NORTHWESTERN UNIVERSITY

Placing a Perceptual Spotlight on Sound: Selective Attention and its Relation to Speech Recognition in Children with Normal Hearing and Hearing Loss

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Communication Sciences and Disorders

By

Kristina M. Ward

EVANSTON, ILLINOIS

September 2020

ABSTRACT

Placing a Perceptual Spotlight on Sound: Selective Attention and its Relation to Speech Recognition in Children with Normal Hearing and Hearing Loss

Kristina M. Ward

Speech recognition in complex acoustic environments is dependent on myriad bottom-up (i.e., peripheral) and top-down (i.e., central) processes. While bottom-up processes remain fairly stable during childhood, the development of top-down processes persists into young adulthood. The immaturity of top-down processes places younger children at considerable risk for poorer speech recognition in complex acoustic environments, which has consequences for academic success and social wellbeing. This is an especially important consideration for children with hearing loss who oftentimes have greater difficulty understanding speech despite restored access to the acoustic characteristics of speech through the use of clinical hearing devices. Therefore, elucidating the factors that influence the development of top-down processes during childhood is important in order to bolster children's speech recognition in complex acoustic environments.

Selective attention is a specific top-down process expected to underlie children's ability to understand speech in complex acoustic environments by placing a perceptual "spotlight" on the target speech to be further processed. While previous research has provided evidence of the development of selective attention during childhood, the effects of disrupted auditory experience on this development remain largely unknown. Additionally, only a few studies have directly tested the relation between selective attention and speech recognition in children, and the results of these studies have been mixed. This dissertation aimed to address these gaps in knowledge by investigating the effects of age and hearing loss on selective attention during childhood, and

quantifying the extent to which selective attention contributes to children's speech recognition in complex acoustic environments.

Children between 5 and 12 years of age with normal hearing and hearing loss participated in two studies to test the hypotheses that: 1) immaturity and disrupted auditory experience impede selective attention during childhood (*Chapter 2*); and 2) children's ability to selectively attend to a target speech stream and inhibit attention to competing auditory input contributes to their speech recognition (*Chapter 3*). In the first study, children performed a behavioral change detection task in the auditory and visual domains during which they were instructed to selectively attend to and detect deviant stimuli within a target stream while inhibiting attention to a distractor stream. Results revealed that younger children and children with hearing loss responded less frequently to deviants in the target stream and more frequently to deviants in the distractor stream, which is indicative of poorer selective attention, than older children and children with normal hearing. Notably, these age- and hearing status-related differences were observed across the auditory and visual domains. In the second study, the same children performed a speech recognition task across acoustic conditions that differed based on reverberation time, masker type, and the spatial location of the masker. Younger children and children with hearing loss demonstrated poorer speech recognition than older children and children with hearing loss. Additionally, children's ability to selectively attend to a target speech stream was significantly predictive of their ability to understand speech, especially under acoustic conditions expected to impose greater attentional demands.

Together, the findings from this dissertation provide novel insight regarding the relations among age, hearing loss, selective attention, and speech recognition during childhood. By demonstrating that age- and hearing status-related differences in selective attention account for observed variability in children's speech recognition, this work has the potential to inform targeted interventions to maximize children's academic success and social wellbeing.

ACKNOWLEDGEMENTS

This dissertation represents the near culmination of a graduate school experience that has been a time of immense academic, professional, and personal growth. The opportunities I have had, the training I have received, the knowledge I have gained, and the skills I have refined through pursuing simultaneous Doctor of Philosophy (Ph.D.) and Doctor of Audiology (Au.D.) degrees at Northwestern University are a direct result of an extraordinary network of individuals to whom I owe my deepest gratitude.

First and foremost, I would like to thank my advisor, Dr. Tina Grieco-Calub, for her steadfast support and encouragement throughout this process. Tina, I am incredibly honored to have had the experience of being your first Ph.D. student, and am forever grateful for the countless hours you have spent mentoring me, the knowledge you have bestowed upon me, and the ways in which you have continually challenged me. Thank you for always believing in my abilities and constantly supporting me in the pursuit of my goals. In addition to your direct influence on my development as a scientist and clinician, your levelheadedness, resilience, and humility are admirable traits that I aspire to hone both personally and professionally.

I would also like to extend my sincerest gratitude to the other members of my Ph.D. committee – Dr. Pamela Souza, Dr. Beverly Wright, and Dr. Elizabeth Norton – for their constructive feedback and unwavering dedication to my growth as a researcher. Our many discussions have profoundly impacted the ways in which I think about and approach the research process. Thank you for the time and effort you have dedicated to sharing your insight, expertise, and words of encouragement, and for consistently pushing me beyond my current limits.

Additionally, thank you to those who provided me with external support and mentorship throughout the duration of this dissertation project. To Scott Pfeiffer and the team at Threshold

Acoustics who expanded my knowledge of room acoustics and modeling through individualized training in acoustical engineering and soundscape design. Also, to Dr. Jungwha (Julia) Lee who offered ongoing guidance on statistical analyses. Lastly, to Dr. Peggy Nelson who provided helpful feedback on the theoretical and methodological design aspects of this project. Your involvement and input have been invaluable throughout this process.

To the members of the Hearing and Language Lab, including the many undergraduate and graduate students I have had the pleasure of mentoring and working alongside, thank you for your contributions to this project and the multiple projects that preceded it. From the many days spent collecting data to the frequent brainstorming sessions and epic games of Pictionary, I am thankful to have been a part of such a hardworking and genuinely delightful group of people. To Katherine, my academic sibling and office mate, thank you for thinking, planning, and executing alongside me every step of the way. I cannot fathom what this experience would have been like without you, and am eternally grateful for our collegiality and friendship.

Furthermore, I am very fortunate to have completed my graduate training in a department comprised of faculty, staff, and students who have provided me with a steady foundation of ongoing support. I would like to specifically acknowledge Dr. Kristine Riley who has been a continuous source of motivation and inspiration, and whose passion for evidence-based clinical practice is contagious. To Cindy Coy and Celeste Lee, thank you for your efforts in navigating the administrative complexities of my simultaneous doctoral programs. Lastly, to my fellow graduate students, thank you for your camaraderie both within and outside of the department.

Most importantly, I would like to thank my family and friends for their continuous love and encouragement. Mom and Dad, thank you for never stifling my tendency to ask questions, be curious, and openly communicate my thoughts. You have always set the precedent of hard work, dedication, and delayed gratification – all traits that served me well in this pursuit – and for that I am incredibly grateful. Katie, thank you for being a constant pillar of support and a reliable source of laughter when I need it most. To my wonderful friends near and far, thank you for keeping me grounded and reminding me that it's okay to take a break. Finally, to Kyle, thank you for being my rock through the triumphs and tribulations of graduate school and life in general. Your patience, generosity, love, and endless support mean the world to me.

The following sources of funding provided essential support for this project: NIH NIDCD F31 DC017055 to Kristina Ward and NIH NIDCD R56 DC015492 to Tina Grieco-Calub.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 | INTRODUCTION

1.1 BACKGROUND

1.1.1 Understanding Speech in Real-World Environments

On a daily basis, children must understand speech in environments containing target speech and competing auditory input. Competing auditory input can originate from an isolated source, such as a television or a single person talking, or multiple sources constituting a complex acoustic environment. Regardless of the origin, there is vast variability in children's ability to understand speech in the presence of competing auditory input (e.g., Corbin et al., 2016; Leibold & Buss, 2013). While the reasons for this variability are only marginally understood, it is apparent that difficulty understanding speech amidst competing auditory input has significant implications for children's ability to listen and learn in real-world environments.

Elementary school classrooms represent a specific complex acoustic environment within which children spend a considerable amount of time. These environments often consist of coincident speech from the teacher and students as well as miscellaneous noise generated by heating and air conditioning systems, the shuffling of papers and other supplies, the movement of students, and diffuse sounds from the hallway and adjacent classrooms. When present concurrently, these various sources of sound contribute to the high levels of background noise observed in elementary school classrooms, which range from 30 to 70 dBA (Crandell & Smaldino, 2001; Crukley et al., 2011; Nelson et al., 2008; Picard & Bradley, 2001). In addition to background noise, reverberation within the classroom generates reflections of the sound present in the environment, which provides additional sources of acoustic competition (Crukley et al., 2011; Knecht et al., 2002). Previous studies have documented significant deficits in children's ability to understand speech under the acoustic conditions of typical elementary school classrooms (Bradley & Sato, 2008; Finitzo-Hieber & Tillman, 1978; Jamieson et al., 2004; Neuman & Hochberg, 1983). The consequences of poorer speech recognition include children missing out on information presented orally, having difficulty following verbal instructions, and needing to expend additional effort to communicate with their peers, all of which compromise their academic success and social wellbeing (Anderson, 2004; Dockrell & Shield, 2006; Klatte et al., 2010; Nelson & Soli, 2000; Shield, 2008; Shield & Dockrell, 2003).

Together, the observation that children have difficulty understanding speech in the presence of competing auditory input combined with the fact that the classroom environments within which children spend a majority of their time are acoustically complex underscores the need to better understand the mechanisms that contribute to children's speech recognition in real-world environments.

1.1.2 Processes that Underlie Speech Recognition in Children

Decades of research have aimed to elucidate the processes that underlie children's ability to understand speech in complex acoustic environments. Converging evidence demonstrating that older children consistently achieve better speech recognition in the presence of background noise and reverberation than younger children suggests that these processes continue to develop during childhood. Additionally, the fact that children of the same age oftentimes demonstrate a range of speech recognition abilities provides evidence that there are individual differences in the rate at which these processes develop during the elementary school years (e.g., Corbin et al., 2016; Leibold & Buss, 2013).

The range of processes expected to contribute to children's ability to understand speech in complex acoustic environments can be generally grouped into bottom-up (i.e., peripheral) processes and top-down (i.e., central) processes (e.g., Alain et al., 2001; Mattys et al., 2012; Moore, 2012; Rönnberg et al., 2013; Stenfelt & Rönnberg, 2009; Sussman & Steinschneider, 2009). Bottom-up processes reflect children's sensory access to the frequency, timing, and level characteristics of the target speech and competing auditory input and are based on patterns of activation in the peripheral auditory system. Examples of bottom-up processes include frequency resolution, temporal integration, and intensity discrimination. Top-down processes refer to the involvement of cognitive abilities and linguistic knowledge to facilitate perception by extracting and encoding information from the target speech in line with behavioral goals. Examples of top-down processes include those related to language ability (e.g., phonological awareness, lexical access, semantic knowledge), working memory, and attention. While speech recognition in complex acoustic environments is expected to involve the dynamic interplay between bottom-up and top-down processes, the observed improvement in speech recognition throughout childhood is thought to primarily reflect the development of cognitive and linguistic processes as the peripheral representation of sound remains fairly constant during childhood (Litovsky, 2015; Moore & Linthicum, 2007).

A top-down process that is of particular interest in considering how children understand speech in complex acoustic environments is selective attention, which aids in the segregation, encoding, and processing of specific auditory streams, such as target speech (Alain & Arnott, 2000; Leibold, 2012; Sussman, 2017). Specifically, while all auditory input evokes activation in the peripheral auditory system, selectively attending to the target speech stream places a perceptual spotlight on the components of the input to be further processed by auditory- and language-based cortical areas. Therefore, children's ability to understand speech in complex acoustic environments is expected to be, in part, dependent on their ability to selectively attend to the target speech stream from which they aim to extract information and inhibit attention to competing auditory input (Cherry, 1953; Wightman & Kistler, 2005).

In considering the anticipated role of selective attention in speech recognition, a potential reason why children have greater difficulty understanding speech in complex acoustic environments is their cognitive immaturity. The development of executive functions, including attention, follows an ascending trajectory during childhood that does not plateau until early adulthood (e.g., Brocki & Bohlin, 2004; Klenberg et al., 2001). Therefore, children are expected to have a reduced attentional capacity as well as greater difficulty selectively allocating attentional resources relative to older children and adults. This does not bode well for children's speech recognition in real-world environments, as the degradation of the acoustic signal via background noise and reverberation increases the attentional demands associated with understanding speech. While selective attention has been shown to contribute to performance on basic psychophysical tasks (e.g., Jones et al., 2015; Sussman et al., 2007) and modulate the neural representation of speech (Forte et al., 2017; Lehmann & Scho, 2014; Rimmele et al., 2015; Wild et al., 2012; Yoncheva et al., 2010), research examining the direct link between selective attention and speech recognition in children remains scarce.

1.1.3 Considerations for Children with Hearing Loss

An in-depth understanding of the processes that contribute to speech recognition in complex acoustic environments is especially important for children with hearing loss, who are frequently educated alongside their peers with normal hearing in mainstream classrooms. Despite the use of clinical hearing devices, which provide children with improved access to the acoustic characteristics of speech, children with hearing loss are at significantly greater risk for poor social and academic outcomes (Bess et al., 1998; Porter et al., 2013; Sarant et al., 2015; Walker et al., 2020). A potential reason for this observation is that children with hearing loss have significant difficulty understanding speech in the presence of background noise and reverberation despite excellent speech recognition under optimal acoustic conditions (Crandell, 1993; Crandell & Smaldino, 2000; Finitzo-Hieber & Tillman, 1978; Inglehart, 2016; Leibold et al., 2013; McCreery et al., 2019). While clinical hearing devices improve children's access to auditory input, including speech, they do not alleviate the attentional demands associated with understanding speech in complex acoustic environments. Thus, poorer speech recognition in children with hearing loss may partially reflect their inability to selectively attend to the target speech stream.

Previous research has demonstrated that disrupted auditory experience early in life may alter the development of executive functions, including attention, in children with hearing loss (Beer et al., 2014; Conway et al., 2009; Dye & Hauser, 2014; Figueras et al., 2008; Houston & Bergeson, 2014; Kronenberger et al., 2020; Monroy et al., 2019; Oberg & Lukomski, 2011). Consistent with these observations, children with hearing loss have been shown to perform more poorly on tasks requiring selective attention in the auditory and visual domains than children with normal hearing (Mitchell & Quittner, 1996; Quittner et al., 1994; Smith et al., 1998). In conjunction with the established immaturity of cognitive processes during childhood, poorer selective attention in children with hearing loss may contribute to their difficulty understanding speech in the presence of competing auditory input. However, the relations among these processes are only marginally understood. Therefore, examining how hearing loss during childhood affects selective attention and the implications of this for speech recognition in complex acoustic environments is a critical first step toward bolstering social and academic outcomes in children with hearing loss.

1.2 SPECIFIC AIMS AND HYPOTHESES

A primary objective of this dissertation is to investigate how age and hearing loss influence children's ability to selectively attend to a target speech stream amidst competing auditory input. Furthermore, we aimed to understand the extent to which individual differences in selective attention during childhood contribute to differences in speech recognition in complex acoustic environments. While previous studies have examined aspects of these relations, this project is the first, to our knowledge, to investigate selective attention, speech recognition, and the associations between these processes in the same cohorts of children with normal hearing and children with hearing loss.

The central hypothesis of this dissertation is that poor selective attention during childhood, due to immaturity and disrupted auditory experience, contributes to the difficulties younger children and children with hearing loss have understanding speech in complex acoustic environments. The experiments reported herein tested this hypothesis through the following specific aims:

Aim 1: To determine how age and hearing loss influence selective attention during childhood (*Chapter 2*). If immaturity and disrupted auditory experience impede selective attention during childhood, then younger children and children with hearing loss should have greater difficulty selectively attending to a target stream and inhibiting attention to competing input than older children and children with normal hearing. Furthermore, if the effect of hearing loss on selective attention is domain general, children with hearing loss should demonstrate poorer selective attention than children with normal hearing regardless of the sensory domain of the task.

Aim 2: To measure the extent to which selective attention contributes to children's ability to understand speech in complex acoustic environments (*Chapter 3*). If children's ability to selectively attend to a target speech stream and inhibit attention to competing auditory input contributes to their speech recognition, then children who have greater difficulty selectively attending to a target speech stream should have poorer speech recognition, especially under acoustic conditions that impose greater attentional demands.

1.3 OVERVIEW OF DISSERTATION

Chapters 2 and *3* of this dissertation describe the findings relevant to these aims. Specifically, a multi-experiment within-subjects design was used to investigate the relations among age, hearing status, selective attention, and speech recognition within the same cohorts of children. Children between 5 and 12 years of age with normal hearing and varying degrees of hearing loss were recruited to participate.

Chapter 2 describes a set of experiments where children performed behavioral change detection tasks in the auditory and visual domains. Children's ability to selectively attend and respond to deviants in the designated target stream and inhibit attention to a competing distractor stream was assessed. Linear quantile mixed-effects modeling was used to test for the effects of age and hearing status on children's ability to selectively attend to the target stream across the auditory and visual domains. Additionally, correlation analyses were used to quantify the relations among children's performance on the auditory and visual change detection tasks, their hearing sensitivity, and their performance on standardized measures of executive function.

Chapter 3 describes an experiment with the same cohorts of children that involved quantifying speech recognition across a variety of acoustic conditions that differed based on reverberation time, masker type, and the spatial location of the masker. Linear mixed-effects modeling was used to test for the effects of age, hearing status, and acoustic condition on children's ability to understand speech. In addition, hierarchical linear regression analyses were used to assess the relation between children's ability to selectively attend to a target speech stream and their speech recognition, as well as how these relations changed based on the attentional demands imposed by the acoustic conditions under which speech recognition was measured.

The final chapter (*Chapter 4*) summarizes the findings from *Chapters 2* and *3* and discusses the potential implications of these findings for speech recognition in real-world environments and evidence-based clinical practice.

CHAPTER 2 | EFFECTS OF AGE AND HEARING LOSS ON SELECTIVE ATTENTION DURING CHILDHOOD

2.1 ABSTRACT

Younger children have greater difficulty understanding speech in complex acoustic environments than older children and adults. A potential reason for this may be that children's ability to selectively attend to a target speech stream and inhibit attention to competing input continues to develop during childhood. This is an especially important consideration for children with hearing loss, who have difficulty understanding speech despite the use of their clinical hearing devices. However, the effects of immaturity and disrupted auditory experience on selective attention during childhood are only marginally understood. The purpose of the present study was to investigate these relations. Children between 5 and 12 years of age with normal hearing and hearing loss performed a behavioral change detection task comprised of a target and distractor stream in the auditory and visual domains. Results revealed that younger children and children with hearing loss had greater difficulty selectively attending to a target stream and inhibiting attention to a distractor stream than older children and children with normal hearing. Notably, these age- and group-related differences in selective attention were observed across the auditory and visual domains. Together, these findings suggest that children's ability to selectively attend to a target stream is influenced by their age and hearing status.

2.2 INTRODUCTION

The ability to understand speech in complex acoustic environments continues to develop during childhood. Specifically, younger children require more favorable signal to noise ratios (Corbin et al., 2016; Litovsky, 2005; Wightman & Kistler, 2005;), shorter reverberation times

(Neuman et al., 2010), and background noise void of speech content (Hall et al., 2002; Leibold & Buss, 2013) to achieve similar performance on measures of speech recognition as older children and adults. In addition, there is often considerable variability in speech recognition among children of the same chronological age regardless of the target speech material and the content of competing background noise (e.g., Corbin et al., 2016; Johnstone & Litovsky, 2006; Lewis et al., 2014), suggesting that age, while significant, is not a sole determinant of speech recognition ability. As peripheral auditory function is fully developed by early childhood in children with normal hearing (Litovsky, 2015; Moore & Linthicum, 2007), it is expected that non-sensory factors related to other aspects of development, such as the immaturity of cognitive and linguistic processes, contribute to children's difficulty understanding speech in complex acoustic environments (McCreery et al., 2019; Thompson et al., 2019).

An in-depth understanding of the developmental factors that contribute to speech recognition is especially important for children with bilateral hearing loss who, like children with normal hearing, spend a considerable amount of time listening and learning in environments containing background noise and reverberation. Providing access to the acoustic characteristics of speech is essential for children with hearing loss and is often achieved by fitting these children with hearing devices, such as hearing aids and cochlear implants. Children with hearing devices that restore the audibility of speech have been shown to achieve speech recognition in quiet that is similar to their peers with normal hearing (Crandell & Smaldino, 2000; Hicks & Tharpe, 2002). However, despite the use of hearing devices that improve access to the acoustic characteristics of speech, many children with even mild degrees of hearing loss demonstrate significantly poorer speech recognition (Crandell, 1993; Lewis et al., 2016) and increased listening effort to understand speech (Hicks & Tharpe, 2002; McGarrigle et al., 2019) than their peers with normal hearing in complex acoustic environments. These findings demonstrate that audibility is necessary, but not sufficient, for understanding speech amidst background noise and reverberation. Given that children's ability to understand speech in complex acoustic environments improves with age, it is important to consider the contribution of cognitive and linguistic processes, which also continue to develop during childhood, to speech recognition.

The present study aimed to determine how *selective attention* – a specific cognitive process that is of particular interest when elucidating how children understand speech in complex acoustic environments – is influenced by immaturity and disrupted auditory experience during childhood. Understanding the factors that influence the development of selective attention will provide novel insight regarding the difficulties younger children and children with hearing loss face understanding speech in complex acoustic environments.

2.2.1 Selective Attention and Speech Recognition

Within the context of speech recognition in complex acoustic environments, selective attention refers to a listener's ability to direct attention to a target speech stream and inhibit attention to competing auditory streams. The role of selective attention in speech recognition is often considered within the framework of auditory scene analysis, which is the proposed model for how listeners identify, segregate, and extract information from a target speech stream amidst competing auditory input (e.g., Bregman, 1990, 1993; see Leibold, 2012 for a review). Specifically, individual sound sources generate acoustic signals, the spectrotemporal structures of which are encoded by the listener's peripheral auditory system to form distinct auditory objects (e.g., Shinn-Cunningham, 2008). As the acoustic signals from individual sound sources unfold over time, auditory objects that share frequency, timing, and/or spatial location cues are perceptually segregated from one another by the listener. This segregation is necessary to form auditory streams with each stream corresponding to a distinct environmental sound source (e.g., Moore & Gockel, 2012; Sussman et al., 2001). While these early processing stages of auditory scene analysis are typically thought to occur automatically (e.g., Bregman, 1990), it is generally accepted that, in complex acoustic environments, selective attention can facilitate the formation of auditory objects and segregation of auditory streams (Alain & Bernstein, 2008; Bregman, 1993; Carlyon et al., 2001, 2003; Shinn-Cunningham, 2008; see Sussman, 2017 for a review). Following the formation of auditory streams, a listener must selectively attend to the stream from which they aim to extract information (i.e., target speech stream), and inhibit attention to competing streams (e.g., Alain & Arnott, 2000; Fritz et al., 2007; Gordon-Salant & Fitzgibbons, 1993; Shinn-Cunningham & Best, 2008; Snyder et al., 2012; Sussman, 2017). Therefore, children who have greater difficulty selectively attending to a target speech stream are expected to have poorer speech recognition in complex acoustic environments. In order to directly test this relation, however, it is necessary to first establish a solid foundation of understanding for how selective attention changes with age during childhood and how disrupted auditory experience, as in children with hearing loss, influences this development.

2.2.2 Selective Attention During Childhood

Previous studies have demonstrated that the ability to selectively attend to a target auditory stream continues to develop during childhood in children with normal hearing (Gomes, 2000; Sussman & Steinschneider, 2009). For instance, a study by Jones, Moore, & Amitay (2015) revealed that children’s ability to inhibit attention to noise that was similar in frequency to a target stimulus during a behavioral tone-in-noise detection task improved between 4 to 11 years of age, with performance becoming adultlike around 9 to 11 years of age. Similar developmental trends have been observed with the use of speech stimuli (Coch et al., 2005; Doyle, 1973; Sanders et al., 2006; Wightman & Kistler, 2005). For instance, Wightman & Kistler (2005) had children 4 to 16 years of age and adults perform a closed-set speech recognition task during which they were instructed to attend to a target speech stream in one ear while ignoring competing speech streams presented in the same ear (i.e., ipsilateral) or the opposite ear (i.e., contralateral). Compared to adults, children were less able to inhibit attention to the competing speech streams. Notably, a substantial proportion of children's response errors reflected words contained in the competing speech streams. In conjunction with this response pattern, the observation that children demonstrated cross-stream competition even when the target and competing speech streams were presented dichotically – and were, therefore, peripherally isolated – is suggestive of a reduced ability to selectively attend to the target speech stream and inhibit attention to competing auditory streams. This interpretation is further supported by findings demonstrating that the development of children's ability to understand speech in the presence of background noise that is expected to be more attentionally engaging (e.g., two-talker speech) is delayed relative to speech recognition amidst background noise that is less attentionally engaging (e.g., speech-shaped noise; Corbin et al., 2016; Hall et al., 2002; Leibold & Buss, 2013). Together, these findings suggest that, for children with normal hearing, the ability to selectively attend to a target speech stream and inhibit attention to competing auditory streams continues to develop during childhood and parallels the development of selective attention.

Considerably less is known about the development of selective attention in children with hearing loss who experience reduced access to robust auditory input during the formative early years of cognitive development. Even with the use of properly fitted hearing devices that increase access to auditory input, the quality of this input is degraded in children with hearing relative to children with normal hearing. Several studies investigating the effect of childhood hearing loss on neurocognitive development have found that children with hearing loss perform more poorly on auditory and non-auditory measures of executive function, specifically attention and inhibitory control, than children with normal hearing (Beer et al., 2014; Dye & Hauser, 2014; Figueras et al., 2008; Houston & Bergeson, 2014; Kronenberger et al., 2020; Monroy et al., 2019; Oberg & Lukomski, 2011). There are several proposed mechanisms for these observed relations. For instance, compromised access to auditory input early in life may lead to the reorganization of neural networks in the frontal and prefrontal cortices, which may alter the development of executive function in children with hearing loss (Conway et al., 2009). Relatedly, reduced opportunities to integrate multimodal sensory information in children with hearing loss may alter their ability to selectively allocate attention to a domain-specific task. For instance, children with hearing loss have been shown to demonstrate poorer selective attention during temporal sequencing tasks in the visual domain, perhaps due to their dependence on vision to monitor their environments to a greater extent than children with normal hearing (Mitchell & Quittner, 1996; Quittner et al., 1994; Smith et al., 1998). In addition, children with hearing loss may exhibit a differential allocation of attentional resources to auditory and visual input compared to children with normal hearing, as demonstrated by their reduced preparatory attention to a spatialized target speech stream (Holmes et al., 2017), tendency to direct attention away from auditory tasks (McFadden & Pittman, 2008), and increased responses to distractor stimuli during a temporal sequencing task in the visual domain (Dye & Hauser, 2014). Lastly, the fact that children with hearing loss receive a degraded internal representation of the acoustic signal at the level of the peripheral auditory system increases the demand for top-down attentional processes to aid in auditory object formation and stream segregation. Due to the limited capacity of these processes (Kahneman, 1973), an expected consequence of this increased dependence on top-down processes is the reduced availability of spare attentional resources to be selectively allocated to the target speech stream to be tracked and encoded by higher-level processes (see Shinn-Cunningham & Best, 2008 for a review). If, in fact, the quality and quantity of auditory input a child receives contributes to their ability to allocate attentional resources within and across sensory domains, it is expected that children with hearing loss will demonstrate greater difficulty selectively attending to a target stream than their peers with normal hearing regardless of task domain.

2.2.3 The Present Study

The primary objective of the present study was to test the hypothesis that immaturity and disrupted auditory experience impede selective attention during childhood. If so, younger children and children with hearing loss should have greater difficulty selectively attending to a target speech stream than older children and children with normal hearing. Furthermore, if the effect of hearing loss on selective attention is domain general, children with hearing loss should demonstrate poorer selective attention than children with normal hearing regardless of the sensory domain of the task. Alternatively, if the effect of hearing loss on selective attention is specific to the sensory domain of impairment (i.e., the auditory domain), then children with hearing loss should demonstrate similar selective attention to their peers with normal hearing in non-auditory domain tasks.

