data
and methods to enhance
research reproducibility

DIGITAL PUBLISHING
Articles designed
for optimal readership
across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics
track visibility across
digital media

EXTENSIVE PROMOTION
Marketing
and promotion
of impactful research

LOOP RESEARCH NETWORK
Our network
increases your
article's readership