To test these predictions, children with normal hearing and children with varying degrees of hearing loss performed a behavioral auditory change detection (ACD) task during which they were instructed to selectively attend to a target auditory stream, inhibit attention to a distractor auditory stream, and respond to frequency deviants only when they occurred in the designated target stream. The same children also completed an analogous visual change detection task (VCD) during which they were instructed to selectively attend to a target visual stream, inhibit attention to a distractor visual stream, and respond to color deviants only when they occurred in the designated target stream.

2.3 METHODS

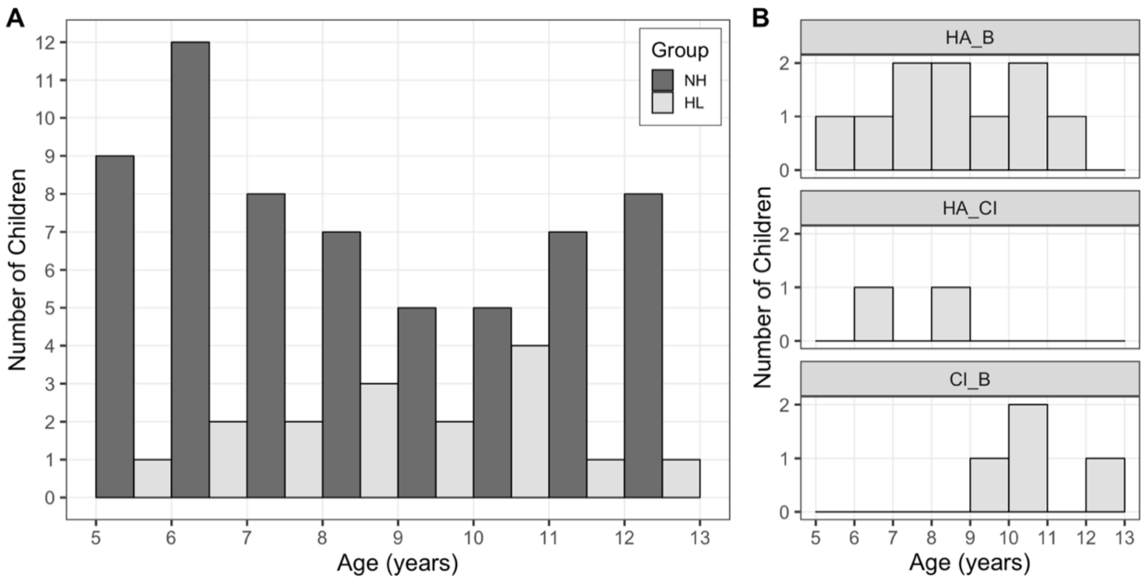
2.3.1 Participants

A total of 77 children 5 to 12 years of age, 61 children with normal hearing (NH) and 16 children with bilateral hearing loss (HL) who used hearing aids and/or cochlear implants, participated in this study. All children were native American-English speakers with no history of speech, language, cognitive, attentional, or neurodevelopmental disorders. In addition, all children had normal or corrected to normal vision per parental report. Children were recruited through the Child Studies Group registry at Northwestern University, informational flyers disseminated throughout the greater Chicagoland area, and word of mouth. Children with HL were additionally recruited through collaborations with Lurie Children's Hospital in Chicago, IL and the Medical College of Wisconsin in Milwaukee, WI. Approval for all study procedures was obtained from the Institutional Review Board at Northwestern University and all children completed an informed consent/assent process prior to participation.

Audiometric and Device Information

Pure-tone audiometry (unaided for ears with acoustic hearing and aided for all ears with hearing devices) at all octave frequencies between 250-8000 Hz and tympanometry were performed at the time of testing if the most recent audiogram was greater than six months old or if the audiogram did not contain aided thresholds. Otherwise, children's most recent audiogram was used to confirm their eligibility for study enrollment. Thresholds at 500, 1000, 2000, and 4000 Hz were averaged to calculate the pure-tone average (PTA) for each ear in the unaided and aided conditions. Unaided PTAs from pre-implant audiograms were obtained for ears with cochlear implants to serve as a proxy for degree of hearing loss. For children who used hearing aids, real-ear measures (REM) were performed to quantify gain across frequencies.

Children with NH had hearing sensitivity within normal limits as indicated by pure-tone thresholds ≤25 dB HL from 250-8000 Hz, bilaterally, as well as normal middle ear function as indicated by Type A tympanograms, bilaterally. Children with HL had bilateral sensorineural (SNHL; $N = 14$) or conductive (CHL; $N = 2$) hearing loss that ranged from mild to profound. Children with SNHL had normal middle ear function as indicated by Type A tympanograms, bilaterally. Children with CHL had abnormal middle ear function due to chronic tympanic membrane perforations (3 ears) and surgical intervention for cholesteatoma (1 ear) as indicated by Type B tympanograms with large ear canal volumes. Children with HL used their own bilateral clinical hearing devices (i.e., hearing aids and/or cochlear implants) fitted by a licensed audiologist throughout testing. Ten children used hearing aids in both ears (HA_B), two children used a hearing aid in one ear and a cochlear implant in the other ear (HA_CI), and four children used cochlear implants in both ears (CI_B). **Figure 2.1** displays the distribution of children across ages for each hearing status and device group. **Table 2.1** contains audiometric and device information for individual children with HL.



<u>Figure 2.1</u>: Distribution of children's ages across hearing and device groups.

Speech Recognition Assessment

In addition, all children completed a sentence recognition task in quiet using 10 Bamford-Kowal-Bench sentences (BKB; Bench et al., 1979) spoken by a male talker and presented at 65 dB SPL in the sound field from 0 degrees azimuth. Performance on this task served as a measure of speech discrimination in the best aided condition for children with HL. Four additional children were recruited and assessed but excluded from analysis for the following reasons: abnormal responses during pure-tone audiometry (child with NH; 9 years of age), scores less than 50% correct on the speech recognition task (children with HL [HA_B]; 7 and 9 years of age) and noncompliance during testing (child with HL [CI_B]; 12 years of age).

Testing occurred within one session that lasted approximately 2 hours, including breaks. Children were compensated at a rate of $10/hour for their participation. In addition, children received a book and a small prize at the conclusion of the session after trading in tickets they earned for completing each task.

Age (Years)	Sex	Type of HL	Etiology	Device Group	Left Device	Right Device	Age at Left Device	Age at Right Device	Daily Wear (Hours)	Left HFPTA Unaided | Aided (dB HL)	Right HFPTA Unaided | Aided (dB HL)
5.12	F	SNHL	Unknown	HA_B	Oticon BTE	Oticon BTE	3 years	3 years	5-9	40 | 23.75	38.75 | 28.75
6.06	M	SNHL	Unknown	HA_CI	Phonak BTE	Cochlear N7	4 months	3 years	10-14	66.25 | 21.25	87.50 | 21.25
6.38	F	SNHL	Unknown	HA_B	Phonak BTE	Phonak BTE	2.5 months	2.5 months	10-14	61.25 | 22.5	67.5 | 22.5
7.85	F	SNHL	Genetic	HA_B	ReSound BTE	ReSound BTE	3 months	3 months	10-14	62.5 | 25	56.25 | 22.5
7.85	M	SNHL	Genetic	HA_B	ReSound BTE	ReSound BTE	3 months	3 months	10-14	65 | 20	62.5 | 20
8.31	F	SNHL	Unknown	HA_B	Phonak BTE	Phonak BTE	5 years	5 years	5-9	45 | 21.25	46.25 | 21.25
8.53	F	SNHL	Unknown	HA_CI	Oticon BTE	Med-El Rondo	3 months	7 years	10-14	83.75 | 23.75	76.25 | 18.75
8.90	M	SNHL	Unknown	HA_B	Phonak BTE	Phonak BTE	3 years	3 years	10-14	21.25 | 15	26.25 | 15
9.32	M	SNHL	Unknown	CI_B	Cochlear N7	Cochlear N7	1 year	1 year	14+	106.25 | 17.5	106.25 | 17.5
9.98	M	CHL	COM (L/R) Chol. (L)	HA_B	Oticon BTE	Oticon BTE	4 years	4 years	10-14	47.5 | 23.75	35 | 21.25
10.46	M	SNHL	Unknown	HA_B	ReSound BTE	ReSound BTE	6 years	6 years	5-9	33.75 | 25	38.75 | 25
10.52	M	SNHL	Genetic	CI_B	Med-El OPUS 2	Med-El OPUS 2	1 year	1 year	14+	102.5 | 28.75	102.5 | 27.5
10.54	F	SNHL	Genetic	HA_B	Phonak BTE	Phonak BTE	1 year	1 year	5-9	35 | 15	31.25 | 15
10.72	F	SNHL	ANSD	CI_B	Cochlear N7	Cochlear N7	4 years	4 years	10-14	68.75 | 22.5	70 | 17.5
11.01	F	CHL	COM	HA_B	ReSound RIC	ReSound RIC	10 years	10 years	5-9	30 | 11.25	30 | 11.25
12.30	M	SNHL	Genetic	CI_B	Cochlear N7	Cochlear N7	7 months	7 months	14+	111.75 | 20	111.75 | 25

Table 2.1: Audiometric and device information for children with hearing loss. COM = chronic otitis media; Chol. = cholesteatoma; ANSD = auditory neuropathy spectrum disorder

2.3.2 Standardized Measures of Executive Function

In addition to the experimental tasks described below, the following standardized measures of attention were completed during testing. **Table 2.2** displays children's performance on these measures as well as additional demographic information.

Flanker Inhibitory Control and Attention Test

The Flanker Inhibitory Control and Attention Test (Weintraub, Anderson, & Manly, 2013) is a subtest of the NIH Toolbox Cognition Battery that assesses the ability to selectively attend to target stimuli and inhibit attention to irrelevant stimuli in the visual domain. The test is administered via an iPad and normed on individuals 3 to 85 years of age. On each trial, a target (i.e., fish for children < 8 years of age and arrows for children ≥ 8 years of age) facing to the left or right is presented in the center of the screen with two similar stimuli (i.e., flankers) positioned on either side. Children were instructed to indicate the direction of the centered stimulus by pressing the leftward or rightward facing arrow on the screen as quickly as possible while ignoring the flankers. The test consists of randomized congruent (i.e., target and flankers facing the same direction) and incongruent (i.e., target and flankers facing opposite directions) trials for a total of 20 trials. Accuracy and reaction time of responses automatically combined to generate uncorrected scores and age-adjusted scaled scores (mean = 100, *SD* = 15). For children 5 to 7 years of age, an additional 20 trials containing arrow stimuli are presented if an accuracy score of at least 90% is achieved on the trials containing fish stimuli. All children completed the Flanker Inhibitory Control and Attention Test.

Test of Everyday Attention for Children

The Test of Everyday Attention for Children 2nd Edition (TEA-Ch2; Manly et al., 2016) is a compilation of tests that assess separable aspects of attention in the auditory and visual domains. The tests are administered via a computer and normed on individuals 5 to 15 years of age. Two versions of the TEA-Ch2 were developed – one for children 5 to 7 years of age and one for children 8 to 15 years of age – with tests that are conceptually similar but contain age-appropriate themes.

The sustained attention battery was utilized in the present study and consists of four subtests as follows: 1) *Barking/Vigil* is an auditory monitoring task during which children mentally count individually presented stimuli separated by long silent gaps with performance quantified by the total number of correct responses; 2) *Hide and Seek Auditory/Cerberus* is an auditory target-detection task during which children listen to short sound clips and press the spacebar as quickly as possible when a target stimulus is presented while ignoring distractor stimuli with performance quantified by the mean reaction time of responses (in ms) weighted for accuracy; 3) *Sustained Attention Response Test* is a visual monitoring task during which children press the spacebar after each shape is presented in the center of the screen but withhold their response when a predetermined no-go shape is presented with performance quantified by the number of no-go trial responses; and 4) *Simple Reaction Time* is a visual target-detection task during which children press the spacebar as quickly as possible when they see the target stimulus appear on the screen with performance quantified by the mean reaction time of responses (in ms). The accuracy and/or reaction time of responses on each subtest are utilized to automatically generate raw and scaled scores (mean = 10, *SD* = 3) for each subtest as well as a composite sustained attention index score (mean = 100, *SD* = 15). Five children (4 with NH and 1 with HL [HA_B]) did not complete the TEA-Ch2 due to its addition to the experimental protocol after their data had been collected. Additionally, three children did not complete at least one subtest of the TEA-Ch2 due to non-compliance or fatigue during testing (1 with NH missing the Sustained Attention Response Test; 1 with NH missing the Simple Reaction Time, and 1 with HL [HA_B] missing the Sustained Attention Response Test and Simple Reaction Time). As a result of these missing subtest scores, the sustained attention index could not be calculated for these children.

Children's Auditory Performance Scale

The Children's Auditory Performance Scale (CHAPS; derived from Smoski, Brunt, & Tannahill, 1998) is a parent questionnaire that assesses children's auditory attention skills across various real-world listening conditions (e.g., in a quiet room; in a room where there is background noise and other distractions). For each item, parents rate (on a scale of +1 indicating less difficulty to -5 indicating inability to function) the amount of difficulty their child has performing an attention-based auditory task under various scenarios relative to their same-age peers. A rating of 0 indicates their child has the same amount of difficulty on a given task as their peers. Parent ratings for each item are added across listening conditions to generate a composite score for each child that ranges from -130 to +36 with more negative values indicating poorer auditory attention skills. The parents of four children (3 with NH and 1 with HL [HA_B]) did not complete the CHAPS due to its addition to the experimental protocol after their children's data had been collected.

	Group	
	Children with NH	**Children with HL**
Number	61	16
Sex	Female = 25 Male = 36	Female = 8 Male = 8
Age (Years)	Mean = 8.64 | *SD* = 2.41 Range = 5.26-12.96	Mean = 8.99 | *SD* = 1.99 Range = 5.12-12.30
Unaided PTAs		
Left Ear (dB HL)	Mean = 5.33 | *SD* = 5.30 Range = -3.75-20.00	Mean = 61.28 | *SD* = 28.06 Range = 21.25-111.75
Right Ear (dB HL)	Mean = 5.66 | *SD* = 5.48 Range = -6.25-20.00	Mean = 61.67 | *SD* = 28.70 Range = 26.25-111.75
Aided PTAs		
Left Ear (dB HL)	-	Mean = 21.02 | *SD* = 4.48 Range = 11.25-28.75
Right Ear (dB HL)	-	Mean = 20.31 | *SD* = 4.29 Range = 11.25-27.50
Maternal Education Level (Years)	Mean = 15.54 | *SD* = 0.87 Range = 13-17	Mean = 14.88 | *SD* = 1.31 Range = 13-17
CHAPS Total Score	Mean = -2.10 | *SD* = 18.57 Range = -73-35	Mean = -17.20 | *SD* = 19.43 Range = -50-25
Flanker Scores		
Raw	Mean = 91.31 | *SD* = 12.76 Range = 39-110	Mean = 84.38 | *SD* = 21.47 Range = 38-112
Scaled	Mean = 103.69 | *SD* = 13.30 Range = 83-135	Mean = 96.19 | *SD* = 17.06 Range = 73-136
TEA-Ch2 Scores		
Barking/Vigil		
Raw (# correct)	Mean = 7.97 | *SD* = 1.91 Range = 3-10	Mean = 7.13 | *SD* = 2.48 Range = 1-10
Scaled	Mean = 8.97 | *SD* = 3.15 Range = 3-15	Mean = 7.67 | *SD* = 3.56 Range = 1-13
Hide and Seek Auditory/Cerberus		
Raw (ms)	Mean = 1747.51 | *SD* = 1286.76 Range = 0-5699.17	Mean = 2181.62 | *SD* = 1649.20 Range = 713.33-6832.00
Scaled	Mean = 12.63 | *SD* = 4.35 Range = 1-19	Mean = 11.13 | *SD* = 4.44 Range = 4-18
Sustained Attention Response Test		
Raw (# incorrect)	Mean = 8.57 | *SD* = 5.16 Range = 0-17	Mean = 9.29 | *SD* = 5.28 Range = 0-17
Scaled	Mean = 10.98 | *SD* = 3.07 Range = 3-18	Mean = 10.14 | *SD* = 2.98 Range = 6-16
Simple Reaction Time		
Raw (ms)	Mean = 732.45 | *SD* = 221.71 Range = 409.85-1335.12	Mean = 706.03 | *SD* = 273.56 Range = 413.66-1316.99
Scaled	Mean = 8.04 | *SD* = 3.14 Range = 1-15	Mean = 7.86 | *SD* = 3.82 Range = 2-13
Sustained Attention Index		
Raw	Mean = 40.80 | *SD* = 8.44 Range = 22-57	Mean = 37.07 | *SD* = 9.29 Range = 25-55
Scaled	Mean = 101.66 | *SD* = 17.65 Range = 65-132	Mean = 93.43 | *SD* = 19.71 Range = 70-132

Table 2.2: Descriptive statistics for demographic information and measures of executive function for children with normal hearing and hearing loss.

2.3.3 Stimuli

Auditory Stimuli

Stimuli for the auditory change detection task consisted of 106 monosyllabic English nouns determined to be within the lexicon of children as young as five years of age that were spoken by a female and a male talker (Cortese & Khanna, 2008; Kuperman et al., 2012). Each talker was a monolingual native American-English-speaking young adult with normal hearing who resided in Chicago, IL. Words were recorded in isolation at least three times, from which the most natural sounding exemplar, as rated by a native American-English-speaking doctoral student with normal hearing, was chosen. Recordings were made in a double-walled sound-attenuating booth at a sampling rate of 44.1 kHz and resolution of 16 bits using a Blue Snowball Microphone connected to a MacBook Pro laptop running Praat open-source software. The recorded words were root-mean-square (RMS) normalized and exported in .wav audio format. The duration of the words ranged from 379 ms to 963 ms (mean duration of 650 ms) for the female talker and from 433 ms to 838 ms (mean duration of 600 ms) for the male talker. Three young adults with normal hearing who were native American-English speakers (two doctoral students and one undergraduate student) and two children with normal hearing (ages 5 and 10 years) who were native American-English speakers listened to the final corpus of recorded words to verify intelligibility and sound quality.

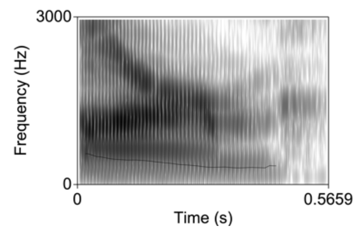
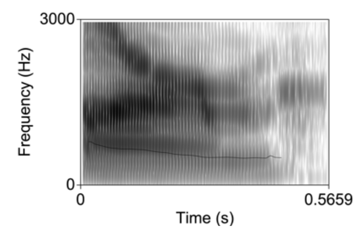
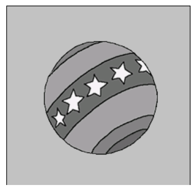
Deviant versions of the recorded (i.e., standard) words were generated by transposing the fundamental frequency (f_0) of the words spoken by the female talker (mean standard f_0 of 198 Hz) and the male talker (mean standard f_0 of 122 Hz) by +3 semitones (ST; a ~19% increase) using Adobe Audition. The resulting average f_0 of the deviant words was 235 Hz for the female talker and 145 Hz for the male talker. Other acoustic aspects of the words (e.g., duration;

intensity) were unaffected by the transposition, ensuring the deviant words differed from the standard words only with regard to f_0 (i.e., pitch). **Table 2.3** displays examples of stimuli used in the auditory change detection task.

Visual Stimuli

Stimuli for the visual change detection task consisted of 106 grayscale line-drawn images of nouns originally developed by Snodgrass & Vanderwart (1980) and determined to be within the lexicon of children as young as five years of age (Cortese & Khanna, 2008; Kuperman et al., 2012). The nouns included as stimuli in the visual change detection task were distinct from those used in the auditory change detection task to eliminate the possible influence of familiarity or learning on performance. Each image was superimposed on a 5” × 5” light gray (30% saturation) background using Adobe Photoshop and exported in .bmp image format. The same three young adults and two children who validated the stimuli for the auditory change detection task reviewed the final corpus of images to confirm nameability and visual quality.

Deviant versions of the original (i.e., standard) images were generated by increasing the saturation of the background by 40% (i.e., to 70%; dark gray) using Adobe Photoshop while leaving the images unaltered. **Table 2.3** displays examples of stimuli used in the visual change detection task.

	Standard	Deviant
Auditory Change Detection Task		
Visual Change Detection Task		

<u>Table 2.3</u>: Examples of stimuli used in the auditory and visual change detection tasks. *Top row*: spectrograms of the word "*arm*" as spoken by the male talker in the standard (*left column*) and deviant (*right column*) versions. F_0 contours for the standard and deviant words are indicated by thin black lines. *Bottom row*: images of the word "*ball*" in the standard (*left column*) and deviant (*right column*) versions.

2.3.4 Testing Apparatus

Testing occurred in a double-walled sound-attenuating booth with the child and a researcher present. Children were seated at a desk located in the center of the booth facing a Dell Ultrasharp U2413 24" monitor connected to an Apple wired USB keyboard. Genelec 8020c loudspeakers were positioned at +45 and -45 degrees azimuth relative to the location of the desk. A custom MATLAB program running on a MacBook Pro laptop located outside of the booth was used to control the randomization and presentation of auditory and visual stimuli inside of the booth. For the auditory change detection task, the recorded words spoken by the female and male talkers were processed through separate analog channels of a Focusrite Saffire PRO 14 audio interface, amplified, and presented at 65 dBA in the sound field via channel-specific loudspeakers positioned as described above. For the visual change detection task, the images were presented from the laptop to the external monitor via a Mini DisplayPort to HDMI connection. For both tasks, children indicated their responses by pressing the spacebar.

2.3.5 Procedures

Familiarization

At the beginning of the session, children completed a familiarization paradigm to ensure they could perceive and discriminate all key contrasts of the auditory and visual stimuli. During the auditory portion of the familiarization paradigm, children completed three blocks of 10 trials each, with each block assessing children's ability to discriminate a different auditory contrast (i.e., female vs. male talker; standard vs. deviant word; left vs. right speaker). Words from the female and male speakers were incorporated into the blocks gauging children's ability to distinguish the standard-deviant contrasts and speaker locations. During the visual portion of the familiarization paradigm, children completed two blocks of 10 trials each, with each block assessing children's ability to discriminate a different visual contrast (i.e., left vs. right side; standard vs. deviant image). The instructions provided by the researcher for each block of the familiarization paradigm are outlined in **Table 2.4**. Children's understanding of the instructions for each block was confirmed prior to proceeding and reinstruction was provided, as needed. If children provided an incorrect response on any given trial, no feedback was provided and the trial was repeated at the end of the block. All children achieved a performance criterion of 100% accuracy on all blocks of the familiarization paradigm prior to completing the experimental tasks.

Block	Domain	Contrast	Instructions
1	Auditory	Female vs. Male	*"Listen to each word. Tell me whether it was spoken by a female or a male."*
2	Auditory	Standard vs. Deviant	*"Listen to each pair of words. Tell me whether they sound the same or different."*
3	Auditory	Left vs. Right	*"Listen to each word. Tell me whether it was presented from the left or right speaker."*
4	Visual	Left vs. Right	*"Look at each image. Tell me whether it was presented on the left or right side of the screen."*
5	Visual	Standard vs. Deviant	*"Look at each pair of images. Tell me whether they look the same or different."*

Table 2.4: Familiarization paradigm completed by all children prior to testing.

Adaptive Auditory and Visual Tasks

To minimize differences in task performance due to developmental factors beyond executive function (e.g., speed of processing, planning and execution of motor movements), individualized rates of stimulus presentation were measured for each child. Specifically, children performed adaptive single-stream versions of the auditory and visual change detection tasks to establish the interstimulus interval (i.e., ISI; the duration between the offset of a stimulus and the onset of the subsequent stimulus in ms) at which they were able to reliably identify deviants (i.e., $ISI_{Auditory}$; ISI_{Visual}). Children were presented with a single stream of images, or of words spoken by the talker of the sex that would be designated as the target stream during the test blocks. They were instructed to press the spacebar as quickly as possible whenever they detected a deviant image or word. Deviants were presented randomly with an approximate occurrence of 10% and at least five consecutive trials containing standard stimuli following a trial containing a deviant.

The adaptive auditory and visual tasks were loosely modeled on the transformed staircase method described by Levitt (1971). While Levitt (1971) describes a "3-down 1-up" tracking procedure whereby a parameter, such as the ISI, is reduced following a correct response on three consecutive trials at the same stimulus level and increased following each single incorrect response, the adaptive procedure employed in the present study utilized an expedited tracking procedure to make the task more child friendly. Specifically, the task began at an ISI of 500 ms with a descending run whereby each correct response (i.e., hit) led to a 50-ms decrease in the ISI until an incorrect response (i.e., miss) occurred, at which point the ISI increased by 20 ms. After this initial reversal, if applicable (i.e., if an incorrect response occurred), three consecutive correct responses at the same ISI were required to begin another descending run of trials for which the step size decreased to 10 ms. Another reversal (i.e., a 10-ms increase in ISI) occurred after the next incorrect response, if applicable. The adaptive procedure continued following this 3-correct-1-incorrect staircase method until either 7 reversals had been made or a descending run resulted in a minimum ISI of 0 ms (i.e., no silent gap between words). If the child did not achieve an ISI of 0 ms, the $ISI_{Auditory}$ and ISI_{Visual} were calculated by averaging the ISI at which the last five reversals occurred. **Figure 2.2** displays **e**xamples of adaptive tracks for an ISI of 0 ms and an ISI of 320 ms.

Children's individual $ISI_{Auditory}$ and ISI_{Visual} values served as the ISI for the dual-stream (i.e., test) versions of the auditory and visual change detection tasks, respectively. Fifteen children were unable to reliably complete the adaptive single-stream versions of the auditory (*N* = 3 with NH; *N* = 3 with HL), visual (*N* = 2 with NH; *N* = 1 with HL), or auditory and visual (*N* = 5 with NH; *N* = 1 with HL) change detection tasks due to non-compliance and/or task difficulty. For these children, an ISI of 300 ms was used for the auditory and/or visual change detection tasks – the approximate average $ISI_{Auditory}$ and ISI_{Visual} for 6 pilot participants ranging from 6 to 12 years of age. **Figure 2.3** displays the relation between $ISI_{Auditory}$ and ISI_{Visual} for individual children.

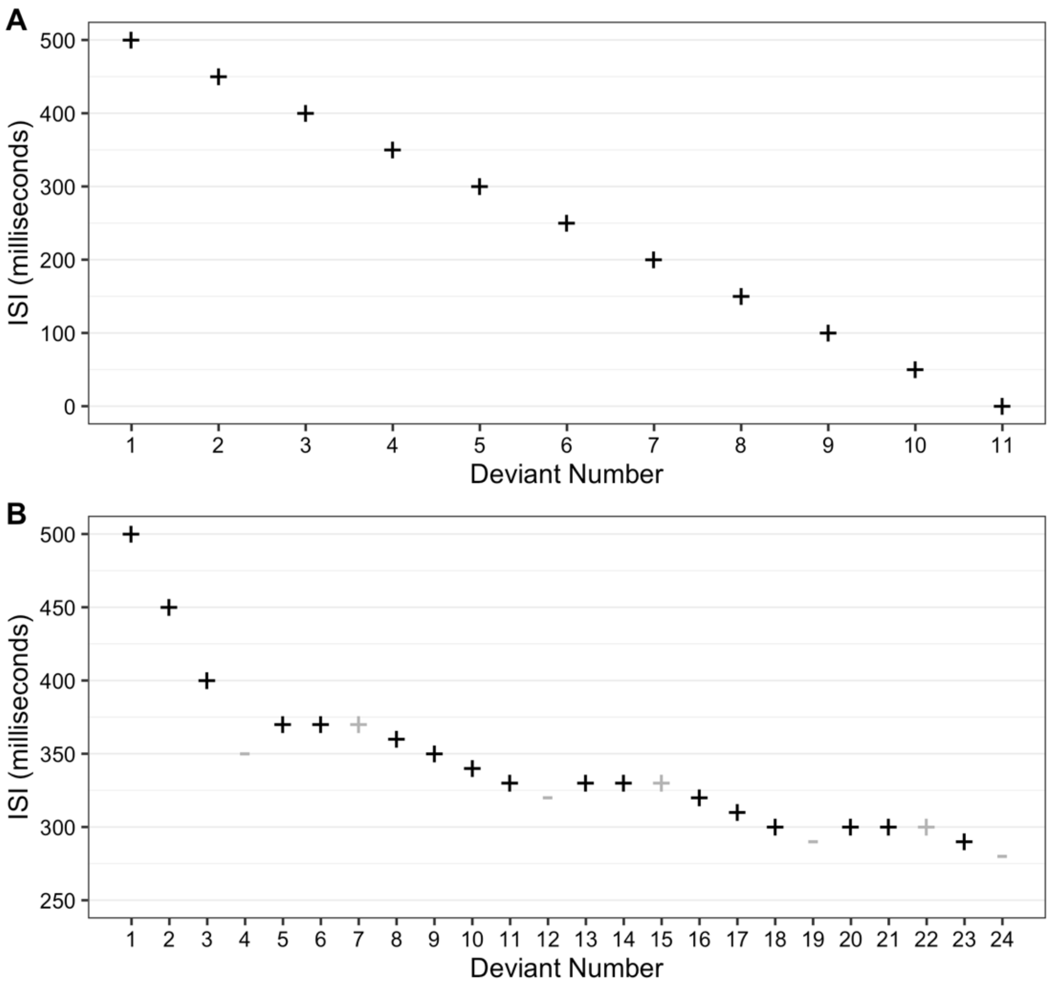
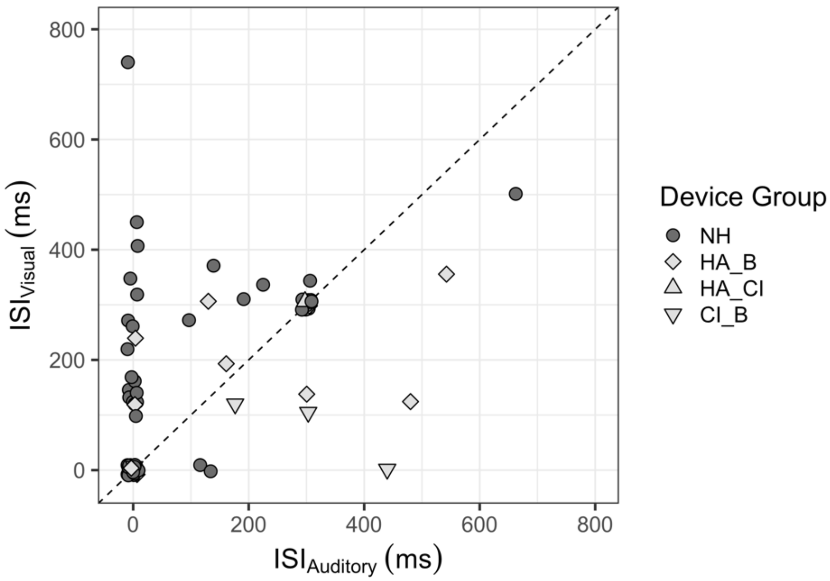


Figure 2.2: Example performance on the adaptive auditory and visual tasks. Adaptive tracks resulting in ISIs of 0 ms (**A**) and 320 ms (**B**). Correct (+) and incorrect (-) responses are displayed with reversals (*light gray*) indicating a change in ISI. The ISIs of the last five reversals were averaged to generate the ISI for each child and task.



<u>Figure 2.3</u>: Relation between adaptive auditory and visual task ISIs. Each child is represented by a single point. Points positioned on the dashed line indicate that the measured ISIs were the same across the auditory and visual tasks. Points falling above or below the dashed line indicate that the ISI_{Visual} was longer than the $ISI_{Auditory}$ or the $ISI_{Auditory}$ was longer than the ISI_{Visual}, respectively. The cluster of points at 300ms includes the children for which adaptive ISIs was unable to be obtained. Points have been slightly jittered to aid the visualization of overlapping points.

Auditory Change Detection Task

Using a behavioral dual-stream oddball paradigm, the auditory change detection task assessed children's ability to selectively attend to a target auditory stream and inhibit attention to a competing auditory stream. Children completed four blocks of the auditory change detection task, each consisting of 150 trials for a total of 600 trials. On each trial, a word spoken by the female talker and a word spoken by the male talker were presented simultaneously from opposite speakers (i.e., female talker from +45 degrees and male talker from -45 degrees or vice versa). It was expected that providing redundant frequency (i.e., female vs. male talker) and spatial (i.e., left vs. right speaker) cues facilitated children's ability to form two distinct auditory streams as the words were presented sequentially across trials (e.g., Sussman et al., 2001). The location from which the female and male talkers were presented was counterbalanced across children.

Consistent with the design of a standard oddball paradigm, each block consisted of deviant words presented with an approximate occurrence of 10% across the target and distractor streams (i.e., 5% occurrence in each stream). Specifically, in any given block of 600 trials, approximately 60 deviants were presented in total: 30 in the target stream and 30 in the distractor stream. The remaining 90% of trials across both streams consisted of standard words. The trials during which a deviant word was presented in either the target or distractor stream was randomized via MATLAB, with the following constraints applied: 1) the number of deviants presented in the target and distractor streams must be approximately equal across the four blocks; 2) a deviant may not be presented in both the target and distractor streams during the same trial; and 3) at least five consecutive trials containing standard stimuli must follow a trial containing a deviant, regardless of whether it was presented in the target or distractor stream. **Figure 2.4 A** displays a schematic of the auditory change detection task.

At the onset of the task, the researcher designated a target and distractor stream based on the sex and location of the talker (e.g., "*Pay attention to the woman's voice coming from the right speaker and ignore the man's voice coming from the left speaker*"). The designated sex and location of the target and distractor streams were counterbalanced across children. Children were instructed to press the spacebar when they detected a deviant word in the target, but not distractor, stream (e.g., "*Press the spacebar as quickly as possible when you hear the woman's voice get higher in pitch. If you hear the man's voice get higher in pitch, do nothing.*"). Children's understanding of the instructions was confirmed by ensuring they could correctly answer questions posed by the researcher about the sex and location of the talker comprising the target and distractor streams (e.g., "*Which talker will you pay attention to*?"; "*Where will that talker's voice come from?*") as well as the method of their response (e.g., "*What will you do when that talker's voice gets higher in pitch?*"). Clarification and reinstruction was provided as needed.

All children, including those with hearing loss, confirmed that the presentation level of the words was audible and comfortable prior to beginning the task. After each block of 150 trials, a cartoon image was presented on the screen until children opted to press the spacebar to proceed to the next block. Children had the option to take breaks between blocks, if needed, though the majority of children decided not to do so. Additional breaks were provided in between tasks, during which children collected a ticket for completing the previous task.

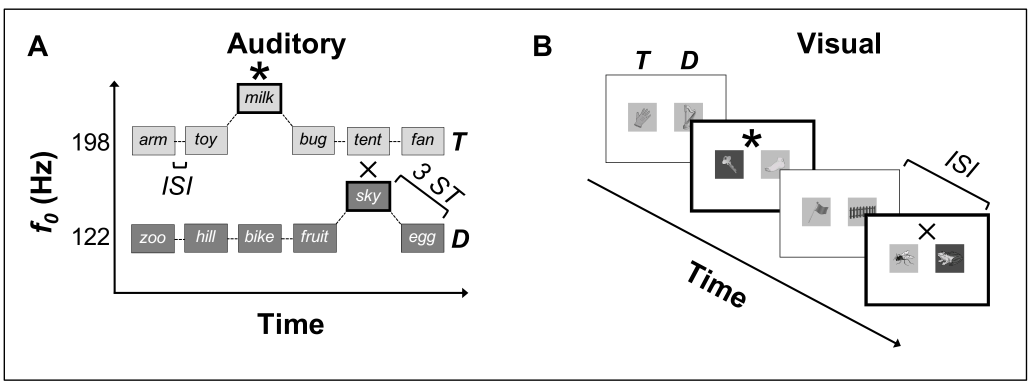


Figure 2.4: **Schematics of the auditory and visual change detection tasks.** Children were instructed to press a key in response to deviants presented in the target stream (*) but not the distractor stream (✕).

Visual Change Detection Task

A dual-stream visual change detection task was used to assess children's ability to selectively attend to a target stream and inhibit attention to a distractor stream in the visual domain. The structure of the visual change detection task (i.e., number of trials and blocks, standard and deviant occurrence rates and presentation restraints) was analogous to that of the auditory change detection task. On each trial, two images were presented simultaneously – one image on the left side of the screen and one image on the right side of the screen – for a duration of 200 ms. It was expected that providing a spatial separation of the images facilitated children's ability to form two distinct visual streams. **Figure 2.4 B** displays a schematic of the visual change detection task. Whether children completed the auditory or visual change detection task first was counterbalanced across participants.

Quantifying Performance on the Auditory and Visual Change Detection Tasks

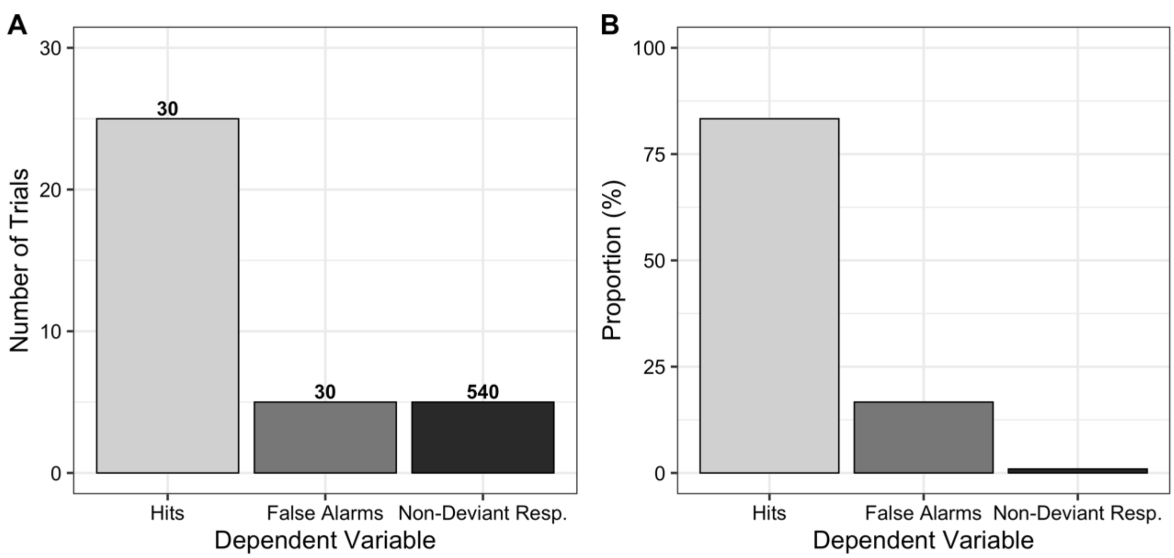
Children's performance on the auditory and visual change detection tasks was quantified by measuring the reaction time and accuracy of responses on each trial. Each trial was classified as one of three trial types that resulted in separate expected outcomes and contributed to different dependent variable(s), as displayed in **Table 2.5**. Each response was scored according to the type of trial during which it occurred. $Deviant_T$-$Standard_D$ trials consisted of a deviant stimulus in the target stream (*T*) and a standard stimulus in the distractor stream (*D*) with an expected response due to the presence of a deviant stimulus in the target stream (approximate $N = 30$ trials; range = 22-37 trials). Responses that occurred during $Deviant_T$-$Standard_D$ trials within the 100-2000 ms window following a deviant stimulus were classified as *hits* to the target stream, which were quantified by the reaction time and proportion of these responses relative to the number of deviants in the target stream across all trials. $Standard_T$-$Deviant_D$ trials consisted of a standard stimulus in the target stream and a deviant stimulus in the distractor stream with no expected response due to the absence of a deviant stimulus in the target stream (approximate $N = 30$ trials; range = 23-38 trials). Responses that occurred during $Standard_T$-$Deviant_D$ trials within the 100-2000 ms window following a deviant stimulus were classified as *false alarms*, which were quantified by the reaction time and proportion of these responses relative to the number of deviants in the target stream across all trials. Lastly, Standard$_T$-Standard$_D$ trials consisted of a standard stimulus in the target and distractor streams with no expected response due to the absence of a deviant stimulus in the target stream (approximate $N = 540$ trials; range = 525-555 trials). Responses that occurred during Standard$_T$-Standard$_D$ (i.e., non-deviant) trials and were outside the 100-2000 ms window following a deviant stimulus in either the target or distractor streams were classified as *non-deviant responses*, which were quantified as the raw number of these responses.

	Trial Type		
	Deviant$_T$-Standard$_D$	**Standard$_T$-Deviant$_D$**	**Standard$_T$-Standard$_D$**
Target Stream Stimulus	Deviant	Standard	Standard
Distractor Stream Stimulus	Standard	Deviant	Standard
Expected Outcome	Response	No Response	No Response
Dependent Variables	Hits	False Alarms	Non-Deviant Responses
	Reaction Time for Hits	Reaction Time for False Alarms	
	Response Sensitivity		

<u>**Table 2.5**</u>**: Trial types comprising the auditory and visual change detection tasks.** Target and distractor stream stimuli, expected outcomes, and resulting dependent variables are specified for each trial type.

In order to address the primary research question of whether children with HL demonstrated a reduced ability to selectively attend to a target stream relative to their peers with NH, performance was additionally quantified by *response sensitivity*, which reflects children's ability to selectively attend to deviants in the target stream (i.e., hit rate; proportion of responses during Deviant$_T$-Standard$_D$ trials) and inhibit attention to those in the distractor stream (i.e., false alarm rate; proportion of responses during Standard$_T$-Deviant$_D$ trials). Response sensitivity was calculated using the signal detection theory formula for d' (i.e., z[Hit Rate] – z[False Alarm Rate]; Green & Swets, 1966).

It is worth noting that, by design, oddball paradigms contain a significantly greater number of standard stimuli than deviant stimuli; therefore, the ratio of the number of non-deviant responses to the number of non-deviant trials tends to be quite small. This is problematic when comparing the frequency of various types of "undesirable" responses. For instance, even if children responded an equal number of times during $Standard_T$-$Deviant_D$ trials (i.e., false alarms) and $Standard_T$-$Standard_D$ trials (i.e., non-deviant responses), the discrepancy in the number of trials contributing to the denominator of each ratio would result in the proportion of non-deviant responses being substantially reduced relative to the proportion of false alarms (**Figure 2.5**). For this reason, non-deviant responses were represented as raw numbers rather than proportions during analysis.



<u>Figure 2.5</u>**: Comparison of raw number of responses versus proportion of responses for hits, false alarms, and non-deviant responses in a sample test block.** Bold numbers above each bar indicate the total number of trials that contributed to each dependent variable (**A**), which served as the denominator when calculating the proportions for each dependent variable in **B**.

2.3.6 Statistical Analyses

Linear Quantile Mixed-Effects Modeling

Linear quantile mixed-effects modeling was employed to assess the predicted relations between performance on the auditory and visual change detection tasks and age for children with NH and HL. As variables with non-normally distributed residuals violate the assumptions for traditional parametric statistical methods, linear quantile mixed-effects modeling – a non-parametric statistical method of analyzing the effect of a set of fixed and random factors on the entire conditional distribution of the outcome – was utilized to assess the relations among age, hearing status, and children's performance on the auditory and visual change detection tasks. Modeling was executed via the 'lqmm' package for R Statistical Software (Geraci, 2014; Geraci & Bottai, 2014; R Core Team, 2019). Similar to parametric linear mixed-effects modeling, linear quantile mixed-effects modeling operates under a linear regression framework, allows for the inclusion repeated measures, and incorporates both fixed and random effects. Unlike parametric linear mixed-effects modeling that compares variation in outcome distribution *means* for fixed effects with random intercepts, linear quantile mixed-effects modeling considers the shape of the outcome distribution in order to estimate conditional quantile functions (e.g., *median*) for fixed effects with random intercepts and/or slopes. Specifically, linear quantile mixed-effects models use an asymmetric Laplace distribution (Hinkley & Revankar, 1977) for maximum likelihood methods and are derived using a skewness parameter (n) that is established *a priori* to define the n^{th} conditional quantile to be estimated. The dependence on the shape of the distribution via preestablished quantiles as opposed to a specific location of the distribution determined by the mean makes linear quantile mixed-effects modeling more appropriate for use with non-normally distributed outcome residuals. Parameter inference was based on a block-bootstrapping method with 50 replications.

To investigate the predicted effects of age and hearing status on children's ability to selectively attend to a target stream in the auditory and visual domains, linear quantile mixed-effects models using the 0.5 quantile (i.e., 50^{th} percentile; median) as the comparison statistic were conducted separately for each dependent variable (i.e., hits [proportion]; false alarms [proportion]; response sensitivity; non-deviant responses [number]; reaction time for hits [in ms]; and reaction time for false alarms [in ms]). Prior to analysis, age was centered to allow for a more meaningful interpretation of model intercepts by subtracting the median age from each observed value. The fixed-effect predictor variables included in each model were age (continuous), group (NH; HL); and task (ACD; VCD). Participant was included as a random factor in each model to account for shared variance due to similarities in performance within a child across tasks. All possible interaction terms were included in each model to test the *a priori* predictions described above.

For each model, the *estimate* refers to the median performance at the centered (i.e., median) age for the intercept term, the change in the slope of median performance for continuous predictor variables (i.e., age), and the difference in median outcome relative to the reference levels for predictor variables with at least two levels (i.e., group; task). For interaction terms, the estimate reflects the difference of the median differences in outcome between the levels of the variables included in the interaction. Treatment contrasts (i.e., dummy codes) were utilized and reference levels were set to children with NH and the ACD task. As such, all estimates, except those reported for paired comparisons, reflect the performance of children with HL relative to children with NH and performance during the VCD task relative to the ACD task. In addition to information pertaining to the model estimates, the *Akaike Information Criterion (AIC)* value provides an estimate of the quality of each model relative to other models with the same dependent variable, with lower values indicating a better model fit for the purposes of model comparison (Akaike, 1974).

In addition to the main models described above, linear quantile mixed-effects modeling was used to assess the contribution of specific individual factors, including unaided and aided audibility and executive function, to performance on the auditory and visual change detection tasks. Specifically, if hearing loss during childhood alters children's ability to selectively attend to a target stream, it is expected that children's hearing sensitivity and performance on standardized measures of executive function will relate to their performance during the auditory and visual change detection tasks. A sequential model-building approach was used to test these relations. Response sensitivity served as the dependent variable for these models, as it accounts for children's responses to deviants in the target and distractor streams. The specific fixed-effect predictor variables included in each model were determined by a combination of *a priori* predictions and *post hoc* correlations, which are described in more detail below with participant included as a random factor. All continuous variables were centered prior to analysis by subtracting the median value from each observed value to allow for a more meaningful interpretation of model intercepts. Interactions were only included to test whether the influence of hearing sensitivity and executive function abilities differed across the auditory and visual domains.

2.4 RESULTS

2.4.1 Effects of Age, Hearing Status, and Task

Table 2.6 displays performance on the auditory and visual change detection tasks for children with NH and HL.

	Group	
	Children with NH	**Children with HL**
Auditory		
Hits (%)	Mean = 85.94 | *SD* = 13.92 Range = 41.94-100	Mean = 64.92 | *SD* = 23.76 Range = 6.45-92
False Alarms (%)	Mean = 4.35 | *SD* = 7.53 Range = 0-38.71	Mean = 7.94 | *SD* = 6.05 Range = 0-19.44
Response Sensitivity	Mean = 3.22 | *SD* = 0.94 Range = 0.76-4.72	Mean = 1.96 | *SD* = 0.92 Range = -0.57-3.31
RT for Hits (ms)	Mean = 864.67 | *SD* = 195.04 Range = 520.22-1330.70	Mean = 1031.67 | *SD* = 251.94 Range = 691.20-1679.60
RT for False Alarms (ms)	Mean = 884.61 | *SD* = 396.32 Range = 162.59-1915.90	Mean = 1104.40 | *SD* = 413.70 Range = 573.40-1790.90
Non-Deviant Responses	Mean = 4.10 | *SD* = 4.41 Range = 0-20	Mean = 10.69 | *SD* = 6.92 Range = 2-28
Visual		
Hits (%)	Mean = 80.62 | *SD* = 16.75 Range = 25-100	Mean = 68.03 | *SD* = 23.45 Range = 25.81-100
False Alarms (%)	Mean = 4.79 | *SD* = 4.04 Range = 0-17.24	Mean = 7.65 | *SD* = 10.83 Range = 0-38.10
Response Sensitivity	Mean = 2.78 | *SD* = 0.72 Range = 1.19-4.29	Mean = 2.33 | *SD* = 1.21 Range = 0.05-4.72
RT for Hits (ms)	Mean = 674.89 | *SD* = 168.42 Range = 418.51-1117.80	Mean = 675.47 | *SD* = 209.67 Range = 321.27-1010.20
RT for False Alarms (ms)	Mean = 817.20 | *SD* = 369.67 Range = 154.65-1860.7	Mean = 913.32 | *SD* = 621.29 Range = 156.87-1919.10
Non-Deviant Responses	Mean = 1.49 | *SD* = 1.99 Range = 0-12	Mean = 4.13 | *SD* = 7.53 Range = 0-25

Table 2.6: Descriptive statistics for performance on the auditory and visual change detection tasks for children with normal hearing and hearing loss.

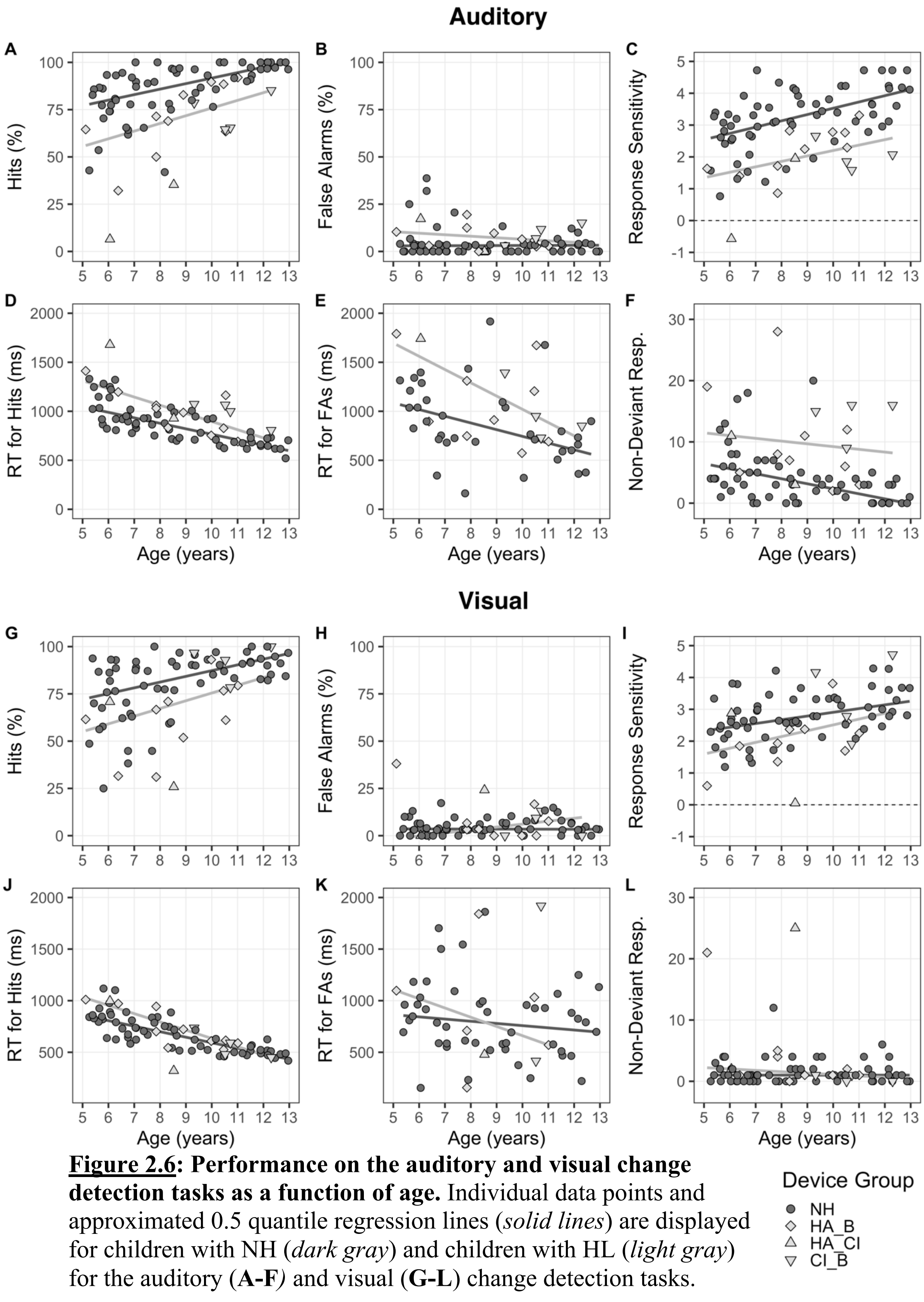
As described in *Section 2.3.6*, linear quantile mixed-effects models were used to examine how age and hearing loss influence children's ability to selectively attend to a target stream in the auditory and visual domains. A total of six models were executed – one for each dependent variable of interest. Three-way interaction terms were initially included in the model for each dependent variable. The purpose of including the three-way interaction terms was to test whether observed differences in children's performance on the auditory and visual change detection tasks differed across age, group, and task. As these interactions were found to be non-significant (see *Appendix A.2.1* for full model output), they were removed from the final models to increase parsimony and improve the interpretability of coefficients for the main effects and two-way interactions of interest. *Appendix A.2.2* displays the remaining main effects and interactions for each dependent variable of interest (**A-F**), which are described in more detail below.

Effect of Age

A primary objective of the present study was to investigate the effect of immaturity on selective attention during childhood. As predicted, significant improvements in performance were observed with increasing age during the auditory and visual change detection tasks (**Figure 2.6**; *Appendix A.2.2*). Specifically, there were age-related improvements in response sensitivity, with older children responding more selectively to deviants in the target stream than younger children (estimate = 0.187, $p < 0.001$; **Figure 2.6 C & I**). Age-related improvements in performance on the auditory and visual change detection tasks were also observed for hits (estimate = 0.034, $p < 0.001$; **Figure 2.6 A & G**), false alarms (estimate = -0.004, $p < 0.05$; **Figure 2.6 B & H**), non-deviant responses (estimate = -0.427, $p < 0.05$; **Figure 2.6 F & L**). reaction time for hits (estimate = -59.013, $p < 0.001$; **Figure 2.6 D & J**), and reaction time for false alarms (estimate = -69.117, $p < 0.001$; **Figure 2.6 E & K**). Together, these findings suggest that, as hypothesized, children's ability to selectively attend to the target stream improved with age, with older children responding more accurately and with fewer non-deviant responses than younger children.

While these findings support the hypothesis that selective attention improves during childhood, the effect of age differed across the auditory and visual change detection tasks for non-deviant responses and reaction time for false alarms, as indicated by a significant age-by-task interaction for non-deviant responses (estimate = 0.427, $p < 0.05$) and a marginally significant age-by-task interaction for reaction time for false alarms (estimate = 53.636, $p = 0.058$). Specifically, paired comparisons revealed that the slope of the relation between age and non-deviant responses was less negative (i.e., more shallow) for the visual change detection task (estimate = -0.00006, $p = 0.999$) than the auditory change detection task (estimate = -0.427, $p < 0.05$), though only the latter relation was significant. Visual inspection of the data in **Figure 2.6 F & L** suggests that these observed differences may be due to the overall fewer non-deviant responses in the visual change detection task relative to the auditory change detection task, which will be discussed in more detail in the subsequent section. Additionally, paired comparisons revealed that the slope of the relation between age and reaction time for false alarms was less negative (i.e., more shallow) for the visual change detection task (estimate = -15.472, $p = 0.415$) relative to the auditory change detection task (-69.117, $p < 0.001$), though only the latter relation was significant.

Lastly, the extent to which the observed age-related improvements in performance differed between children with normal hearing and children with hearing loss was assessed. While visual inspection of the data in **Figure 2.6** suggests that, for some dependent variables, age may differentially affect performance for children with NH and HL, no significant age-by-group interactions were observed. These findings indicate that the slope of the change in median performance across age for each dependent variable, as described above, does not significantly differ between children with NH and HL. Notably, the relatively small sample of children with HL may limit the ability to detect marginal differences in the magnitude of relations between groups across the tested age range.



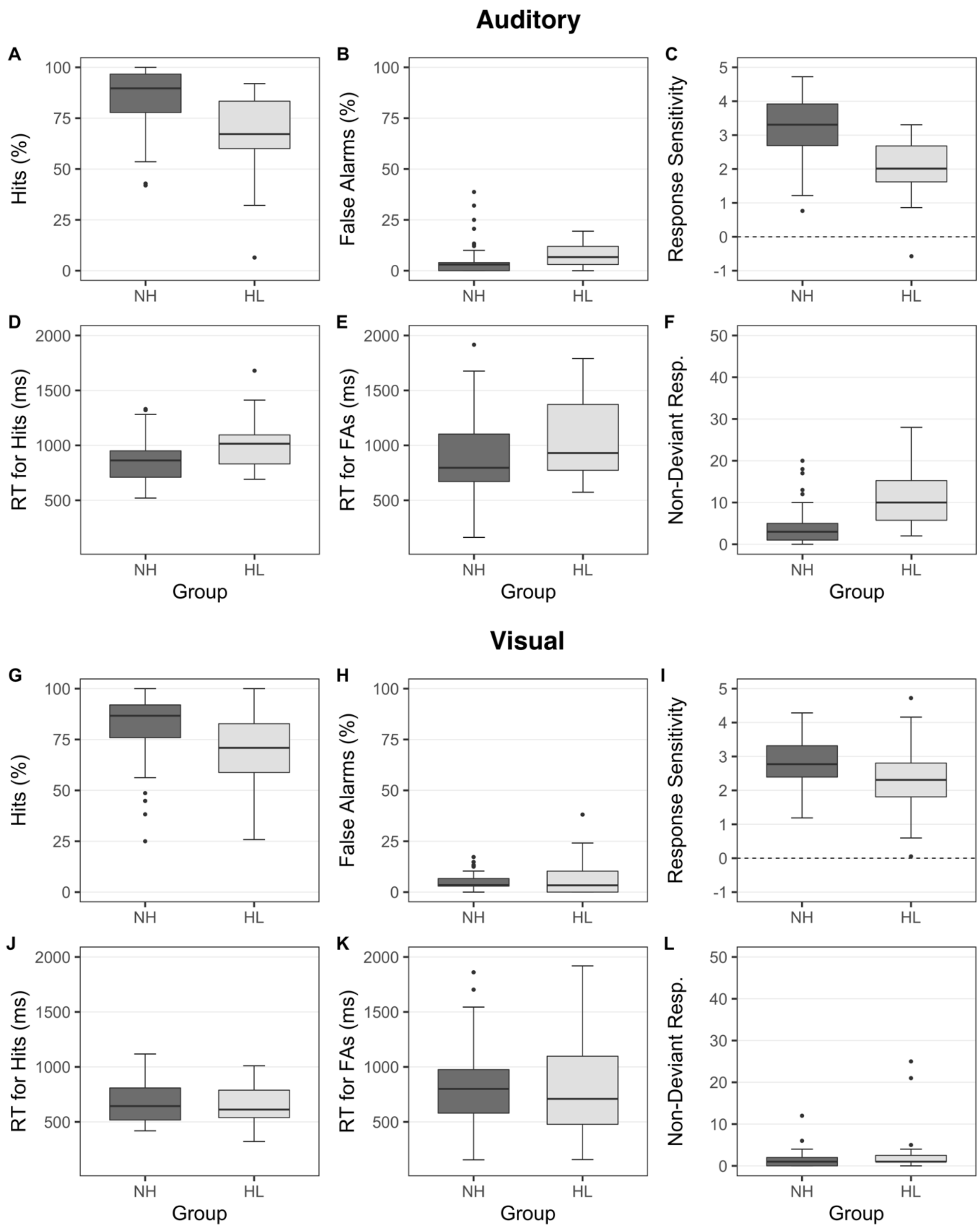
<u>Figure 2.6</u>: Performance on the auditory and visual change detection tasks as a function of age. Individual data points and approximated 0.5 quantile regression lines (*solid lines*) are displayed for children with NH (*dark gray*) and children with HL (*light gray*) for the auditory (**A-F***)* and visual (**G-L**) change detection tasks.

Effects of Hearing Status and Task

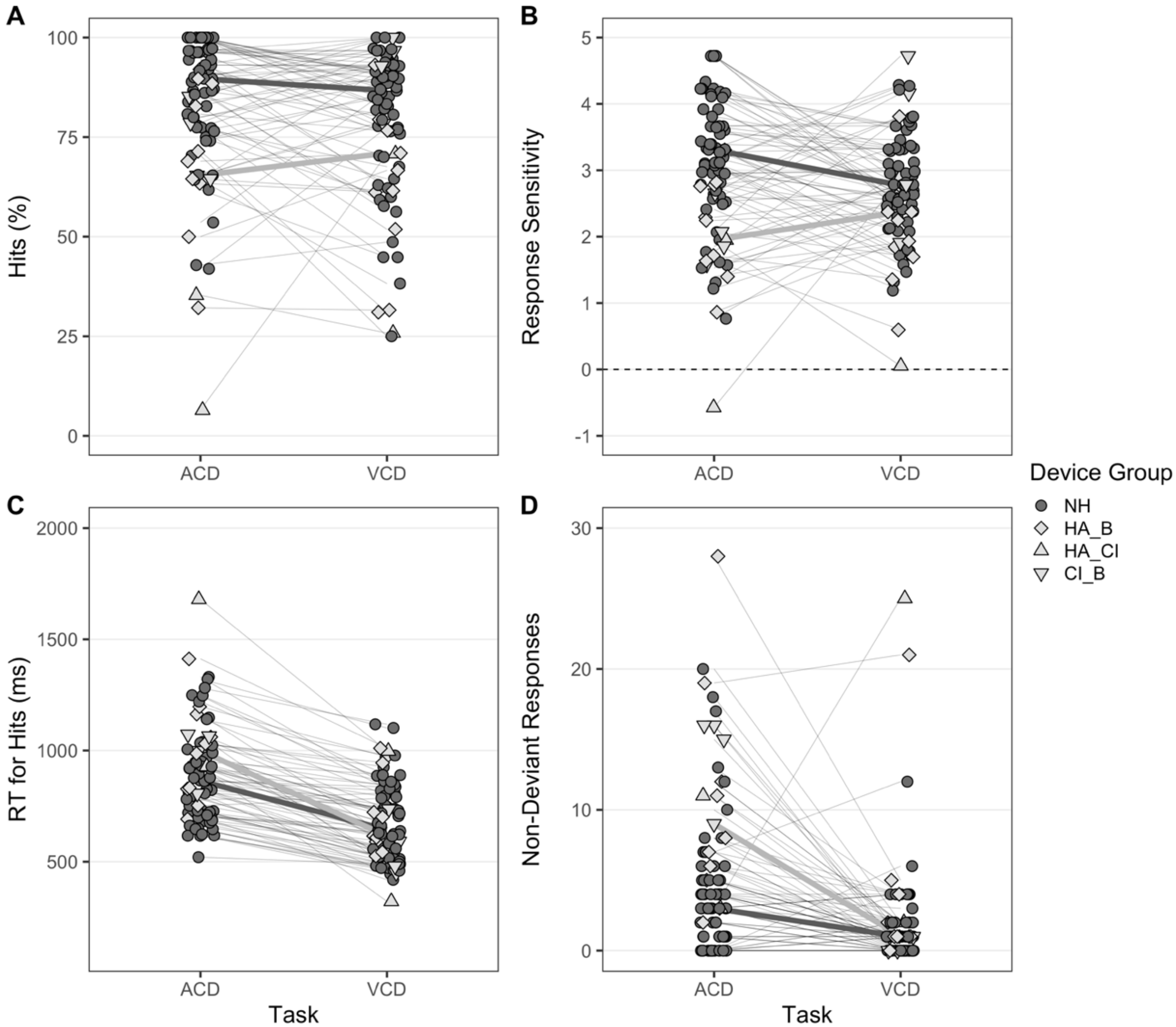
In addition to investigating how children's age affects selective attention, the present study aimed to test the hypothesis that disrupted auditory experience impedes selective attention during childhood. As predicted, significant differences in performance were observed between children with NH and HL. Specifically, significant improvements in response sensitivity were observed such that children with NH demonstrated higher response sensitivity than children with HL during the auditory and visual change detection tasks (estimate = -1.258, $p < 0.001$; **Figure 2.7 C & I**). In addition, significant differences in performance between groups were observed for hits (estimate = -0.242, $p < 0.001$; **Figure 2.7 A & G**), false alarms (estimate = 0.042, $p < 0.05$; **Figure 2.7 B & H**), non-deviant responses (estimate = 6.515, $p < 0.001$; **Figure 2.7 F & L**), reaction times for hits (estimate = 211.065, $p < 0.001$; **Figure 2.7 D & J**), and reaction time for false alarms (estimate = 359.625, $p < 0.01$; **Figure 2.7 E & K**) with children with NH performing better overall. Together, these findings suggest that, as hypothesized, children with HL had greater difficulty selectively attending to deviants in the target stream and inhibiting attention to a competing stream of input compared to children with NH.

If, as hypothesized, hearing loss during childhood alters domain-general selective attention, the difference in performance between children with NH and HL should be consistent across the auditory and visual change detection tasks. Contrary to this hypothesis is the observation that children's performance differed across the auditory and visual change detection tasks for response sensitivity (estimate = -0.491, $p < 0.001$; **Figure 2.8 B**) as well as for hits (estimate = -0.053, $p < 0.05$; **Figure 2.8 A**), reaction time for hits (estimate = -180.296, $p < 0.001$; **Figure 2.8 C**), and, as previously mentioned, non-deviant responses (estimate = -2.182, $p < 0.001$; **Figure 2.8 D**). Additionally, the observed differences in performance across the auditory and visual change detection tasks were inconsistent for children with NH and HL, which is supported by the presence of significant or trending-toward-significant group-by-task interactions for hits (estimate = 0.084, $p = 0.065$), response sensitivity (estimate = 0.700, $p = 0.054$), reaction time for hits (estimate = -171.034, $p < 0.001$), and non-deviant responses (estimate = -6.051, $p < 0.05$). Pairwise comparisons revealed that children with NH demonstrated a significantly lower proportion of hits (estimate = -0.053, $p < 0.05$; **Figure 2.8 A**) and lower response sensitivity (estimate = -0.491, $p < 0.001$; **Figure 2.8 B**) during the visual change detection task relative to the auditory change detection task while children with HL demonstrated no change in the proportion of hits (estimate = 0.037, $p = 0.566$) or response sensitivity (estimate = 0.174, $p = 0.645$) between the auditory and visual change detection tasks.

The presence of a significant group-by-task interaction also raises a question about the extent to which the observed differences in performance between children with NH and children with HL differed based on the task domain. Specifically, if group differences in performance solely reflect the hypothesized influence of hearing loss on selective attention – a domain general cognitive process – then these group differences should be observed to the same extent regardless of task domain. However, this was not found to be the case. Pairwise comparisons revealed that children with HL demonstrated significantly fewer hits and lower response sensitivity than children with NH during the auditory and visual change detection tasks with these differences being slightly greater in the auditory domain (estimate[hits] = -0.242, $p < 0.001$; estimate[response sensitivity] = -1.258, $p < 0.001$) than the visual domain (estimate[hits] = -0.158, $p < 0.01$; estimate[response sensitivity] = -0.571, $p < 0.05$; **Figure 2.8 A & B**). Consistent with this outcome, significant differences in performance between children with NH and HL were observed for reaction time for hits and non-deviant responses during the auditory change detection task (estimate[reaction time] = 211.065, $p < 0.001$; estimate[non-deviant responses] = 6.515, $p < 0.001$) but not the visual change detection task (estimate[reaction time] = 38.803, $p = 0.347$; estimate[non-deviant responses] = 0.474, $p = 0.783$; **Figure 2.8 C & D**). Together, these findings suggest that differences in performance between children with NH and HL tended to be greater during the auditory change detection task than the visual change detection task.



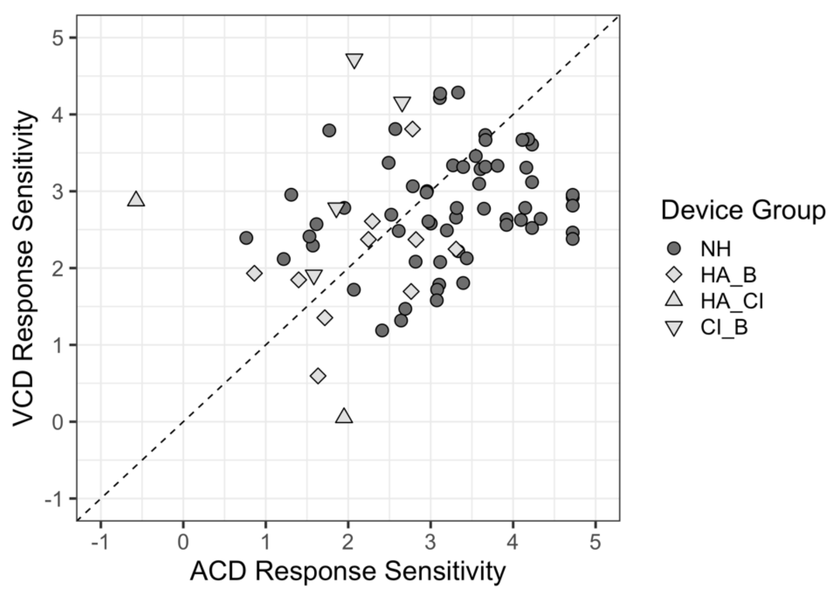
<u>Figure 2.7</u>: Boxplots of children's performance on the auditory and visual change detection tasks grouped by hearing status. Median performance (*black line*) and interquartile range is plotted for children with NH (*dark gray*) and children with HL (*light gray*) for the auditory (**A-F***)* and visual (**G-L**) change detection tasks.



<u>Figure 2.8</u>: Differences in performance between the auditory and visual change detection tasks. Individual differences (*thin lines*) and 0.5 quantile regression lines (*thick lines*) for children with NH (*dark gray*) and children with HL (*light gray*) for hits (**A**), response sensitivity (**B**), reaction time for hits (**C**), and non-deviant responses (**D**).

The observation that group differences in performance varied across task domains indicates the absence of a direct 1:1 correspondence between children's performance on the auditory and visual change detection tasks. A correlation analysis was used to quantify this relation (**Figure 2.9**). For response sensitivity – the primary dependent variable of interest –

there was a significant, albeit modest, positive linear relation between children's performance on the auditory and visual change detection tasks (ρ = 0.33, $p < 0.01$). This finding indicates that, as hypothesized, children who had greater difficulty selectively attending to the target stream in the auditory domain tended to also do so in the visual domain. The absence of a strong correlation between tasks in conjunction with the observed group and task differences described above, however, suggests that the ability of children with NH and HL to meet the attentional demands of the tasks may have differed across domains.



<u>Figure 2.9</u>: Spearman correlation between response sensitivity during the auditory and visual change detection tasks. Individual children's response sensitivity during the visual change detection task positively related to their response sensitivity during the auditory change detection task (ρ = 0.33, $p < 0.01$).

2.4.2 Relation between Hearing Sensitivity and Performance

The observation that children with HL demonstrated poorer selective attention during the auditory and visual change detection tasks than children with NH motivates the exploration of whether children's hearing sensitivity accounted for these group differences. If the quality and quantity of auditory experience influences the development of selective attention during childhood, it is possible that children's hearing sensitivity with and without the use of their hearing devices relates to their task performance. Additional linear quantile mixed-effects models with response sensitivity as the dependent variable were executed to assess these relations. If the observed group differences in response sensitivity are explained by underlying differences in hearing sensitivity between children with NH and HL, unaided PTA should statistically account for the effect of group. This is expected as groups were established based on children's hearing status, which is defined by unaided PTA. If, alternatively, the observed group differences additionally reflect insufficient audibility of experimental stimuli during testing for children with HL despite the use of their hearing devices, the inclusion of aided PTA in the model should also statistically account for the effect of group.

To test these hypotheses, aided and unaided PTA were separately added as predictor variables to the model with response sensitivity as the dependent variable. For children with normal hearing, unaided PTA values were used as the predictor variable for both models as an "aided" condition is not applicable for this group. To account for differences across ears within a child, the right and left ear PTAs in each condition were averaged to generate an average unaided and aided PTA for each child. Due to the strong correlation between average unaided PTA and average aided PTA ($r = 0.89$, $p < 0.001$) and associated risk for multicollinearity, these variables were added to the model sequentially rather than simultaneously. The predictor variables that comprised the base model for this analysis included age, group, and task, which were all found to significantly affect response sensitivity in the above analyses (*Appendix A.2.2* **C**).

Appendix A.2.3 displays the output of the base model and the models testing for the effects of aided and unaided PTA on the observed group differences in response sensitivity. As stated above, the primary objective of this analysis was to measure changes in the main effect of group across tasks with the addition of aided PTA or unaided PTA. In considering that the group-by-task interaction for response sensitivity was only marginally significant and paired comparisons revealed that differences in performance between children with NH and HL persisted across both tasks (*Appendix A.2.2*), the group-by-task interaction term was not included in this model. As a result, the reported estimates for group reflect overall trends across the auditory and visual change detection tasks with NH and ACD as the reference levels.

When neither aided nor unaided PTA were included in the model, the main effect of group was significant (i.e., estimate = -0.987, $p < 0.001$). When aided PTA was added to the model, the significant effect of group remained (estimate = -0.695, $p < 0.01$; *Appendix A.2.3* **B**). This finding suggests that the auditory conditions under which children completed the experimental tasks (i.e., unaided for children with NH; aided for children with HL) did not account for the observed group differences in response sensitivity. However, when unaided PTA was added to the base model, the significant effect of group was removed (estimate = -0.583, $p =$ 0.176; *Appendix A.2.3* **C**). This finding provides evidence that, as expected, group and unaided PTA accounted for significant overlapping variance in response sensitivity. Taken together, these findings suggest that the observed differences in response sensitivity across groups reflect underlying differences in unaided audibility but not differences in their access to stimuli during the experimental tasks.

A separate set of linear quantile mixed-effects models was conducted with group removed from the models to determine the relation between children's hearing sensitivity, as measured by unaided PTA and aided PTA, and response sensitivity across the auditory and visual domains. If disrupted auditory experience impedes domain-general selective attention during childhood, children's hearing sensitivity should relate to their response sensitivity during the auditory and visual change detection tasks. Furthermore, the slope of the relation between hearing sensitivity and response sensitivity may change as a function of age as children's selective attention continues to develop.

To test these hypotheses, aided and unaided PTA were added to separate models along with age and task as well as two-way age-by-PTA and task-by-PTA interactions. Non-significant three-way interactions among age, task, and PTA were removed from both models to increase model parsimony. *Appendix A.2.4* shows the output from these models and **Figure 2.10** displays these relations graphically. Overall, the findings of this analysis suggest that hearing sensitivity (measured in the unaided or aided condition for children with HL) related to response sensitivity during the auditory and visual change detection tasks, and that these relations became more positive as a function of age. Specifically, results revealed that both aided and unaided PTA significantly related to response sensitivity (estimate [aided PTA] = -0.061, $p < 0.001$; estimate[unaided PTA] = -0.024, $p < 0.001$; *Appendix A.2.4* **A & B**). Notably, the slopes of these relations became more positive with increasing age as indicated by significant age-by-PTA interactions (estimate[aided PTA] = 0.014, $p < 0.01$; estimate[unaided PTA] = 0.004, $p < 0.05$; *Appendix A.2.4* **A & B).** Though there was not a task-by-PTA interaction for the aided condition (estimate = 0.021, $p = 0.236$; *Appendix A.2.4* **A)**, there was a significant interaction between task and unaided PTA (estimate = 0.015, $p < 0.01$; *Appendix A.2.4* **B)**. Pairwise comparisons revealed that the slope of the relation between unaided PTA and response sensitivity was significantly less negative in the visual domain (estimate = -0.009, $p < 0.05$) than the auditory domain (estimate = -0.024, $p < 0.001$). These results should be interpreted with caution, however, due to the limited sample of children with HL included here. Additionally, in considering that the distribution of children's hearing sensitivity is non-normal and, in some cases, limited to a restricted range, future research is needed to further explore these relations in a larger sample of children with varying degrees of HL.

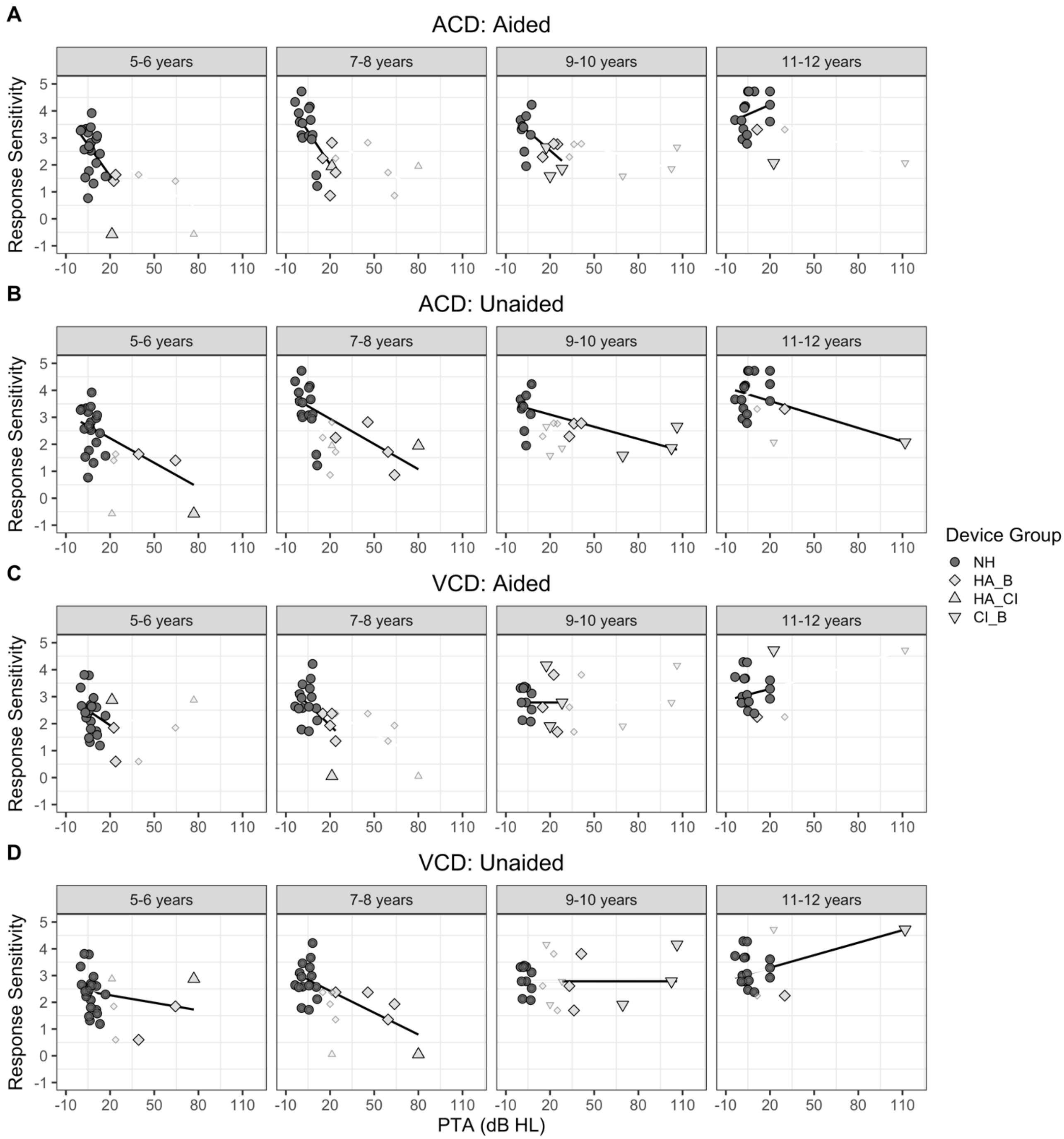


Figure 2.10: Response sensitivity during the auditory and visual change detection tasks as a function of aided and unaided pure-tone averages. Relations between aided (**A & C**) and unaided (**B & D**) pure-tone averages and response sensitivity are displayed for the auditory (**A & B**) and visual (**C & D**) change detection tasks. Individual data points and approximated 0.5 quantile regression lines (*solid lines*) are shown for children with NH (*dark gray*) and children with HL (*light gray*). The small faded points in each figure provide a reference to the other threshold condition (i.e., unaided or aided) not included in each relation for children with HL.

2.4.3 Comparison to Standardized Measures of Executive Function

Lastly, in considering that the primary objective of the present study was to determine whether age and hearing loss influence selective attention during childhood, it is necessary to consider whether the observed effects of age and group on children's performance reflect underlying differences in executive function. Specifically, if children's performance on the auditory and visual change detection tasks reflects their ability to selectively attend to a designated target stream and inhibit attention to competing streams, individual differences in performance on these tasks should relate to individual differences in executive function. To assess these relations, Spearman correlations were measured between children's performance on standardized measures of executive function, as quantified by raw scores to provide a direct comparison across children, and response sensitivity during the auditory and visual change detection tasks (**Table 2.7**).

Overall, the presence of significant relations between response sensitivity during the auditory and visual change detection tasks and children's performance on various standardized measures of executive function suggest that the experimental tasks are taxing cognitive processes of interest. The observation that the same standardized measures did not relate similarly to response sensitivity across auditory and visual domains, however, suggests that disparate task demands may have resulted in a differential dependence on underlying cognitive and/or behavioral mechanisms. (**Table 2.7**). Specifically, for the auditory change detection task, significant relations were observed between response sensitivity and children's performance on 1) the Flanker Inhibitory Control and Attention Task ($\rho = 0.340$, $p < 0.01$); and 2) two subtests of the TEA-Ch2, specifically the Hide and Seek Auditory/Cerberus ($\rho = -0.393$, $p < 0.001$) and Simple Reaction Time ($\rho = -0.387$, $p < 0.001$) subtests. Trending-toward-significant relations were observed for the Sustained Attention Response Test subtest (ρ = -0.210, $p = 0.081$) and the Sustained Attention Index of the TEA-Ch2 (ρ = 0.210, $p = 0.083$). No significant relations were observed between response sensitivity and parental ratings of children's auditory attention skills on the CHAPS questionnaire (ρ = 0.093, $p = 0.433$) or children's performance on the other subtests of the TEA-Ch2, specifically the Barking/Vigil subtest (ρ = 0.114, $p = 0.342$) .

		Response Sensitivity	
	Age	Auditory Change Detection Task	Visual Change Detection Task
CHAPS Total Score	ρ = -0.052, *p* = 0.664	ρ = 0.093, *p* = 0.433	**ρ = 0.329, *p* < 0.01**
Flanker Score Uncorrected	**ρ = 0.717, *p* < 0.001**	**ρ = 0.340, *p* < 0.01**	**ρ = 0.317, *p* < 0.01**
TEA-Ch2 Raw Scores			
Barking/Vigil	ρ = -0.071, *p* = 0.553	ρ = 0.114, *p* = 0.342	ρ = 0.124, *p* = 0.298
Hide and Seek Auditory/Cerberus	**ρ = -0.402, *p* < 0.001**	ρ = **-0.393, *p* < 0.001**	**ρ = -0.249, *p* < 0.05**
Sustained Attention Response Test	**ρ = -0.352, *p* < 0.01**	ρ = -0.210, *p* = 0.081	**ρ = -0.378, *p* < 0.01**
Simple Reaction Time	**ρ = -0.630, *p* < 0.001**	**ρ = -0.387, *p* < 0.001**	**ρ = -0.279, *p* < 0.05**
Sustained Attention Index	ρ = -0.081, *p* = 0.507	ρ = 0.210, *p* = 0.083	ρ = 0.183, *p* = 0.133

Table 2.7: Correlations between performance on the standardized measures of executive function and response sensitivity during the auditory and visual change detection tasks. Spearman correlation coefficients and significance values are provided for each relation.

For the visual change detection task, significant relations were observed between response sensitivity and children's performance on 1) the Flanker Inhibitory Control and Attention Task (ρ = 0.317, $p < 0.01$); 2) three subtests of the TEA-Ch2, specifically the Hide and Seek Auditory/Cerberus (ρ = -0.249, $p < 0.05$), Sustained Attention Response Test (ρ = -0.378, $p < 0.01$), and Simple Reaction Time (ρ = -0.279, $p < 0.05$) subtests, and 3) parental ratings on the CHAPS questionnaire (ρ = 0.329, $p < 0.01$). However, no significant relations were observed between response sensitivity and performance on the Barking/Vigil subtest (ρ = 0.124, $p = 0.298$) or the Sustained Attention Index (ρ = 0.183, $p = 0.133$) of the TEA-Ch2.

In considering that children's performance on the standardized measures of executive function and the auditory and visual change detection tasks was shown to be influenced by age, it is unclear whether the observed relations between executive function and response sensitivity reflect common underlying attentional mechanisms or behavior related to other facets of development. To further elucidate these relations accounting for age, raw scores from the standardized measures that significantly related to response sensitivity (i.e., Flanker Inhibitory Control and Attention Task, Hide and Seek Auditory/Cerberus subtest, and Simple Reaction Time subtest) were included in separate linear quantile mixed-effects models with response sensitivity as the dependent variable and age as a covariate. Separate models were generated for each standardized measure of executive function because, as described above, a slightly different number of children completed each task, which precluded the use of a single within-subjects model. Prior to analysis, each raw score was centered to allow for a more meaningful interpretation of model intercepts by subtracting the median score from each observed value.

Appendix A.2.5 displays the output of each model testing for the relation between children's performance on the Flanker Inhibitory Control and Attention Task (*Appendix A.2.5* **A)**, Simple Reaction Time (*Appendix A.2.5* **B)** subtest, and Hide and Seek Auditory/Cerberus subtest (*Appendix A.2.5* **C**) and response sensitivity. Overall, the findings of this analysis support the interpretation that observed differences in children's performance on the auditory and visual change detection tasks reflect, at least in part, underlying differences in executive function, including selective attention. Specifically, after accounting for age, significant relations with response sensitivity were not observed for performance on the Flanker Inhibitory Control and Attention Task (estimate = 0.003, $p = 0.636$) or the Simple Reaction Time task (estimate = -0.0006, $p = 0.199$). However, a modest but significant effect of performance on the Hide and

Seek Auditory/Cerberus subtest was observed (estimate = -0.0002, $p < 0.05$), suggesting that children's ability to selectively attend to target stimuli and inhibit attention to distractor stimuli significantly related to response sensitivity during the auditory and visual change detection tasks. To test whether this relation differed across the auditory and visual change detection tasks, task was added as a factor to the model with age and performance on the Hide and Seek Auditory/Cerberus subtest. The absence of a task-by-Hide and Seek Auditory/Cerberus subtest interaction (estimate = 0.0001, $p = 0.230$) suggests that the relation between performance on this standardized measure of executive function and response sensitivity was consistent regardless of task domain (*Appendix A.2.5* **D**; **Figure 2.11**).

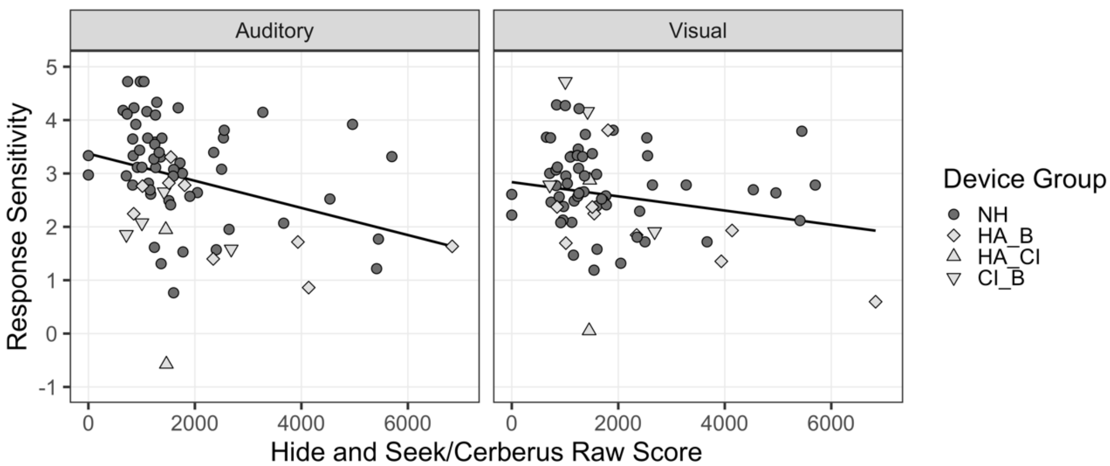


Figure 2.11: **Relation between performance on the Hide and Seek Auditory/Cerberus standardized measure of attention and inhibition and response sensitivity on the auditory and visual change detection tasks.** Individual data points and approximated 0.5 quantile regression lines (*solid lines*) are displayed for children with NH (*dark gray*) and children with HL (*light gray*) for the auditory (*left)* and visual (*right*) change detection tasks.

2.5 DISCUSSION

The primary objective of the present study was to test the hypotheses that immaturity and disrupted auditory experience impede children's selective attention, and that any differences in selective attention related to children's hearing status are domain general. Children between 5 and 12 years of age with NH and HL completed behavioral change detection tasks in the auditory and visual domains during which they were instructed to selectively attend and respond to deviant stimuli within a target stream while inhibiting attention to a distractor stream. Children's performance on the auditory and visual change detection tasks was quantified by the frequency and speed of their responses to stimuli in the target and distractor streams as well as response sensitivity, which was used to quantify their ability to selectively attend to the target stream. The separate and interacting effects of age, hearing status, and task (i.e., auditory vs. visual) on performance were assessed. To examine the potential mechanisms contributing to children's performance on the auditory and visual change detection tasks, the relations among response sensitivity, hearing sensitivity, and performance on standardized measures of executive function were considered using linear quantile mixed-effects modeling.

As predicted, age and hearing loss significantly affected children's performance on the auditory and visual change detection tasks. Specifically, younger children and children with HL demonstrated lower response sensitivity than older children and children with NH, which is suggestive of greater difficulty selectively attending to the target stream. Notably, the differences in the ability of children with NH and HL to selectively attend to the target stream were observed across the auditory and visual domains, albeit to different extents.

The differences in response sensitivity between children with NH and HL were driven by differences in unaided PTA across groups, though aided and unaided PTA were both found to negatively relate to performance such that children with poorer hearing sensitivity demonstrated lower response sensitivity. The relations between unaided PTA, aided PTA, and response sensitivity during the auditory and visual change detection tasks became significantly less negative as a function of age, providing preliminary evidence that the extent to which hearing loss alters children's ability to selectively attend to a target stream may diminish with age. Lastly, children's performance on a standardized measure of executive function was shown to relate to their response sensitivity during the auditory and visual change detection tasks, demonstrating that the observed differences in selective attention are not isolated to the sensory domain of impairment for children with HL.

2.5.1 Measuring Selective Attention in Children

Selective attention is a multifaceted and dynamic cognitive process that is difficult to measure empirically. While the tasks implemented in the present study did not capture children's selective attention in its entirety, they did reflect the efficacy and efficiency with which children selectively allocated attention to a designated target stream. Specifically, the benefit of the auditory and visual change detection tasks employed in the present study is that various aspects of children's responses could be quantified, including the type (i.e., hit; false alarm; non-deviant response) and speed of their responses. The pattern of these responses provided a useful indication of how and where children deployed attention during the tasks.

The primary dependent variable of interest for the present study was response sensitivity, which reflected the proportion of children's responses to deviants in the target stream (i.e., hits) relative to the proportion of their responses to deviants in the distractor stream (i.e., false alarms). Higher response sensitivity represented a higher proportion of hits and a lower proportion of false alarms and was, therefore, indicative of an intact ability to selectively attend to the target stream. Lower response sensitivity represented a reduced ability to selectively attend to the target stream, which was reflected as either: 1) a higher proportion of hits and false alarms; 2) a lower proportion of hits and false alarms; or 3) a lower proportion of hits and a higher proportion of false alarms. Therefore, it is necessary to consider response sensitivity within the context of the underlying proportions of responses to deviants in the target and distractor streams.

The subsequent sections will discuss the following: 1) the observed effects of age and hearing status on children's ability to selectively attend to the target stream; 2) the observed differences in performance across the auditory and visual change detection tasks and what these differences may indicate about children's response patterns; and 3) the potential mechanisms underlying the effect of hearing status on selective attention and potential implications.

2.5.2 Child-Specific Factors that Influence Selective Attention

Results revealed that age and hearing status influenced children's ability to selectively attend to the target stream across the auditory and visual domains.

Age

As expected, response sensitivity during the auditory and visual change detection tasks improved with age for children with NH and HL. Specifically, younger children demonstrated a lower proportion of responses to deviants in the target stream and a higher proportion of responses to deviants in the distractor stream than older children across the auditory and visual domains. This pattern of responses suggests that younger children may have allocated attention more diffusely across the target and distractor streams or switched attention between the target and distractor streams as a result of immature attentional control rather than selectively attending to the target stream and inhibiting attention to the distractor stream. Consistent with these possibilities, previous research has demonstrated that younger children tend to distribute attention across multiple aspects of sensory input, even those not relevant to the task at hand (Hanania & Smith, 2010; Leon-Carrion et al., 2004; Plebanek & Sloutsky, 2017). For instance, Plebanek & Sloutsky (2017) showed that children 4 to 5 years of age attended to irrelevant features of stimuli during a visual search task to a significantly greater extent than adults. In that study, this diffuse allocation of attention benefitted children during the visual search task by allowing them to accurately recognize the same number of task-relevant features and more task-irrelevant features than adults. This attention allocation strategy, however, is detrimental to performance on tasks that require selective attention, such as the auditory and visual change detection tasks implemented in the present study.

In addition to lower response sensitivity, younger children responded more slowly to deviants in the target stream than older children during the auditory and visual change detection tasks. These findings are consistent with previous studies that have shown similar age-related changes in reaction time for tasks in the auditory domain (e.g., Fuchigami et al., 1993) and visual domain (e.g., Grieco-Calub et al., 2017). Age-related changes in reaction time are generally thought to be due to the development of processing speed, which has been shown to parallel the development of executive functions (Chevalier et al., 2015; Kail, 1991; Kail & Ferrer, 2007; Kail & Salthouse, 1994) and motor reaction time (Goodenough, 1935; Thomas et al., 1981) during childhood. In the present study, interpreting individual differences in children's reaction times within the context of their response patterns may provide additional insight as to how children allocated attention during the auditory and visual change detection tasks. Specifically, the observation that younger children demonstrated slower and more variable reaction times to deviants in the target stream in addition to lower response sensitivity than older children across domains is consistent with the interpretation that younger children had a reduced ability or willingness to selectively allocate attention to the target stream and respond to deviants in an efficient manner.

Overall, the finding that children's ability to selectively attend to the target stream improved with age is consistent with previous research, and is expected to reflect age-related changes in neurocognitive development (Davidson et al., 2006; De Luca et al., 2003; Kronenberger et al., 2014; Machinskaya, 2006; Plude et al., 1994). Behaviorally, younger children have reliably shown reduced selective attention compared to older children, as demonstrated by greater distractibility from competing auditory and visual streams (Doyle, 1973; Elliott, 2002; Enns & Cameron, 1987; Jones et al., 2015; Klatte, Lachmann, Schlittmeier, et al., 2010), a reduced ability to allocate attention in accordance with task demands (Irwin-Chase & Burns, 2000; Karatekin, 2004; Pearson & Lane, 1991), and a greater tendency to process task-irrelevant information within auditory and visual streams (Hanania & Smith, 2010; Leon-Carrion et al., 2004; Plebanek & Sloutsky, 2017). Coincident evidence of the prolonged maturation of selective attention is evidenced by studies showing differential activation patterns in cortical regions comprising attention-based neural networks between children and adults (e.g., Abundis-Gutiérrez et al., 2014; Booth et al., 2003). These findings reinforce those of the present study, and provide converging evidence of the development of selective attention during childhood.

Hearing Status

As predicted, response sensitivity during the auditory and visual change detection tasks was lower for children with HL compared to children with NH, suggesting that they had greater difficulty selectively attending to the target stream. Specifically, children with HL demonstrated a lower proportion of responses to deviants in the target stream and a higher proportion of responses to deviants in the distractor stream than children with NH across the auditory and visual domains. In addition, children with HL responded significantly more slowly to deviants in the target stream during the auditory change detection task and the distractor streams during the auditory and visual change detection tasks than children with NH. This response pattern parallels that of younger children, as described above, and suggests that children with HL may have allocated attention more diffusely across the target and distractor streams than children with NH.

While the main effect of hearing status (i.e., group) was present regardless of task domain, the observed differences in response sensitivity between children with NH and HL were significantly greater in the auditory domain than the visual domain. As displayed in **Figures 2.6-2.8**, this difference appears to be driven by a reduction in the proportion of deviants in the target stream to which children with NH responded during the visual change detection task as children with HL maintained their performance across the auditory and visual domains. There are at least two possible interpretations of this observation. Firstly, if considering the response patterns of children with NH as the developmental standard, children with HL of the same age did not achieve the same level of performance as children with NH during the auditory change detection task.

Secondly, children with NH may have performed more poorly during the visual change detection task as a result of a methodological constraint. Specifically, a greater portion of children with NH were tested at an ISI of 0 ms, as determined during the adaptive visual change detection task, than children with HL (i.e., 48% of children with NH vs. 25% of children with HL). In considering the overall faster presentation speed of visual stimuli for the children with NH relative to children with HL, the reduced proportion of responses to deviants in the target stream for children with NH may have been due to momentary lapses in attention or other perceptual phenomena, such as the attentional blink, rather than reduced attention to the target stream per se (Heim et al., 2015; Martens & Wyble, 2010; Shapiro et al., 1997). To further explore the effect of ISI on children's performance, post hoc correlation analyses were conducted and revealed significant or trending-toward-significant negative relations between adaptive ISI and children's performance during the auditory and visual change detection tasks (see *Appendix A.2.6* for correlation plots). Specifically, children with NH and HL who performed better on the adaptive task – and, as a result, were tested at a faster ISI – demonstrated higher response sensitivity during the experimental tasks despite the faster stimulus presentation rate. Also notable in these correlation plots is the downward shift in response sensitivity during the visual change detection task relative to the auditory change detection task for children with NH who were tested at an ISI of 0 ms, as described above. These findings provide evidence against the notion that the performance of children who completed the auditory and visual change detection tasks at a faster ISI was disadvantaged by the rate at which stimuli were presented. Therefore, it is possible that other factors contributed to the observed differences in hits between children with NH and HL during the visual change detection task.

Critically, the observation that children with HL still performed more poorly, as quantified by hits as well as response sensitivity, than children with NH regardless of potential methodological constraints is consistent with the interpretation that children with HL had greater difficulty selectively attending to the target stream than children with NH. Furthermore, the fact that these differences were observed across the auditory and visual domains suggests that hearing loss during childhood alters selective attention beyond the sensory domain of impairment. These findings expand upon those of previous studies showing that children with HL who use hearing aids or cochlear implants allocate attention differently than their peers with NH across auditory and visual domains. In the visual domain, children with HL have been shown to perform more poorly on measures of attentional control using a nonverbal visual search task (Kronenberger et al., 2020) as well as visual continuous performance tasks that require selective and sustained attention over time (Dye & Hauser, 2014; Quittner et al., 1994). In the auditory domain, children with HL have been shown to direct reduced preparatory spatial attention to speech following a visual cue (Holmes et al., 2017) and perform more poorly on selective attention-based dichotic listening task (Asbjørnsen & Ma, 2000). The present study contributes to this existing body of literature by demonstrating that children with HL had greater difficulty selectively attending to a target stream than children with NH across the auditory and visual domains within the same cohorts of children.

2.5.3 Differences in Non-Deviant Responses Across Auditory and Visual Domains

In addition to response sensitivity, another aspect of children's performance on the auditory and visual change detection tasks was the number of responses to non-deviant (i.e., standard) stimuli to which children responded (i.e., non-deviant responses). While the differences in children's performance described thus far are consistent across the auditory and visual domains, cross-domain differences were observed in the number of non-deviant responses children produced. Specifically, during the auditory change detection task, younger children and children with HL produced a greater number of non-deviant responses than older children and children with NH. These trends were not observed in the visual domain. In considering that the proportion of responses to deviants in the target and distractor streams did not significantly differ across the auditory and visual domains, these findings support the observation that younger children and children with HL provided more responses overall during the auditory change detection task.

A potential reason for the observation that younger children had a greater number of non-deviant responses is that there were differences in children's ability to detect deviants across the auditory and visual domains. Specifically, younger children and children with HL may have had greater difficulty detecting the standard-deviant contrasts during the auditory change detection task than older children and children with NH. There is precedent for this in the literature, as it is well known that the development of auditory perceptual skills, including frequency discrimination, continues during childhood (Moore et al., 2008; see Sanes & Woolley, 2011 for a review) and that some children with HL, especially those who use cochlear implants, have poorer frequency discrimination than children with NH (Carroll & Zeng, 2007; Cleary et al., 2005; see Oxenham, 2008 for a review). However, all children demonstrated the ability to successfully discriminate between standard and deviant words for the female and male talker during the familiarization paradigm. Therefore, the cross-domain differences in the number of non-deviant responses are not expected to be due to insufficient bottom-up access to the standard-deviant contrasts in the auditory domain.

Rather, it is possible that the standard-deviant contrasts during the auditory change detection task were less perceptually salient than those during the visual change detection task. Specifically, greater attentional demands may have been associated with detecting a deviant in the auditory domain, resulting in younger children and children with HL – the children who demonstrated greater difficulty selectively attending to the target stream – adopting a more liberal response strategy due to their inability or unwillingness to selectively allocate attention to meet these demands. Another reason for a more liberal response strategy may have been that children were less certain about their responses and developed a compensatory strategy by responding when they thought a deviant had been presented even if they were not fully confident. Children with HL may have been especially prone to adopting this response pattern during the auditory change detection task in considering they have a sensory impairment in the auditory domain and, as a result, may be less confident of their perception of a deviant (or lack thereof). In fact, the data in the present study suggest that children's uncertainty regarding whether they detected a deviant when one was *not* present as opposed to when one *was* present contributed the most to the increase in the number of responses during the auditory change detection task. Specifically, the observation that the proportion of responses to deviants in the target and distractor streams was consistent for children with HL across the auditory and visual domains suggests that the increase in the number of responses was isolated to non-deviant stimuli as opposed to a general increase in responses across stimulus types.

2.5.4 Possible Mechanisms Underlying the Effect of Hearing Loss on Selective Attention

The finding that children with HL demonstrated a reduced ability to selectively attend to the target stream compared to children with NH across the auditory and visual domains suggests that the mechanism(s) underlying these differences are not specific to the sensory domain of impairment for children with HL. This finding is consistent with extant research showing deficits in executive function in children with HL in the auditory and/or visual domains with multiple possible mechanisms underlying these observations. One potential mechanism underlying these observations is that disrupted access to auditory input early in life alters the development of domain-general executive functions (Conway et al., 2009; Quittner et al., 1994; Smith et al., 1998). This may reflect functional reorganization in auditory cortical pathways due to disrupted stimulus-driven learning and a reduction of cross-domain sensory redundancy or concomitant delays in language development, all of which have downstream effects on neurocognitive development and executive function (Botting et al., 2017; Figueras et al., 2008; Kral & Eggermont, 2007). Another possible explanation for the results of the present study is that children with HL may be delayed in their development of selective attention relative to their peers with NH due to altered sensory experience early in life. Consistent with this idea is the observation of trending-toward-significant age-by-group interactions for hits in both the auditory and visual domains. However, additional research is needed to further explore this possibility.

An alternate mechanism suggests that children with HL display a differential allocation of attentional resources to auditory and visual input resulting in greater distractibility and more diffuse attention in children with HL (Dye & Hauser, 2014; Holmes et al., 2017; McFadden & Pittman, 2008). In the auditory domain, this may reflect a capacity-limited inability to meet the attentional demands of the task due to a degraded peripheral representation of auditory input and reduced access to spectrotemporal cues (Bernstein & Oxenham, 2006; Oxenham, 2008). Therefore, children with HL may experience increased attentional demands associated with forming and tracking auditory objects as well as segregating auditory streams. Due to the increased attentional demands associated with these processes, children with HL may have greater difficulty selectively attending to the target speech stream and inhibiting attention to competing auditory streams (see Shinn-Cunningham & Best, 2008 for a review). In the visual domain, children with HL may allocate attention more diffusely due to reduced sensory redundancy across domains and, therefore, greater dependence on vision to monitor the environment and direct attention to events in their surroundings (see Tharpe et al., 2002 for a review). Some evidence suggests that providing children with enhanced access to sound via cochlear implantation may aid in reducing this dependence on visual attention and result in concomitant improvements in performance on temporal sequencing tasks in the visual domain (Quittner et al., 1994; Smith et al., 1998). Inherently, these mechanisms reflect differences in children's hearing sensitivity and the downstream effects of this disrupted auditory experience on the availability and allocation of attentional resources across auditory and visual domains.

In considering the results of the present study within the context of these proposed mechanisms, it is expected that children's response sensitivity during the auditory and visual change detection tasks may relate to their underlying hearing sensitivity and the development of their executive function. To account for the range of degrees of hearing loss and hearing device configurations of children with HL, aided and unaided PTAs were considered. Aided PTA provides an estimate of children's hearing sensitivity under optimal conditions (i.e., with the use of properly fitted hearing devices) whereas unaided PTA provides an estimate of their degree of hearing loss and associated auditory experience without the use of hearing devices. In addition, to capture various aspects of executive function, children's performance on multiple standardized measures were compared to their response sensitivity during the auditory and visual change detection tasks.

Both aided and unaided PTA were found to negatively relate to response sensitivity in the auditory and visual domains. These relations became less negative with increasing age, suggesting that the extent to which hearing sensitivity alters children's ability to selectively attend to a target stream in the auditory and visual domains may diminish with age. Therefore, older children with HL may demonstrate similar abilities to allocate attention to a target stream as their peers with NH. Results from previous studies have provided some evidence that this is the case. For instance, Dye & Hauser (2014) showed that younger (i.e., 6 to 8 year-old) children with HL demonstrated poorer selective attention than their peers with NH during a temporal sequencing task in the visual domain, though these differences were not observed in older (i.e., 9 to 13 year-old) children with HL. In the present study, while visual inspection of children's performance during the auditory and visual change detection tasks suggests the emergence of these trends, no significant age-by-group interactions were observed. The observation that the ability of children with HL to selectively attend to a target stream was still poorer than that of children with NH at 12 years of age suggests that the effect of hearing loss on children's ability to selectively attend to a target stream amidst competing input may persist into adolescence. One reason for this may be the continued maturation of attentional control or general executive function during adolescence and into adulthood, which may compensate for the differences in selective attention observed between children with NH and HL during childhood (Karns et al., 2015; Plude et al., 1994).

Additionally, children's performance on standardized measures of executive function were found to relate to response sensitivity during the auditory and visual change detection tasks. Specifically, children who had better performance on the standardized measures of executive function achieved higher response sensitivity. However, of the standardized measures included, children's performance on only one – the measure of their ability to selectively attend to and detect a target stimulus within an auditory stream – was found to significantly relate to response sensitivity after accounting for age. Notably, the extent of this relation was consistent across the auditory and visual change detection tasks. This finding suggests that the observed differences in performance between children with NH and HL may be due, at least in part, to underlying differences in the capacity or allocation of attentional resources in the auditory and visual domains. However, future research is needed to further explore the mechanisms contributing to the development of selective attention in children with HL.

Together, these findings suggest that the mechanism by which children's hearing sensitivity contributes to their ability to selectively allocate attention is not rooted solely in the auditory domain. Specifically, the present study provides evidence of a domain-general mechanism as differences in the ability of children with NH and HL to selectively attend to a target stream were observed across the auditory and visual domains. Therefore, it is possible that reduced or inconsistent access to robust auditory input early in life may alter the development of children's attentional capacity as well as their ability to effectively allocate attentional resources across the auditory and visual domains. Increasing the attentional demands in either the auditory or visual domain, such as by adding distractor streams that compete for attentional resources, may impede the ability of children with HL to meet those demands, resulting in greater difficulty selectively attending to a target stream than their peers with NH, as demonstrated in the present study.

2.5.5 Limitations

While the present study contributes novel insight regarding how age and hearing status influence children's ability to selectively attend to a target stream in the auditory and visual domains, there are a few limitations to consider. Firstly, consistent with all behavioral paradigms, the method of quantifying selective attention was based on children's overt responses, which can be influenced by dynamic processes such as their physiological state, degree of compliance, and motivation to complete the task. Additional factors related to the experimental design, such as the attentional demands imposed by the tasks, are also expected to modulate children's performance. Therefore, while children's performance on these tasks reflected their ability to selectively attend to a target stream in the context of this experiment, it is possible that variations in this ability would be observed under alternate conditions. Future research should consider investigating children's selective attention using a variety of experimental and ecologically-valid tasks that differ in regard to their dependence on overt responses from the child in order to explore this possibility.

Specific to the design of the change detection tasks employed in the present study, it proved challenging to develop analogous tasks to probe selective attention across the auditory and visual domains. While all methodological attempts were made to equate task demands and ensure equal access to stimuli across ages and groups, it is possible that differences in perceptual and, therefore, attentional demands persisted. As described above, this may have contributed to the differences in response patterns observed between the auditory and visual change detection tasks. Potentially beneficial modifications to this methodology for use in future studies may include the addition of a behavioral threshold-based measure of frequency and saturation discrimination specific to the auditory and visual domains, respectively, and the implementation of individualized standard-deviant contrasts to ensure equal detectability of deviants across domains.

A final limitation is the difference in sample size between the groups of children with NH and HL. The overall limited sample size of children with HL relative to children with NH was accounted for during the planning and implementation of statistical analyses; however, examining certain group-specific relations was precluded due to reduced statistical power in the group of children with HL relative to the group of children with NH. In addition, there was considerable heterogeneity in the audiologic profiles of children with HL. Some factors that varied within this group included the etiology, onset, type, and degree of hearing loss as well as the type of hearing devices and the duration of their usage. Degree of hearing loss was accounted for during the statistical analyses when possible, as described in *Section 2.4 Results*. From a mechanistic perspective, while these factors likely contribute to the effect of hearing loss on selective attention, previous literature has documented similar trends of reduced selective attention in children with even mild degrees of hearing loss (Asbjørnsen & Ma, 2000). Future studies can build upon the groundwork provided by the present study and others to further examine selective attention in a more robust sample of children with HL.

2.5.6 Implications and Future Directions

The results of the present study have potential implications for children’s speech recognition, especially in considering that the real-world environments within which children spend a considerable amount of time often contain multiple sources of auditory and visual input. Understanding how children’s age and hearing status contribute to their ability to selectively attend to a target stream and inhibit attention to competing streams may provide additional insight into the difficulties children have listening and learning in such environments. For instance, the observed age-related improvements in selective attention provide a possible empirical explanation as to why younger children are poorer at understanding speech in complex acoustic environments than older children (Bradley & Sato, 2008; Jones et al., 2015; Leibold & Buss, 2013; Neuman et al., 2010). In addition, the finding that children with HL had greater difficulty selectively attending to a target stream than children with HL offers a potential explanation as to why children with HL have difficulty understanding speech in complex acoustic environments despite the use of clinical hearing devices. From a clinical perspective, the results from the present study reinforce the importance of habilitation programs aimed at strengthening top-down “listening skills” in children with HL and emphasize the need to direct selective attention to the talker or sound source as part of this intervention. More generally, apprising adults who frequently interact with young children about the development of selective attention may foster better communication strategies that encourage attending to the talker and, as a result, bolster academic success and social wellbeing.

While these implications are noteworthy, further research is needed to delineate the role of selective attention in children’s ability to understand speech. There is theoretical reason to believe that children who have greater difficulty selectively attending to a target stream may have poorer speech recognition in complex acoustic environments, such as classrooms, which may hinder their social and academic success. However, only a few studies have investigated the relation between selective attention and speech recognition in children with NH and HL, and the extent to which the acoustic characteristics of the environment modulate this relation remains unknown. The study described in *Chapter 3* aimed to address these knowledge gaps.

2.6 CONCLUSION

In summary, the present study revealed that immaturity and disrupted auditory experience impede children’s ability to selectively attend to a target stream in the auditory and visual domains. Specifically, younger children and children with HL demonstrated lower response sensitivity than older children and children with NH during the auditory and visual change detection tasks. Additionally, children’s hearing sensitivity and performance on a standardized measure of selective attention significantly related to their response sensitivity in the auditory and visual domains. Together, these findings suggest that children with NH and HL differentially allocate attention to sensory input, with children with HL attending less selectively than children with NH across the auditory and visual domains.

CHAPTER 3 | RELATION BETWEEN SELECTIVE ATTENTION AND CHILDREN'S SPEECH RECOGNITION IN COMPLEX ACOUSTIC ENVIRONMENTS

3.1 ABSTRACT

Previous research has demonstrated that children's ability to selectively attend to a target speech stream is influenced by their age and hearing status (*Chapter 2*), yet the implications of this for speech recognition in complex acoustic environments remain largely unknown. The purpose of the present study was to investigate the relation between selective attention and speech recognition in children. Children between 5 and 12 years of age with normal hearing and hearing loss performed a speech recognition task under acoustic conditions that varied based on reverberation time, masker type, and the spatial location of the masker. Results revealed that younger children and children with hearing loss achieved poorer speech recognition than older children and children with normal hearing. As hypothesized, children's ability to selectively attend to a target speech stream significantly related to their speech recognition, especially under acoustic conditions expected to impose greater attentional demands.

3.2 INTRODUCTION

Children spend a considerable amount of time in complex acoustic environments that contain background noise and reverberation. Elementary school classrooms, for example, have been shown to contain high levels of background noise ranging from 30 to 70 dBA and long reverberation times between 0.4 and 1.2 seconds (Crandell & Smaldino, 2000; Crukley et al., 2011). The presence of background noise and reverberation as well as their synergistic effects pose significant challenges for children's ability to understand speech, which is detrimental for communication, socialization, and academic achievement (Bradley & Sato, 2008; Dockrell &

Shield, 2006; Jamieson et al., 2004; Klatte et al., 2010; Nelson & Soli, 2000; Shield & Dockrell, 2003). Although children's ability to understand speech in complex acoustic environments varies widely, some groups of children are considered to be significantly more at-risk for poor speech recognition in these environments. For instance, young children and children with hearing loss consistently demonstrate greater difficulty understanding speech amidst background noise and reverberation (Fallon et al., 2000). One explanation for these observations is that immaturity and disrupted auditory experience during childhood alter the development of cognitive and linguistic processes, such as selective attention, working memory, and vocabulary knowledge, that are thought to underlie speech recognition in complex acoustic environments.

Consistent with this idea, the findings from *Chapter 2* provided evidence that younger children and children with hearing loss had greater difficulty selectively attending to a target speech stream and inhibiting competing input than older children and children with normal hearing. However, the implications of this for speech recognition in complex acoustic environments remain unknown. Therefore, the present study aimed to investigate whether poorer selective attention may be one such factor contributing to the difficulties younger children and children with hearing loss have understanding speech in the presence of background noise and reverberation. An in-depth understanding of the internal and external factors that contribute to speech recognition in children with normal hearing and hearing loss is necessary in order to maximize their academic success and social wellbeing.

3.2.1 Factors that Influence Speech Recognition

Fundamentally, background noise and reverberation alter the bottom-up sensory representation of the target speech signal in the ascending auditory system. Depending on the specific acoustic characteristics of the environment, the audibility of the target speech signal may be reduced due to increased overlapping energy between the background noise and the target speech across spectral and temporal domains – a process called *energetic masking* (Fletcher, 1940). As a result, greater demands may be placed on top-down cognitive and linguistic processes to resolve and extract information from the target speech signal (Camos & Barrouillet, 2011, 2014; Magimairaj & Montgomery, 2012a, 2012b; Mattys et al., 2012). The perceptual consequences resulting from the content of the masker imposing greater demands associated with the higher-level processing of the target speech is often referred to as *informational masking* (e.g., Brungart et al., 2001; Pollack, 1975; Wightman & Kistler, 2005; see Brungart, 2005 and Kidd et al., 2008 for reviews). Thus, background noise and reverberation may detriment speech recognition by degrading the bottom-up (i.e., peripheral) representation of the target speech signal as well as increasing the demand for top-down processes (e.g., cognitive and linguistic processes) to aid in resolving the degraded target speech (Fallon et al., 2002; Nittrouer & Boothroyd, 1990; Wightman & Kistler, 2005; see Leibold & Buss, 2019 for a review). While the peripheral encoding inherent to bottom-up processing has been shown to remain stable throughout childhood (see Eggermont & Moore, 2012 for a review), there is prolonged development of cognitive and linguistic processes (e.g., Luna et al., 2004; Nippold, 1998). Therefore, it is suspected that immature cognitive and linguistic processes contribute to the observed variability in children's ability to understand speech in complex acoustic environments (Leibold & Buss, 2019; Wightman & Allen, 2004).

Consistent with this idea, multiple cognitive and linguistic processes have been shown to contribute to the ability to understand speech in complex acoustic environments. For instance, as postulated by the Ease of Language Understanding (ELU) model (Rönnberg et al., 2013), working memory is a limited-capacity cognitive process thought to contribute to speech recognition by facilitating a listener's ability to store information and integrate it with stored knowledge during ongoing processing (Baddeley, 2003, 2010; Baddeley & Hitch, 1974;). Previous research has demonstrated that children who have higher working memory capacities are better at understanding speech in the presence of background noise and reverberation (MacCutcheon et al., 2019; McCreery et al., 2017; Sullivan et al., 2015). In addition, children who have stronger expressive vocabulary skills and better language abilities have been shown to have better speech recognition in noise, which is inherently a linguistically-demanding process (Thompson et al., 2019; Walker et al., 2019). Similar relations have been observed in children with hearing loss (Klein et al., 2017; McCreery et al., 2019; Pisoni & Cleary, 2003; Torkildsen et al., 2019), suggesting that the restored audibility of speech provided by the use of properly fitted hearing devices is necessary, but not sufficient, for speech recognition. However, unexplained variance in speech recognition remains even after accounting for the contribution of cognitive and linguistic factors, such as working memory, language ability, and vocabulary size, in children with normal hearing (Klein et al., 2017; McCreery et al., 2019; Pisoni & Cleary, 2003; Torkildsen et al., 2019) and hearing loss (Lewis & Hoover, 2010). These findings indicate that a multitude of cognitive and linguistic processes contribute to children's speech recognition in complex acoustic environments. Thus, the variability in children's speech recognition is likely due, in part, to their cognitive and linguistic immaturity.

3.2.2 Selective Attention and Speech Recognition

When considering the cognitive processes that contribute to children's ability to understand speech in complex acoustic environments, attention is a frequently mentioned, but rarely tested, factor. Selective attention is a specific type of attention that, as discussed in *Chapter 2*, underlies a listener's ability to allocate attention to the target speech to be further processed while inhibiting attention to competing auditory input (see Gomes, 2000 for a review). Within the context of auditory scene analysis, selective attention contributes to a listener's ability to track, resolve, and encode information from a target speech stream as it unfolds over time (Alain & Bernstein, 2008; Carlyon et al., 2001, 2003; Shinn-Cunningham, 2008; see Sussman, 2017 for a review). In addition, selective attention may reduce the demands on other cognitive processes, such as working memory, by specifying the aspects of the auditory input to be encoded by higher-level processes (Blamey et al., 2001; Caldwell & Nittrouer, 2013). Within the context of the ELU model, selectively attending to a target speech stream and inhibiting attention to competing auditory input reduces the occurrence of mismatches between the perceived phonological content of the speech stream and children's language knowledge that need to be resolved by working memory (Rönnberg et al., 2013). If, in fact, the ability to effectively allocate attention partially relieves the dependence on downstream cognitive processes for resolving degraded speech, children with stronger selective attention abilities should demonstrate better speech recognition in complex acoustic environments.

Selective attention has also been shown to alter the neural processing and perceptual representation of both target and distractor streams in the ascending auditory pathway. From a neurocognitive perspective, selectively attending to a target speech stream modulates the neural representation of the target speech signal in the brainstem (Forte et al., 2017; Lehmann & Scho, 2014) and language-based cortical areas (Rimmele et al., 2015; Wild et al., 2012; Yoncheva et al., 2010). From a behavioral perspective, previous research suggests that listeners' ability to process and recall the content of an auditory stream is modulated by the selective allocation of attention toward or away from that stream. This was initially demonstrated in a seminal study by Cherry (1953). Specifically, adults performed a dichotic listening paradigm during which they were presented with two streams of speech and instructed to listen to and repeat the content of the designated target stream while ignoring the other (i.e., distractor) stream. While the recognition of the words in the target stream was excellent, adults were not able to reliably report the language characteristics, semantic content, or individual words contained within the distractor stream (Cherry, 1953). This finding supports the early filter theory of attention proposed by Broadbent (1958), which posits that incoming peripheral input is held in a sensory buffer and selectively processed through a filter of attention if the physical characteristics of the input are relevant to the task at hand. Together, these findings suggest that a listener's ability to selectively attend to a target speech stream facilitates their ability to process, encode, and act upon the content of that stream.

If the observation that selective attention alters the processing of target and distractor streams generalizes to real-world listening situations, children with poor selective attention should demonstrate poorer speech recognition in complex acoustic environments. Consistent with this idea, the study described in *Chapter 2* revealed that younger children and children with hearing loss had greater difficulty selectively attending to a target auditory stream and inhibiting attention to a competing (i.e., distractor) auditory stream. This finding provides evidence that auditory selective attention is a dynamic cognitive process influenced by age and auditory experience during childhood that contributes to the observed variability in children's speech recognition.

Only a few studies to date have considered selective attention in the auditory domain within the context of speech recognition in children with normal hearing and children with hearing loss, and the results of these studies have been mixed. For instance, a recent study by Holmes et al. (2017) showed that 7 to 16 year-old children with hearing loss demonstrated less neural activity in anticipation of an upcoming auditory stimulus than their peers with normal hearing when visually prompted to selectively attend to a target speech stream amidst two competing speech streams that differed in spatial location and fundamental frequency. Consistent with this result, children with hearing loss performed significantly more poorly than their peers with normal hearing on a behavioral closed-set task measuring their recognition of words contained within the target speech stream. Notably, the same results were obtained regardless of whether children with hearing loss used their clinical hearing aids, suggesting that the observed differences in preparatory attention were not due to insufficient acoustic access to the speech stimuli. Together, these findings suggest that children with hearing loss may deploy selective attention to auditory input differently than children with normal hearing, which has downstream implications for speech recognition.

Contrary to these findings, a recent study by McCreery et al. (2019) revealed that individual differences in auditory attention did not account for differences in speech recognition between children with normal hearing and children with hearing loss. Specifically, 7 to 9 year-old children with normal hearing and children with hearing loss performed an adaptive speech recognition task in the presence of unmodulated speech-shaped noise and a 0.6-second reverberation time. There were no significant differences in speech recognition based on auditory attention after accounting for other factors, such as visuospatial working memory and receptive vocabulary. A potential reason for this finding is related to the standardized measure of auditory attention used in the study. Specifically, the Auditory Attention subtest of the Developmental Neuropsychological Assessment (NEPSI-II; Brooks et al., 2009) measures children's ability to monitor a single stream of words for a designated key word. In considering the absence of competing auditory streams in this measure, children’s performance may more so reflect the sustained, rather than selective, deployment of attention. Therefore, it is expected that a measure of children’s ability to selectively attend to a target signal amidst competing auditory input would be more predictive of speech recognition in complex acoustic environments. The present study will directly test this relation.

3.2.3 Speech Recognition in Complex Acoustic Environments

An inherent challenge of investigating the relation between selective attention and speech recognition is that children’s ability to understand speech varies widely based on myriad factors, including the reverberation characteristics of the environment as well as the presence, content, intensity, and spatial location of the target speech and background noise contained within. Numerous studies have investigated the independent effects of these acoustic characteristics on speech recognition in children, and converging evidence suggests the involvement of both bottom-up and top-down processes. Thus, in order to assess the extent to which selective attention contributes to speech recognition in complex acoustic environments, the individual mechanisms underlying the effects of the acoustic characteristics that comprise these environments must be well understood.

Children’s ability to understand speech depends, in part, on the acoustic characteristics of the environment within which the target speech is presented. Reverberation is one such characteristic that refers to the reflections of acoustic energy off of surfaces (e.g., floors, ceilings, windows, objects) in a room. It is often quantified as reverberation time, which refers to the amount of time required for the intensity of an emitted acoustic signal at a specific frequency to decrease by a certain amount (typically 30 dB SPL [T30] or 60 dB SPL [T60]) after the offset of the signal. In addition to the size and shape of the room, reverberation time is influenced by the absorption properties of the materials covering the surfaces: the less energy that is absorbed, the more energy that is reflected. These reflections interact with direct energy to varying extents depending on the location of the listener relative to the source (e.g., talker; Assmann & Summerfield, 2004; Boothroyd, 2004). When the listener is located within close proximity to the source, the listener receives mostly direct energy. As the distance between the listener and the source increases, the amount of direct energy decreases until the critical distance is surpassed, after which the reverberant energy exceeds the direct energy. Reverberation impairs speech understanding by smearing the spectrotemporal cues (e.g., formant transitions, temporal fine structure) as well as masking speech frequencies, both of which degrade the quality of the signal (Assmann & Summerfield, 2004; Boothroyd, 2004). Consistent with this, previous research has shown that speech recognition decreases as a function of increasing reverberation time in quiet for children with normal hearing (Neuman & Hochberg, 1983), and to an even greater extent for children with hearing loss (Crandell & Smaldino, 2000; Finitzo-Hieber & Tillman, 1978). These findings are especially noteworthy as the reverberation times included in these studies reflect the range of those observed in elementary school classrooms (i.e., 0.4-1.2 seconds), and impaired speech recognition was observed even at the recommended classroom reverberation time of 0.6 seconds (ANSI S12.60-2010; Crukley et al., 2011).

While reverberation impairs speech recognition in quiet, the effects have been shown to be even more detrimental in the presence of concomitant background noise (Klatte, Lachmann, & Meis, 2010; Neuman et al., 2010; Neuman & Hochberg, 1983; Yacullo & Hawkins, 1987). Background noise comprises any acoustic signal in the environment that competes with – or masks – the target speech. Empirically, the effects of background noise are investigated by presenting masker(s) that overlap in time and space with a concurrent task, typically one that involves speech recognition. Various factors influence the extent to which a masker impedes speech recognition. For instance, the intensity of the masker relative to the intensity of the target speech signal defines the signal to noise ratio (SNR), which provides a general indication of the audibility of the target speech. Extant research has demonstrated that children's ability to understand speech decreases as a function of SNR, with younger children and children with hearing loss requiring more positive SNRs than older children and children with normal hearing to achieve similar speech recognition (Corbin et al., 2016; Elliott, 1979; Gravel et al., 1999; Neuman et al., 2010). More relevant to the present study is consideration of the masker content, which has been shown to alter the extent to which the presence of masker impedes speech recognition (Corbin et al., 2016; Hall et al., 2002; Johnstone & Litovsky, 2006; Leibold et al., 2013; Leibold & Buss, 2013). Specifically, maskers that contain non-speech steady-state content (e.g., speech-shaped noise) are considered to impose primarily energetic masking, whereas maskers that contain speech content with amplitude modulations (e.g., multitalker speech) impose both energetic and informational masking. As cognitive and linguistic processes are known to remain immature during childhood, it is sensible that children experience greater difficulty understanding speech amidst maskers that impose greater informational masking. Consistent with this, previous research has demonstrated that younger children require more favorable SNRs to recognize speech in the presence of a two-talker speech masker as compared to a speech-shaped noise masker (Corbin et al., 2016; Hall et al., 2002; Leibold et al., 2013). When the target speech and masker are presented at a fixed SNR, younger children have been shown to achieve significantly poorer recognition of speech (i.e., consonant-vowel tokens) compared to older children and adults, with larger age-related differences observed for a two-talker speech masker than a speech-shaped noise masker (Leibold & Buss, 2013).

While the mere presence of a masker is expected to interfere with children's speech recognition, the spatial location of the target and masker have been shown to modulate the extent to which this occurs. Specifically, numerous studies have documented the benefit in speech recognition associated with increased spatial separation between a target and masker in children with normal hearing (Johnstone & Litovsky, 2006; Litovsky, 2005). This benefit is referred to as *spatial release from masking* and occurs due to interaural timing and frequency-specific level differences of the signal arriving at the two ears, which provides access to monaural head shadow and binaural interaction cues (Hawley et al., 2004; Zurek, 1993). In addition, spatial separation is expected to facilitate auditory object formation (Bregman, 1990), and therefore stream segregation, due to the reduced overlap of spectrotemporal cues between the target and masker. From a top-down perspective, spatial separation decreases informational masking, and therefore reduces dependence on attentional resources to aid in segregating, tracking, and encoding information from the target stream (Arbogast et al., 2002; Freyman et al., 1999, 2001). In fact, Arbogast et al., 2002 and others have documented that the speech recognition benefits from spatial separation are larger for primarily informational maskers (18 dB) than primarily energetic maskers (7 dB). However, the benefit of spatial separation between target and maskers is prone to the environment within which the stimuli are presented. Specifically, regardless of masker content, the benefit of spatial separation has been shown to diminish as reverberation time increases due to the degradation of interaural timing and level difference cues by the reflected energy (Hirsh, 1950; Kidd et al., 2005; Koehnke & Besing, 1996; MacKeith & Coles, 1971; Plomp, 1976).

In summary, reverberation time, masker type, and the spatial location of the masker influence speech recognition by imposing varying demands on bottom-up and top-down processes. Therefore, it is possible that the extent to which selective attention contributes to speech recognition differs based on the acoustic conditions under which speech recognition is measured and the attentional demands they impose. For instance, selective attention may contribute to speech recognition to a greater extent under conditions where it is more difficult to segregate and track target and masker streams, such as occurs with increased reverberation times or reduced spatial separation. In addition, selective attention may contribute to a greater extent to speech recognition in conditions where the masker poses greater competition for attentional resources, such as when the masker contains meaningful speech content or when the number of masker streams to ignore increases. While children's ability to selectively attend to a target speech stream is expected to be related to their age and hearing sensitivity (unaided and aided), as described in *Chapter 2*, it is possible that individual differences in selective attention may provide additional predictive value for speech recognition than age or hearing sensitivity alone, especially under acoustic conditions that impose greater attentional demands. Therefore, an enhanced understanding of the relation between selective attention and speech recognition is needed in order to provide additional insight as to why younger children and children with hearing loss are at greater risk for poor speech recognition in complex acoustic environments.

3.2.4 The Present Study

The primary objective of the present study was to test the hypothesis that children's ability to selectively attend to a target speech stream contributes to their speech recognition in complex acoustic environments. If so, children who have greater difficulty selectively attending to a target speech stream, including younger children and children with hearing loss, should demonstrate poorer speech recognition in the presence of competing auditory input. Furthermore, selective attention should relate to speech recognition to a greater extent under acoustic conditions that impose greater attentional demands, such as environments with longer reverberation times, maskers containing speech content, and co-located target and masker streams. To test these relations, children with normal hearing and children with varying degrees of hearing loss performed a speech recognition task under acoustic conditions that differed based on reverberation time, masker type, and the spatial location of the masker.

3.3 METHODS

3.3.1 Participants

Sixty-six children 5 to 12 years of age, 51 children with normal hearing (NH) and 15 children with bilateral hearing loss (HL), participated in this study. The children who participated in this study comprise a majority subset of the children who completed the study described in *Chapter 2*, and therefore meet the same eligibility and audiometric criteria as described therein. In the present study, nine children used hearing aids in both ears (HA_B), two children used a hearing aid in one ear and a cochlear implant in the other ear (HA_CI), and four children used cochlear implants in both ears (CI_B). **Figure 3.1** displays the distribution of children across ages for each hearing status and device group. The audiometric and device information for individual children with HL is consistent with that displayed in *Chapter 2, Table 2.1,* with the exception of the 10.5 year-old child with bilateral hearing aids who did not participate in this study.

Approval for all study procedures was obtained from the Institutional Review Board at Northwestern University and all children completed an informed consent/assent process prior to participation. Testing occurred within one session that lasted approximately 2 hours, including breaks. Children were compensated at a rate of $10/hour for their participation. In addition, children received a book and small prize at the conclusion of the session after trading in tickets they earned for completing each task.

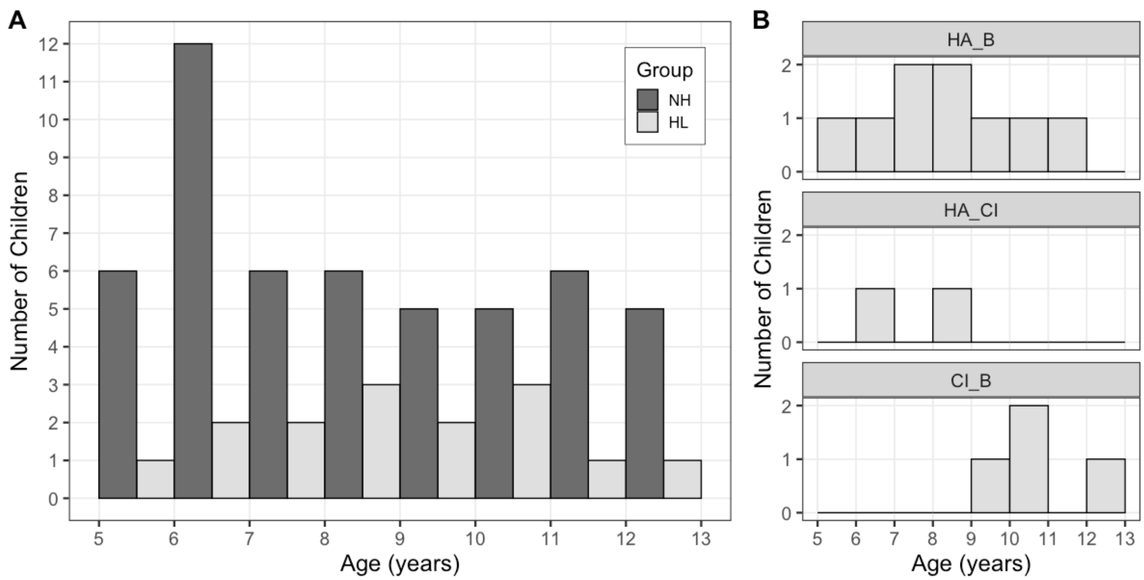


Figure 3.1: Distribution of children's ages across hearing and device groups.

3.3.2 Demographic Information and Measures of Executive Function

Children with NH and HL completed standardized measures of executive function as part of a larger study, as described in *Chapter 2, Section 2.3.2*. **Table 3.1** displays performance on these measures and demographic information for the children who participated in this study. In addition, children's response sensitivity on the auditory and visual change detection tasks used to quantify selective attention, as described in *Chapter 2*, is summarized in **Table 3.1**.

	Group	
	Children with NH	**Children with HL**
Number	51	15
Sex	Female = 21 Male = 30	Female = 7 Male = 8
Age (Years)	Mean = 8.59 | *SD* = 2.33 Range = 5.39-12.96	Mean = 8.89 | *SD* = 2.02 Range = 5.12-12.30
Unaided PTAs		
Left Ear (dB HL)	Mean = 4.71 | *SD* = 3.93 Range = -3.75-12.50	Mean = 63.03 | *SD* = 28.12 Range = 21.25-111.75
Right Ear (dB HL)	Mean = 5.07 | *SD* = 4.02 Range = -3.75-15.00	Mean = 63.70 | *SD* = 28.50 Range = 26.25-111.75
Aided PTAs		
Left Ear (dB HL)	-	Mean = 21.42 | *SD* = 4.33 Range = 11.25-28.75
Right Ear (dB HL)	-	Mean = 20.67 | *SD* = 4.20 Range = 11.25-27.50
Maternal Education Level (Years)	Mean = 15.53 | *SD* = 0.83 Range = 14-17	Mean = 14.87 | *SD* = 1.36 Range = 13-17
CHAPS Total Score	Mean = -1.86 | *SD* = 18.16 Range = -73-35	Mean = -17.20 | *SD* = 19.43 Range = -50-25
Auditory Change Detection (ACD) Task Response Sensitivity	Mean = 3.17 | *SD* = 0.95 Range = 0.76-4.72	Mean = 1.94 | *SD* = 0.95 Range = -0.57-3.31

Table 3.1: Descriptive statistics for demographic information and measures of executive function for children with normal hearing and hearing loss.

3.3.3 Acoustic Conditions

All children completed an open-set sentence recognition task in acoustic conditions that followed a crossed repeated-measures design: all levels of each acoustic condition co-occurred with all levels of the other two acoustic conditions. **Table 3.2** summarizes the acoustic conditions and the associated levels of each, which are described in the subsequent section. To measure how the content of the masker affected speech recognition, masker type was altered to consist of either speech (i.e., two-talker speech) or non-speech (i.e., speech-shaped noise) content. Reverberation time was manipulated (i.e., low; moderate; high), as described in more detail below, to quantify the extent to which reverberation disrupts speech recognition in the presence or absence of a concomitant masker. Lastly, the spatial location of the masker, when present, was modified to be either co-located with the target speech at 0 degrees or spatially separated. For the spatially separated conditions, the masker was either presented only on the right (i.e., +90 degrees) or decorrelated single-stream maskers were presented on both the right and left (i.e., +90/-90 degrees) sides of the listener. These specific acoustic conditions were selected due to the predicted differences in the dependence on top-down cognitive resources, including selective attention, expected to be observed between the levels of each condition.

Reverberation Time	Masker Type	Spatial Location (of Masker)
Low (0.49 seconds)	No Masker	-
	Speech-Shaped Noise	0 degrees
		+90 degrees
		+90/-90 degrees
	Two-Talker Speech	0 degrees
		+90 degrees
		+90/-90 degrees
Moderate (0.75 seconds)	No Masker	-
	Speech-Shaped Noise	0 degrees
		+90 degrees
		+90/-90 degrees
	Two-Talker Speech	0 degrees
		+90 degrees
		+90/-90 degrees
High (1.16 seconds)	No Masker	-
	Speech-Shaped Noise	0 degrees
		+90 degrees
		+90/-90 degrees
	Two-Talker Speech	0 degrees
		+90 degrees
		+90/-90 degrees

Table 3.2: Acoustic conditions under which children performed the speech recognition task.

3.3.4 Stimuli

Target Stimuli

Target stimuli for the speech recognition task consisted of Bamford-Kowal-Bench (BKB; Bench et al., 1979) phonetically balanced short sentences spoken by a male talker (Etymotic Research, 2010). BKB sentences were chosen because they contain vocabulary derived from language samples of young children with hearing loss as young as 5 years of age (Bamford & Bench, 1979) and are commonly used empirically and clinically to assess speech recognition in children (Boothalingam et al., 2019; Grieco-Calub et al., 2017; Magimairaj et al., 2018; Ricketts et al., 2007). Individual sentences were scaled to 65 dB SPL, root-mean-square normalized, concatenated into lists of 10 sentences, and exported in .wav audio format.

Masker Stimuli

Masker stimuli consisted of a two-talker speech masker and a speech-shaped noise masker consistent with previous studies investigating the effects of energetic and informational masking on speech recognition and cognitive processing in children (e.g., Corbin et al., 2016; Grieco-Calub et al., 2018; Leibold & Buss, 2013). The two-talker speech masker was generated by temporally overlaying streams of concatenated recorded passages from 16 child-directed science articles spoken by two female talkers. Each talker was a monolingual native American-English-speaking young adult with normal hearing who resided in Chicago, IL. Recordings were made separately in an anechoic chamber at a sampling rate of 44.1 kHz and resolution of 16 bits using a Blue Snowball Microphone connected to a MacBook Pro laptop. The recorded passages were scaled to 65 dB SPL, root-mean-square normalized, exported in .wav audio format, and randomized prior to being concatenated and overlaid. To reduce potential glimpsing opportunities, silent gaps in the 24 minute-long two-talker masker were limited to 500 ms or less.

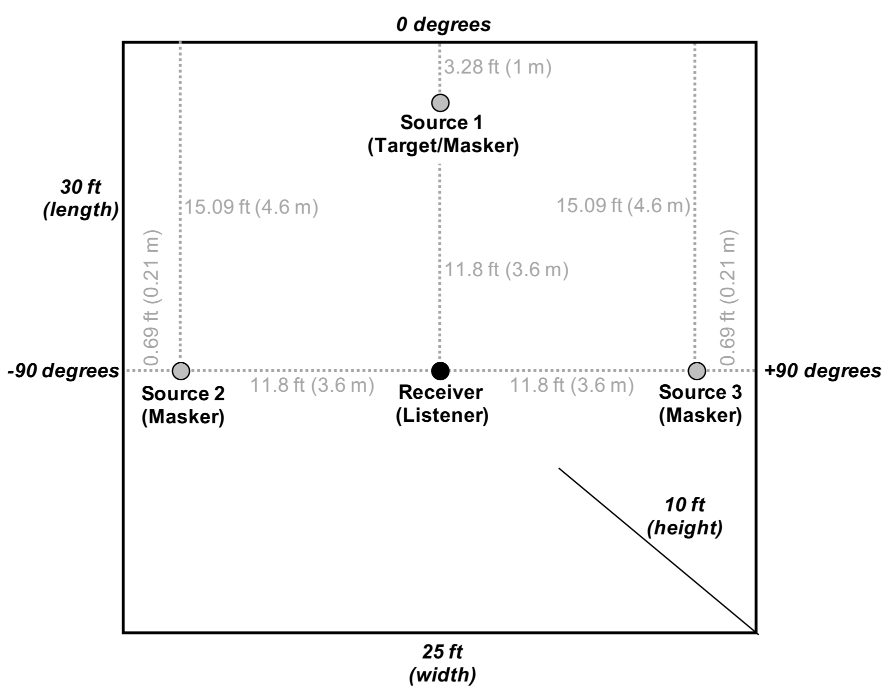
The speech-shaped noise masker was generated by multiplying the spectral envelope of the two-talker speech masker by broadband Gaussian noise. As a result, the speech-shaped noise masker had the same long-term average speech spectrum as the two-talker speech masker but contained no meaningful speech content.

Reverberation Simulation

To simulate the reverberation characteristics of real-world complex acoustic environments, target and masker stimuli were separately subjected to a multi-step modeling and processing method. The objective of the modeling approach was to simulate the reverberation characteristics, as quantified by reverberation time (T30; in seconds), of three separate rooms with low, moderate, and high broadband reverberation times to reflect fully-treated, partially-treated, and untreated real-world environments, respectively (ANSI S12.60-2010).

The modeling approach employed in the present study simulated the reverberation characteristics of real-world environments by generating soundscapes convolved with computed impulse responses to be presented in the sound field (i.e., external to the listener). This ensured that each child's unique head-related transfer function (HRTF) was accounted for without the need for stimuli to be presented over headphones. This contrasts with other reverberation simulation methods, which use binaural room impulse responses based on standard or individualized HRTFs to represent the acoustic characteristics of the convolved signal arriving at the listener's ear. As such, stimuli must be presented over headphones as sound field presentation is confounded by the additive effects of the simulated HRTF and room acoustics with those encountered in the environment during playback. Therefore, a significant benefit of the reverberation modeling approach used in this study is that children with HL were able to wear their clinical hearing devices during the experiment to listen to speech in the sound filed as they would in real-world environments.

The first step of the modeling approach used SketchUp software (Trimble Inc., 2000) to establish the physical layout of the virtual base room from which the rooms with low, moderate, and high reverberation would be simulated. The base room was 30’ × 25’ × 10’ and contained four walls, a floor, and a flat ceiling covered in acoustic panels. Next, the location of the sound sources, which reflect the location of the target and masker stimuli, as well as the receiver, which reflects the location of the listener, within each modeled room were defined using Odeon software (Odeon A/S; Naylor, 1993). To simulate the spatialization of sound sources within the reverberant environments, virtual sources were placed at -90, 0, and +90 degrees azimuth relative to the receiver. Each sound source and receiver was positioned 3.28 ft (1 m) above the floor, which approximates the ear height of a seated person. **Figure 3.2** displays an overhead schematic of the modeled base room, including the location of the sound sources and receiver.



<u>Figure 3.2</u>: Overhead schematic of the modeled base room.

Once the physical layout of the modeled base room, source locations, and receiver location were established, the absorption coefficients of the acoustic panels were manipulated in Odeon to achieve the desired broadband reverberation time for each room. **Table 3.3** displays the absorption coefficients for each modeled room, which are consistent with those of materials used to acoustically treat surfaces in real-world environments (e.g., Doelle, 1972). Applying the same absorption coefficients to all surfaces in a given room provided additional experimental control by ensuring the reverberation characteristics were consistent regardless of the modeled location of the source and receiver within the room.

Modeled Absorption Coefficients Across Rooms							
Low Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
0.26	0.26	0.26	0.26	0.27	0.28	0.26	0.22
Moderate Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
0.17	0.16	0.17	0.17	0.18	0.18	0.17	0.13
High Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
0.1	0.1	0.1	0.11	0.12	0.12	0.12	0.1

Table 3.3: Absorption coefficients for modeled rooms with low, moderate, and high reverberation.

After the absorption coefficients for each room were determined, 2-dimensional B-format impulse responses that characterize early, mid, and late azimuthal reflections from each sound source location (i.e., -90, 0, and +90 degrees) relative to the receiver were generated in Odeon for each room. Impulse responses were then separately convolved with target and masker stimuli and decoded to reflect the relative locations of the speakers in the sound field array used for testing, as described in more detail below. Finally, the convolved target and masker stimuli were combined using REAPER software (Cockos, 2006) at 0 dB SNR. The list for each test condition was exported as a separate 13-channel .wav file to be routed to specific speakers in the array within the testing environment, as described in *Section 3.3.5* below. This step was repeated with the intensity of the masker decreased by 5 dB SPL to generate stimuli at +5 dB SNR. **Table 3.4** displays the frequency-specific and broadband T30 values for each of the modeled rooms. **Figure 3.3** displays waveforms and spectrograms of a target sentence in the unprocessed and processed conditions.

Modeled Reverberation Times (T30) Across Rooms							
Low Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
0.52	0.52	0.52	0.52	0.51	0.47	0.45	0.39
Broadband T30 = 0.49 seconds							
Moderate Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
0.84	0.85	0.84	0.82	0.78	0.74	0.66	0.52
Broadband T30 = 0.75 seconds							
High Reverberation							
63 Hz	125 Hz	250 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	8000 Hz
1.43	1.43	1.42	1.28	1.15	1.08	0.88	0.59
Broadband T30 = 1.16 seconds							

Table 3.4: Frequency-specific and broadband reverberation times (T30) for modeled rooms with low, moderate, and high reverberation. Broadband T30 reflects the average reverberation time across all measured frequency bands (63 to 8000 Hz).

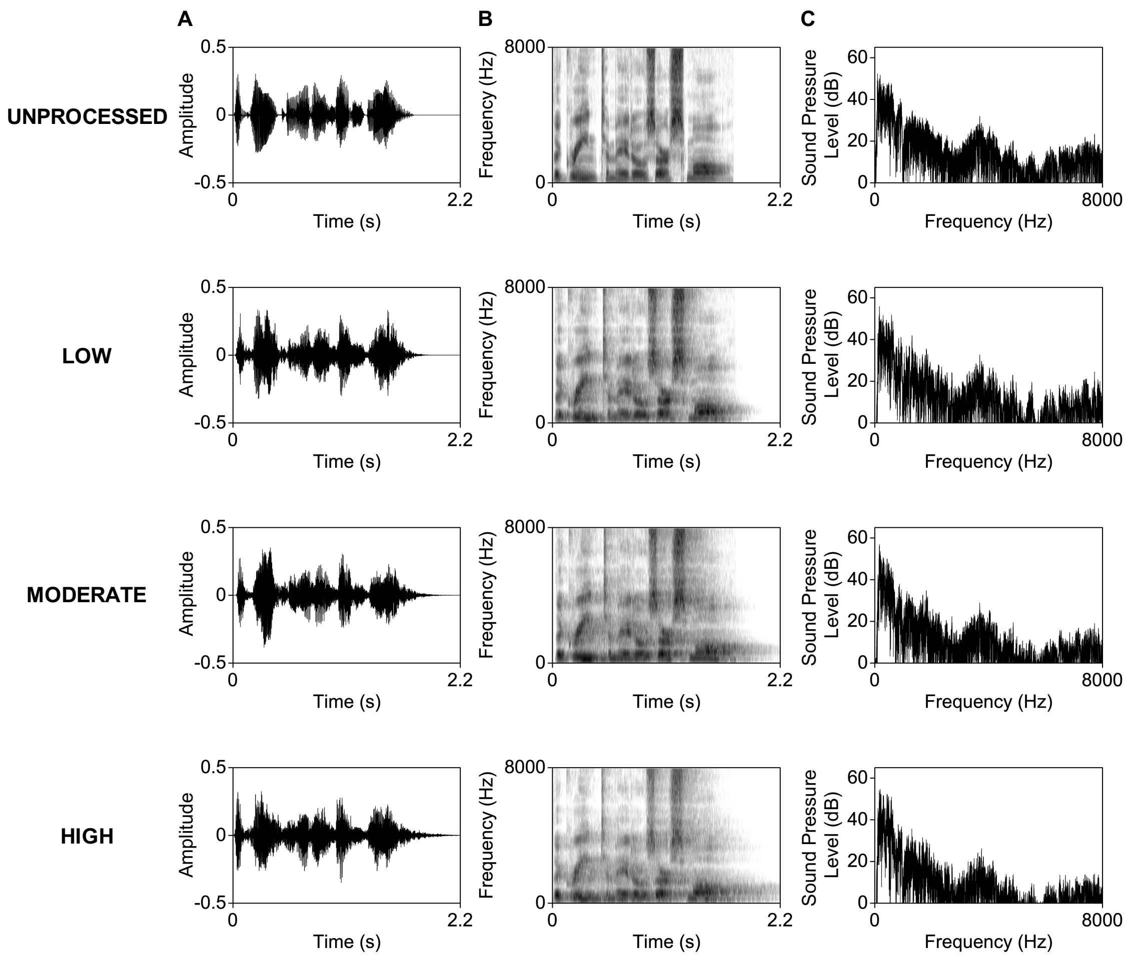


Figure 3.3: Example of target speech in the unprocessed and processed conditions. Waveforms (**A**), spectrograms (**B**), and spectrums (**C**) are displayed for the sentence "*The green tomatoes are small*" in the unprocessed condition (*top*) as well as processed conditions simulating low (*second from top*), moderate (*second from bottom*), and high (*bottom*) reverberation times.

3.3.5 Testing Apparatus

Testing occurred in a sound-attenuating room with the child and a researcher present. The room was 16'6" × 14'6" × 9'0" and housed a custom 37-speaker array (Constellation; Meyer Sound Laboratories). The speaker array was designed such that each of the four walls contained three compact loudspeakers on the lower horizontal plane (2 ft above the floor), three full-range surround loudspeakers on the upper horizontal plane (7 ft above the floor), and a low-frequency loudspeaker (subwoofer) below the lower horizontal plane (1 ft above the floor). An additional 9 loudspeakers were located on the ceiling, however these were not used in the present study. The floor was carpeted and the walls were composed of a material with a broadband absorption coefficient of 0.67. The unoccupied room acoustics were measured using sound field impulse responses using Easera software (AFMG Technologies, 2011) and exhibited a mid-band reverberation time of 0.13 seconds.

To generate the simulated acoustic environments within the room, the processed target and masker stimuli were routed to the 24 speakers in the lower- and upper-horizontal plane arrays as well as the 4 subwoofers via CueStation software version 5.5 (Meyer Sound Laboratories, 2015). While the location of the speakers in the testing environment corresponded to the modeled azimuthal output locations in Odeon, there were no physical speakers in the testing environment that reflected the elevation of the average ear height of a seated person (i.e., 3.28 ft [1 m] above the floor) as established in the model. To account for this, the first 12 channels of each .wav file were routed to the lower- *and* upper-horizontal plane speakers positioned at each azimuthal location in the array and the output level was panned to 60%-40%, respectively, to generate "virtual speakers" at the desired elevation. The remaining channel of each .wav file was routed to each of the four subwoofers. While the relative intensity of the channels within each .wav file were maintained, the overall output of the speaker array was adjusted for each model in CueStation such that all target stimuli were presented at 65 dBA in the sound field. The purpose of this intensity equalization step was to ensure that any observed differences in performance across conditions were not due to differences in the audibility of the target stimuli.

During testing, children were seated on a chair in the center of the room facing the front wall, which corresponds to the relative location of the receiver in **Figure 3.2**. A condenser microphone connected to a Tascam DR-40 Portable Digital Recorder was positioned approximately 1 ft in front of the chair to record the child's verbal responses. The researcher was positioned in the back corner of the room approximately 7 feet behind the child so as to not alter the transmission of sound from the speakers to the child.

3.3.6 Procedures

Familiarization

Prior to testing, children completed a familiarization block of 10 sentences in the same experimental setup as used during testing. Children were instructed to listen to each sentence and repeat it aloud as accurately as possible, even if it required guessing. Children were given as much time as needed to respond and no feedback was provided. Responses were scored live with a point awarded for each of the 3 or 4 predetermined key words correctly repeated for each sentence. The purpose of the first portion of the familiarization block was to ensure that the audibility of test stimuli was sufficient and that children were reliably able to listen to and repeat sentences in an open-set format. Children were presented with 4 sentences that were unprocessed and presented from 0 degrees azimuth in quiet with the baseline reverberation characteristics of the testing environment (i.e., T30 = 0.13 seconds).

In order to avoid potential floor effects for young children and children with HL, the SNR at which target and masker stimuli would be presented for each child during the speech recognition task was determined during the second portion of the familiarization block. Children were presented with six unprocessed sentences located at 0 degrees azimuth amidst a co-located masker at 0 dB SNR with the baseline reverberation characteristics of the testing environment (i.e., T30 = 0.13 seconds). The first three sentences occurred in the presence of a speech-shaped noise masker and the last three sentences were presented with a two-talker speech masker. If children scored greater than 50% correct at 0 dB SNR, they proceeded to the test blocks of the speech recognition task at 0 dB SNR. If children scored less than 50% correct at 0 dB SNR, they repeated this portion of the familiarization block with stimuli presented at +5 dB SNR. All children who completed the +5 dB SNR familiarization block scored at least 50% and completed the test blocks of the speech recognition task at +5 dB SNR (i.e., 3 children with NH; 12 children with HL).

Speech Recognition Task

Following familiarization, children completed the test blocks of the speech recognition task in the acoustic conditions outlined in **Table 3.2**. Test blocks were structured such that children completed the speech recognition task in all possible combinations of the acoustic conditions. Specifically, there were 18 conditions with a competing masker (i.e., 2 [masker type: speech-shaped noise; two-talker speech masker] × 3 [reverberation time: low, moderate, high] × 3 [spatial location: separated (+90); separated (+90/-90); co-located]) and three conditions without a masker (i.e., quiet with low, moderate, or high reverberation), resulting in a total of 21 test blocks. The instructions and scoring procedure for the test blocks were consistent with the familiarization block. Eight children with normal hearing (ages 6.26, 6.85, 6.91, 8.42, 9.34, 11.5, 11.9, and 12.66 years of age) did not complete the no-masker blocks for the sentence recognition task due to its addition to the testing protocol after the initial onset of testing. Children were presented with one list of 10 sentences in each test block for a total of 210 sentences. The sentence list assigned to each acoustic condition and the presentation order of these conditions was randomized and counterbalanced across participants. Within each test block, children's performance on the speech recognition task was quantified as the proportion of key words correctly identified out of 61 or 62 possible key words, depending on the number of predetermined key words per list. Breaks were provided after every 3 blocks during which children collected a ticket for completing the previous blocks.

3.3.7 Statistical Analyses

Linear Mixed-Effects Modeling

Before testing for the primary relation of interest between selective attention and speech recognition, linear mixed-effects modeling was used to test for the effects of age, hearing status, and acoustic condition on speech recognition (*Section 3.4.1*). Modeling was executed using the 'lme4' package for R Statistical Software (Bates et al., 2015; R Core Team, 2019). Significance was calculated using the 'lmerTest' package (Kuznetsova et al., 2017), which applies Satterthwaite's method to estimate degrees of freedom and generate *p*-values for mixed models.

There were two primary benefits of using linear-mixed effects modeling over other mean-comparison based statistical models (e.g., analyses of variance) for data that follow a repeated-measures design. Firstly, it is possible to specify random effects in addition to fixed effects (Pinheiro & Bates, 2000). Random effects are variables with levels that represent a random sample from a population (e.g., participants). Including the intercept of a random effect in a linear mixed-effects model partitions the variance associated with how a participant's score relates to the population group mean rather than building this variance into an error term (e.g., Baayen et al., 2008; Magezi, 2015; Walker et al., 2019). Secondly, linear mixed-effects modeling better accommodates data with small sample sizes or unbalanced groups by using maximum likelihood estimation and allowing varying intercepts and slopes across participants (Gałecki & Burzykowski, 2013; Muth et al., 2016; Wu, 2009). For these reasons, linear mixed-effects modeling was deemed appropriate for use in the present study.

Specifically, two linear mixed-effects models were executed to assess how age, hearing status, and acoustic condition affect speech recognition in children. The first model tested these relations for speech recognition in quiet (i.e., without a masker) while the second model did so for speech recognition in the presence of a competing masker. The dependent variable for each model was speech recognition transformed into rationalized arcsine units (RAU; Studebaker, 1985). A common characteristic of proportion-based data is that the standard deviations of group scores are strongly correlated with the mean, which violates the homogeneity of variance assumption of parametric statistical tests. Transforming the data into RAU overcomes this violation by first converting the raw data to arcsine-transformed scores and then linearly transforming the scores to units that approximate percentages to maintain interpretability (Oleson et al., 2019; Studebaker, 1985; Studebaker et al., 1999). To aid in interpretation, all figures display the non-transformed (i.e., raw) speech recognition scores.

The fixed-effect predictor variables included in each model were age (continuous), group (NH; HL), and reverberation time (low; moderate; high) with masker type (speech-shaped noise; two-talker speech) and spatial location (co-located; separated [+90/-90]; separated [+90]) added to the model with speech recognition in the presence of a masker. For this model, a top-down model comparison procedure was utilized to determine the random effects structure that resulted in the most parsimonious model. In considering the crossed repeated-measures design of this study, random effects of reverberation time (for the no-masker and masker models) as well as masker type and spatial location (for the masker model) were individually removed from each model and, through successive model comparison, were not found to significantly improve model fit. Therefore, participant was the only factor included as a random factor in each model to account for shared variance due to similarities in speech recognition within a child across acoustic conditions.

Finally, to increase the interpretability of estimated coefficients in each model, age was centered prior to analysis by subtracting the mean value from each observed value (Gelman & Hill, 2007). Orthogonal sum to zero contrast codes (i.e., [1, 0, -1] or [1, -1] where -1 is the reference level and 1 is the comparison level) were applied to all categorical variables (i.e., group; reverberation time; masker type; spatial location). This method of contrast coding allowed for main effect estimates to be evaluated as the grand mean (i.e., collapsed across all levels of the other categorical variables) rather than the mean of the reference levels alone. Reference levels for each model were set as the level of each categorical variable expected to result in the poorest speech recognition (i.e., HL; high reverberation time; two-talker speech masker; co-located target and masker). Therefore, model estimates with positive values indicated relations that were consistent with these predictions.

Hierarchical Linear Regression Analyses

Following the linear mixed-effects modeling, hierarchical linear regression analyses were used to directly test the primary hypothesis of this study – that children's ability to selectively attend to a target speech stream contributes to their speech recognition in complex acoustic environments (*Section 3.4.2*). The observed differences in speech recognition between children with NH and HL across the tested age range in conjunction with the observed age- and group-related differences in selective attention observed in *Chapter 2* provide support for this hypothesis. Regressions were executed using the 'stats' package for R Statistical Software (R Core Team, 2019).

Two regression models were implemented to investigate the factors that contributed to children's speech recognition in each acoustic condition. To test the hypothesis that children's ability to selectively attend to a target stream significantly relates to their ability to understand speech in complex acoustic environments, **Model 1** included children's performance on the auditory change detection task from *Chapter 2* (i.e., ACD response sensitivity). ACD response sensitivity is an empirical measure of children's ability to selectively attend to a target speech stream and inhibit attention to a competing stream. The presence of ACD response sensitivity as a significant predictor in Model 1 provided evidence of the hypothesized relation between selective attention and speech recognition.

Additionally, **Model 2** assessed whether any of the variance in speech recognition accounted for by ACD response sensitivity in Model 1 persisted after accounting for age and aided PTA, which were shown to relate to ACD response sensitivity in *Chapter 2* ($r = 0.45$, $p < 0.001$ and $r = -0.54$, $p < 0.001$, respectively; *Appendix A.3.3*). If ACD response sensitivity no longer remained a significant predictor in Model 2, it suggested that age and/or aided PTA accounted for similar variance in speech recognition as ACD response sensitivity due to the expected interrelations among these processes. If, however, ACD response sensitivity remained a significant predictor in Model 2 despite the inclusion of age and aided PTA, it suggested that ACD response sensitivity was uniquely predictive of speech recognition over and above age and aided PTA.

In addition to assessing the individual contributions of predictor variables, models were successively compared to determine whether ACD response sensitivity, age, and aided PTA

(**Model 2**) accounted for significantly more variance in speech recognition than ACD response sensitivity alone (**Model 1**). Variance inflation factors (VIFs) were computed for each predictor variable and were found to be well within acceptable limits (i.e., less than 5) for each model indicating the absence of multicollinearity (Robinson & Schumacker, 2009; Yoo et al., 2014).

3.4 RESULTS

3.4.1 Linear Mixed-Effects Modeling: Effects of Age, Hearing Status, Reverberation Time, Masker Type, and Spatial Location on Speech Recognition

The primary objective of this study was to assess the relation between children's selective attention and their speech recognition in complex acoustic environments. As an initial step in this analysis, linear mixed-effects modeling was used to test the predicted effects of age, hearing status, and the acoustic characteristics of the environment on speech recognition in quiet (i.e., no masker) and in the presence of a competing masker.

Factors that Influence Speech Recognition in Quiet

Consistent with previous studies, children's age and hearing status, as well as the amount of reverberation in the environment, affected children's speech recognition (**Figure 3.4 A**; *Appendix A.3.1*). Specifically, better speech recognition scores were observed with increasing age (estimate = 3.17, $p < 0.001$) as well as for children with NH as compared to children with HL (estimate = 18.27, $p < 0.001$). Reverberation time significantly affected speech recognition overall, with greater differences observed between the low and moderate reverberation times (estimate = 14.60, $p < 0.001$) than the moderate and high reverberation times (estimate = 5.16, $p < 0.01$). However, the presence of a significant group-by-reverberation time interaction suggests that the extent to which increasing reverberation time affected speech recognition differed for children with NH and HL. Specifically, increasing the reverberation time from low to moderate resulted in a larger decrease in speech recognition for children with HL than children with NH (estimate [NH] = 9.50, $p < 0.001$; estimate [HL] = 19.70, $p < 0.01$). When reverberation time increased from moderate to high, further detriments in speech recognition were observed for children with NH (estimate = 7.26, $p < 0.01$) but not children with HL (estimate = 3.06, $p = 0.999$), though the extent of this difference was less than that observed between low and moderate reverberation times.

The absence of a significant three-way interaction among age, group, and reverberation time suggests that these relations were consistent across the tested age range. Together, these findings indicate that increased reverberation time negatively affected speech recognition in quiet, and that the greatest detriment occurred when reverberation time increased from low (i.e., broadband T30 of 0.49 seconds) to moderate (i.e., broadband T30 of 0.75 seconds) for children with NH and HL regardless of age.

Factors that Influence Speech Recognition in the Presence of a Masker

The experimental design of this study provided a way to evaluate both the independent and synergistic effects of reverberation time, masker type, and the spatial location of the masker on children's speech recognition. The following analyses investigated the variable effects of these acoustic characteristics as well as children's age and hearing status on speech recognition.

Main Effects

Consistent with previous results, a linear mixed-effects model revealed significant main effects of age, hearing status, reverberation time, masker type, and spatial location on speech recognition in the presence of a simultaneous masker (*Appendix A.3.2*; **Figure 3.4 B-D** and **Figure 3.5**). Specifically, there were significant main effects of age (estimate = 3.048, $p < 0.001$)

and group (estimate = 6.069, $p < 0.001$), suggesting that speech recognition improved with age and was better for children with NH compared to children with HL. Results also revealed a main effect of reverberation time (F = 103.036, $p < 0.01$). Differences in speech recognition were observed between low and high reverberation times (estimate = 10.670, $p < 0.001$) as well as between moderate and high reverberation times (estimate = 9.690; $p < 0.001$) but not between low and moderate reverberation times (estimate = 0.980, $p = 0.744$). Additionally, as predicted, a main effect of masker type was observed (F = 228.301, $p < 0.01$) such that children achieved better speech recognition in the presence of a speech-shaped noise masker than in the presence of a two-talker speech masker (estimate = 5.070, $p < 0.001$). Lastly, there was a main effect of spatial location (F = 174.916, $p < 0.01$). Specifically, children achieved significantly better speech recognition when one masker was separated from the target speech (+90) relative to when two maskers were separated from the target speech ([+90/-90]; estimate = 14.600, $p < 0.001$) and when the target speech and masker were co-located (estimate = 11.600, $p < 0.001$). Notably, children were unable to benefit from spatial separation when two masker streams were present, as demonstrated by significantly lower speech recognition in the separated (+90/-90) condition relative to the co-located (estimate = -3.000, $p < 0.01$) condition.

In addition to the observed main effects, the presence of significant interactions between reverberation time, masker type, and spatial location suggests that, as predicted, the effects of these factors on children's speech recognition were more complex. The following analyses were restricted to interactions with age and group because these factors have been previously shown to relate to selective attention (see *Chapter 2*), which pertains to the prediction of the present study that selective attention contributes to children's speech recognition in complex acoustic environments.

Interactions with Age

The effect of masker type was found to differ as a function of age, as demonstrated by a significant age-by-masker type interaction (F = 10.074, $p < 0.01$). Specifically, pairwise comparisons revealed a greater difference in speech recognition between the speech-shaped noise and two-talker speech maskers for the youngest end of the age range (estimate = 13.860, $p < 0.001$) as compared to the oldest end of the age range (estimate = 5.620, $p < 0.01$), which is consistent with previous studies (Corbin et al., 2016; Hall et al., 2002; Leibold & Buss, 2013). There were no significant interactions between age and reverberation time or spatial location.

Interactions with Group

The effect of masker type was also found to differ as a function of group (F = 43.740, $p < 0.001$). Pairwise comparisons revealed that children with HL demonstrated a greater difference in speech recognition between the speech-shaped noise and two-talker speech maskers than children with NH (estimate [NH] = 5.700, $p < 0.001$; estimate [HL] = 14.600, $p < 0.001$; **Figure 3.4 C**). These findings suggest that the speech recognition of children with HL was more adversely affected by the presence of the two-talker speech masker relative to children with NH. Furthermore, the presence of an age-by-group-by-masker type interaction (F = 5.329, $p < 0.05$) suggests that the degree to which the difference in speech recognition between masker types changed across the age range varied across groups. Specifically, while the difference in speech recognition between the speech-shaped noise and two-talker speech maskers appeared to decrease with age for children with NH, this age-related change was not observed for children with HL.

Additionally, the effect of spatial location was found to differ as a function of group (F = 14.006, $p < 0.001$). Pairwise comparisons revealed that children with NH benefitted from the spatial separation of the target speech and masker positioned at +90 degrees relative to the co-located condition to a greater extent than children with HL (estimate[NH] = 15.812, $p < 0.001$; estimate[HL] = 7.308, $p < 0.001$). The differences between children's speech recognition in the co-located condition and the spatially-separated condition containing two masker streams presented at +90/-90 degrees were either not significant (estimate[NH] = -1.670, $p = 0.573$) or trending toward significant (estimate[HL] = -4.325, $p = 0.058$). Together, these results suggest that speech recognition improved for all children when the target speech and masker were spatially separated; however, this improvement was greater for children with NH than children with HL and was eliminated for all children when the two masker streams were spatially separated at +90/-90 degrees (**Figure 3.4 D**).

Lastly, the extent to which reverberation time affected speech recognition in the presence of a competing masker was similar for children with NH and HL (**Figure 3.4 B**). Specifically, while there was a significant group-by-reverberation time interaction for speech recognition in quiet, this relation was trending toward significant but not significant for speech recognition in the presence of a competing masker ($F = 2.909$, $p = 0.055$).

Interaction Between Group, Reverberation Time, Masker Type, and Spatial Location

In addition to the two- and three-way interactions with age and group described above, there was a significant four-way interaction between group, reverberation, masker type, and spatial location ($F = 0.438$, $p < 0.05$; *Appendix A.3.2*). The presence of this interaction is of theoretical interest as it underscores the complexity of the relations among reverberation time, masker type, spatial location, and speech recognition in children with NH and HL. Specifically, the extent to which one aspect of the acoustic environment – such as the reverberation time, the content of the masker, or the spatial location of the target speech and masker – impacted speech recognition differed based on the concomitant acoustic characteristics of the environment as well as children's hearing status. While fully exploring the multiplicity of pairwise comparisons between levels of each acoustic condition that underlie this interaction spans beyond the primary objective of this study, considering whether the various combinations of reverberation time, masker type, and spatial location differentially affect speech recognition in children with NH and HL was of interest.

Therefore, pairwise comparisons were completed to compare the speech recognition of children with NH and HL within each acoustic condition, which represented all possible combinations of reverberation time, masker type, and spatial location. These relations are shown in **Figure 3.5**, which displays the average speech recognition scores for children with NH and HL for each acoustic condition. Results revealed no significant differences in speech recognition between children with NH and HL in the acoustic conditions containing a co-located masker after Bonferroni-correcting for multiple comparisons. However, differences in speech recognition between NH and HL were observed for the speech-shaped noise and two-talker speech maskers for a subset of acoustic conditions during which the target speech and masker were spatially separated and presented at +90 or +90/-90 degrees (*Appendix A.3.2*). In these conditions, children with NH achieved better speech recognition than children with HL, which may partially reflect underlying differences in the ability of children with NH and HL to selectively attend to a target speech stream and inhibit attention to spatially-separated competing maskers.

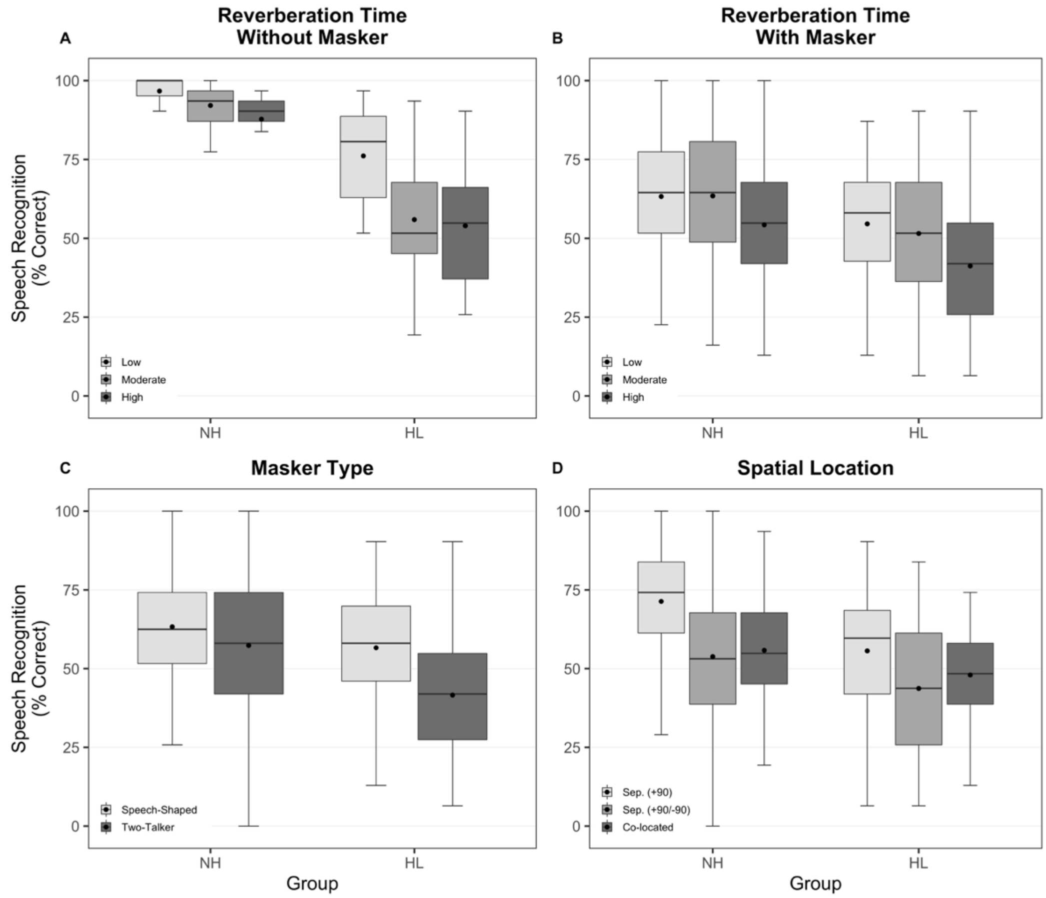


Figure 3.4: Performance on the speech recognition task across acoustic conditions and groups. Each panel (**A-D**) reflects the effect of a specific acoustic manipulation collapsed across all levels of the other conditions for children with NH (*left*) and HL (*right*).

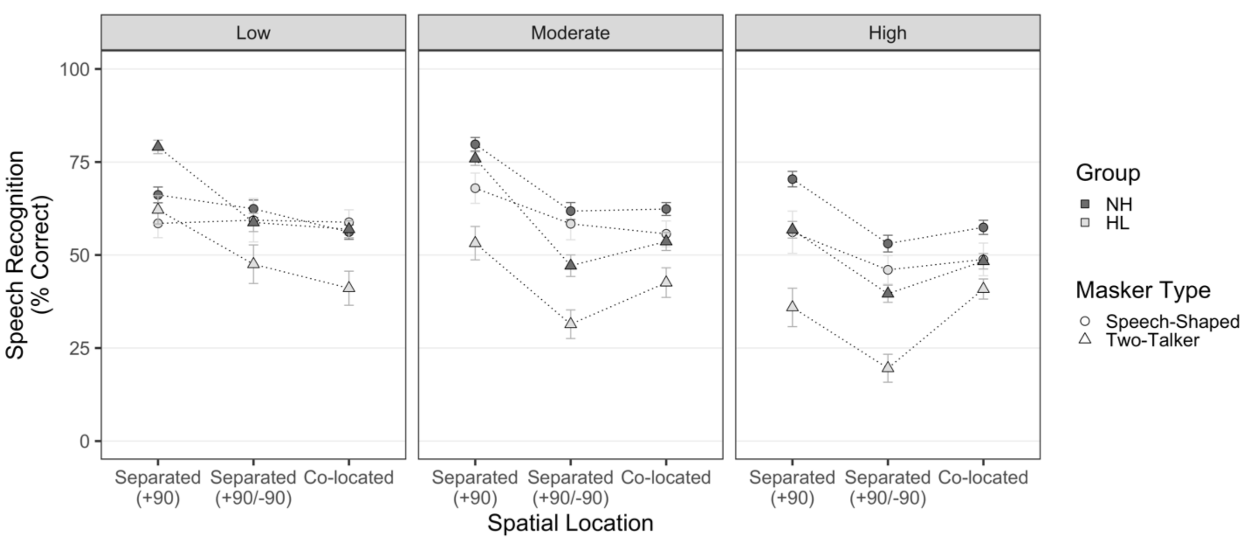


Figure 3.5: Average performance on the speech recognition task for each group across acoustic conditions. Points reflect average speech recognition for children with NH (*dark gray*) and HL (*light gray*) across spatial locations for low, moderate, and high reverberation times. Error bars reflect ± 1 standard error.

3.4.2 Hierarchical Linear Regression: Relation Between Selective Attention and Speech Recognition

The results of the linear mixed effects models described above indicate that, as predicted, reverberation time, masker type, and spatial location significantly affected children's speech recognition, and that the extent to which this occurred differed based on the other acoustic characteristics of the environment. Furthermore, in considering that these acoustic conditions affected speech recognition to a greater extent in certain groups of children (i.e., younger children and children with HL) than others (i.e., older children and children with NH), it is possible that children's reliance on selective attention for speech recognition was modulated by the acoustic characteristics of the environment and the attentional demands they imposed. Specifically, as predicted, children who have greater difficulty selectively attending to a target speech stream should have poorer speech recognition, especially under acoustic conditions that impose greater attentional demands. The following analyses directly tested these relations.

Overview of Model Outputs

Table 3.5 displays the results of the individual parameters within each regression model (i.e., Model 1 and Model 2) for each acoustic condition. The unstandardized estimate (*B*) reflects the strength and directionality of the change in speech recognition associated with a one unit increase in the predictor variable. **Table 3.6** displays the results of each model as well as the model comparisons for each acoustic condition. R^2 reflects the cumulative proportion of variance in speech recognition accounted for by the predictor variables in each model while ΔR^2 represents the additional proportion of variance accounted for by Model 2 relative to Model 1. Model 1 and Model 2 were statistically compared to determine whether age, aided PTA, and ACD response sensitivity accounted for greater variance in speech recognition than ACD response sensitivity alone. If age and aided PTA accounted for significant variance in speech recognition over and above their effect on selective attention, then Model 2 should account for significantly more variance in speech recognition than Model 1.

		Low Reverberation				Moderate Reverberation				High Reverberation			
		Estimate	SE	t	*p*	Estimate	SE	t	*p*	Estimate	SE	t	*p*
A. No Masker													
M1	*ACD Resp. Sensitivity*	9.398	1.732	5.425	**<0.001**	10.392	2.371	4.383	**<0.001**	10.237	2.011	5.090	**<0.001**
M2	*ACD Resp. Sensitivity*	1.368	1.563	0.875	0.385	0.477	2.232	0.214	0.832	1.251	1.701	0.735	0.466
	Age	1.894	0.670	2.827	**<0.01**	1.670	0.957	1.746	0.087	1.431	0.729	1.962	0.055
	Aided PTA	-1.734	0.203	-8.541	**<0.001**	-2.301	0.290	-7.936	**<0.001**	-2.105	0.221	-9.524	**<0.001**
B. Speech-Shaped Noise Masker; Co-located													
M1	*ACD Resp. Sensitivity*	3.136	1.458	2.151	**<0.05**	4.281	1.301	3.291	**<0.01**	7.213	1.359	5.306	**<0.001**
M2	*ACD Resp. Sensitivity*	3.736	1.913	1.953	0.055	3.288	1.770	1.858	0.068	6.402	1.847	3.466	**<0.001**
	Age	1.095	0.772	1.419	0.161	0.859	0.714	1.204	0.233	0.922	0.745	1.237	0.221
	Aided PTA	0.414	0.238	1.737	0.087	-0.048	0.220	-0.219	0.828	0.013	0.230	0.056	0.955
C. Speech-Shaped Noise Masker; Separated (+90)													
M1	*ACD Resp. Sensitivity*	4.771	1.631	2.926	**<0.01**	6.135	1.824	3.365	**<0.01**	7.055	1.921	3.673	**<0.001**
M2	*ACD Resp. Sensitivity*	1.769	2.130	0.831	0.409	2.875	2.431	1.183	0.242	3.476	2.553	1.362	0.178
	Age	2.227	0.859	2.592	**<0.05**	1.120	0.981	1.142	0.258	1.815	1.030	1.762	0.083
	Aided PTA	-0.234	0.265	-0.884	0.380	-0.564	0.303	-1.864	0.067	-0.479	0.318	-1.509	0.137
D. Speech-Shaped Noise Masker; Separated (+90/-90)													
M1	*ACD Resp. Sensitivity*	7.875	1.879	4.191	**<0.001**	4.394	1.777	2.472	**<0.05**	4.830	1.648	2.930	**<0.001**
M2	*ACD Resp. Sensitivity*	7.276	2.475	2.940	**<0.05**	3.696	2.402	1.538	0.129	3.653	2.255	1.620	0.110
	Age	1.927	0.998	1.930	0.058	1.294	0.969	1.335	0.187	0.732	0.910	0.805	0.424
	Aided PTA	0.306	0.308	0.994	0.324	0.130	0.299	0.435	0.665	-0.125	0.281	-0.447	0.657
E. Two-Talker Speech Masker; Co-located													
M1	*ACD Resp. Sensitivity*	8.676	1.640	5.291	**<0.001**	9.132	1.574	5.803	**<0.001**	7.343	1.243	5.909	**<0.001**
M2	*ACD Resp. Sensitivity*	3.811	2.039	1.868	0.066	5.685	1.940	2.931	**<0.01**	5.150	1.549	3.325	**<0.01**
	Age	2.744	0.823	3.335	**<0.01**	3.058	0.783	3.907	**<0.001**	2.268	0.625	3.630	**<0.001**
	Aided PTA	-0.586	0.254	-2.307	**<0.05**	-0.150	0.242	-0.621	0.537	-0.018	0.193	-0.095	0.925
F. Two-Talker Speech Masker; Separated (+90)													
M1	*ACD Resp. Sensitivity*	6.782	1.636	4.144	**<0.001**	10.157	1.568	6.476	**<0.001**	11.185	1.616	6.920	**<0.001**
M2	*ACD Resp. Sensitivity*	2.488	2.095	1.188	0.240	5.195	1.903	2.730	**<0.01**	6.729	2.059	3.268	**<0.01**
	Age	1.414	0.845	1.673	0.100	1.172	0.768	1.527	0.132	2.049	0.831	2.467	**<0.05**
	Aided PTA	-0.757	0.261	-2.904	**<0.01**	-0.985	0.237	-4.158	**<0.001**	-0.647	0.256	-2.525	**<0.05**
G. Two-Talker Speech Masker; Separated (+90/-90)													

M1	*ACD Resp. Sensitivity*	9.919	1.700	5.835	**<0.001**	9.642	2.040	4.726	**<0.001**	9.259	1.731	5.350	**<0.001**
M2	*ACD Resp. Sensitivity*	5.654	2.110	2.679	**<0.01**	3.971	2.516	1.579	0.120	2.451	2.008	1.221	0.227
	Age	3.179	0.851	3.734	**<0.001**	3.842	1.015	3.786	**<0.001**	2.825	0.810	3.489	**<0.001**
	Aided PTA	-0.330	0.263	-1.254	0.214	-0.530	0.313	-1.691	0.096	-1.061	0.250	-4.248	**<0.001**

Table 3.5: Results of the hierarchical linear regression models for conditions containing a speech-shaped noise masker.

	Model 1 (M1)				**Model 2 (M2)**				**M2 vs. M1**	
	R^2	**ΔR^2**	**F**	***p***	**R^2**	**ΔR^2**	**F**	***p***	***F***	***p***
A. Low Reverberation										
No Masker	0.345	-	29.43	**<0.001**	0.722	0.377	46.8	**<0.001**	36.713	**<0.001**
SSN Masker										
Co-located	0.067	-	4.625	**<0.05**	0.153	0.086	3.736	**<0.05**	3.138	0.050
Sep. (+90)	0.118	-	8.561	**<0.01**	0.206	0.088	5.356	**<0.01**	3.428	**<0.05**
Sep. (+90/-90)	0.215	-	17.56	**<0.001**	0.282	0.067	8.099	**<0.001**	2.856	0.065
TTS Masker										
Co-located	0.304	-	28.00	**<0.001**	0.432	0.128	15.71	**<0.001**	6.959	**<0.01**
Sep. (+90)	0.212	-	17.18	**<0.001**	0.318	0.106	9.634	**<0.001**	4.835	**<0.05**
Sep. (+90/-90)	0.347	-	34.05	**<0.001**	0.469	0.122	18.25	**<0.001**	7.103	**<0.01**
B. Moderate Reverberation										
No Masker	0.255	-	19.21	**<0.001**	0.657	0.402	34.4	**<0.001**	31.529	**<0.001**
SSN Masker										
Co-located	0.145	-	10.83	**<0.01**	0.164	0.019	4.063	**<0.05**	0.725	0.489
Sep. (+90)	0.150	-	11.32	**<0.01**	0.203	0.053	5.259	**<0.01**	2.043	0.138
Sep. (+90/-90)	0.087	-	6.111	**<0.05**	0.120	0.033	2.812	**<0.05**	1.149	0.324
TTS Masker										
Co-located	0.345	-	33.67	**<0.001**	0.474	0.129	18.66	**<0.001**	7.648	**<0.001**
Sep. (+90)	0.396	-	41.94	**<0.001**	0.531	0.135	23.35	**<0.001**	8.892	**<0.001**
Sep. (+90/-90)	0.259	-	22.34	**<0.001**	0.405	0.146	14.06	**<0.001**	7.617	**<0.01**
C. High Reverberation										
No Masker	0.316	-	25.900	**<0.001**	0.745	0.429	52.650	**<0.001**	45.456	**<0.001**
SSN Masker										
Co-located	0.306	-	28.15	**<0.001**	0.323	0.017	9.871	**<0.001**	0.814	0.448
Sep. (+90)	0.174	-	13.49	**<0.001**	0.230	0.056	6.171	**<0.001**	2.248	0.114
Sep. (+90/-90)	0.118	-	8.587	**<0.01**	0.129	0.011	3.05	**<0.05**	0.366	0.695
TTS Masker										
Co-located	0.353	-	34.92	**<0.001**	0.469	0.116	18.28	**<0.001**	6.799	**<0.001**
Sep. (+90)	0.428	-	47.89	**<0.001**	0.510	0.082	21.51	**<0.001**	5.189	**<0.01**
Sep. (+90/-90)	0.309	-	28.63	**<0.001**	0.509	0.200	21.44	**<0.001**	12.640	**<0.001**

Table 3.6: Comparisons between Model 1 and Model 2 for all acoustic conditions. SSN = speech-shaped noise masker; TTS = two-talker speech masker

Relation between Selective Attention and Speech Recognition (Model 1)

As described in *Section 3.3.7*, Model 1 directly tested the hypothesis that children's ability to selectively attend to a target speech stream contributes to their speech recognition in complex acoustic environments. The results of Model 1 were consistent with this hypothesis by demonstrating that ACD response sensitivity accounted for significant variance in speech recognition regardless of the reverberation time, masker presence or type, and spatial location of the masker, if present (**Table 3.5; M1)**. Notably, the proportion of variance in speech recognition accounted for by ACD response sensitivity, as quantified by R^2, appeared to differ across acoustic conditions. Specifically, for the no-masker conditions, ACD response sensitivity accounted for 34%, 26%, and 32% of the variance in speech recognition in the low, moderate, and high reverberation conditions, respectively. The variance in speech recognition accounted for by ACD response sensitivity across reverberation times and spatial locations ranged from 7% to 31% for conditions containing a speech-shaped noise masker and from 21% to 43% for conditions containing a two-talker speech masker. The wide range in variance accounted for by ACD response sensitivity across acoustic conditions containing a masker is consistent with the idea that the attentional demands associated with understanding speech differ based on the acoustic characteristics of the environment.

Specifically, ACD response sensitivity appeared to account for greater variance in speech recognition in conditions containing a two-talker speech masker (**Figure 3.6**; *dark gray triangles*) than those containing a speech-shaped noise masker (**Figure 3.6**; *light gray circles*) regardless of reverberation time or spatial location. This may reflect greater attentional demands associated with selectively attending to target speech amidst a masker that imposes greater informational masking. Second, across masker types, the proportion of variance in speech recognition accounted for by ACD response sensitivity appeared to differ based on the reverberation time and spatial location of the masker. For instance, in the low reverberation condition, ACD response sensitivity accounted for the largest proportion of variance when the target and either masker (i.e., speech-shaped noise masker or two-talker speech masker) were spatially separated at +90/-90 degrees. These findings may reflect the additional attentional demands associated with inhibiting attention to two spatially-separated single-stream maskers compared to a co-located masker or a masker presented at only +90 degrees when reverberation is low. In contrast, ACD response sensitivity accounted for the smallest proportion of variance in speech recognition when the target and either masker were separated at +90/-90 degrees in the moderate and high reverberation conditions. A possible reason for this may be that, in the presence of longer reverberation times, the spectrotemporal content of the two spatially-separated single-stream maskers “smear” and become perceptually more diffuse, which may relieve competition for attentional resources. An opposite pattern of results was observed for conditions containing a two-talker speech masker that was spatially separated at +90 degrees in the presence of moderate and high reverberation: the amount of variance in speech recognition accounted for by ACD response sensitivity was greater relative to the low reverberation condition. This may reflect an increase in the attentional demands associated with resolving target speech degraded by reverberation while simultaneously inhibiting attention to a localized and, therefore, more perceptually salient, competing masker, especially one that contains speech content.

Though observational in nature, considering the trends in R^2 provides valuable insight regarding how the extent to which children’s ability to selectively attend to a target speech stream differed across acoustic conditions and what this may suggest about attentional demands.

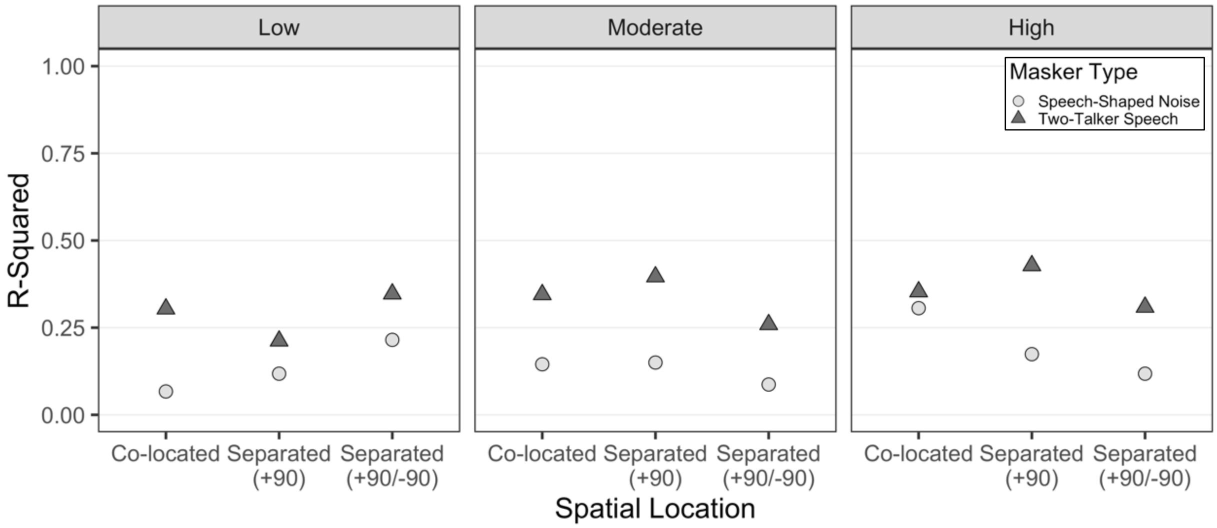


Figure 3.6: Comparison of R^2 values for Model 1 across acoustic conditions containing a masker. Acoustic conditions differed based on reverberation time, masker type, and the spatial location of the masker.

Relation between Selective Attention and Speech Recognition Accounting for Age and Aided PTA (Model 2)

While the results of Model 1 demonstrated that, as hypothesized, children's ability to selectively attend to a target speech stream accounted for variance in speech recognition across acoustic conditions, the results of Model 2 indicated whether the variance in speech recognition accounted for by age and aided PTA overlapped with the variance accounted for by ACD response sensitivity. Additionally, the results of the model comparisons for each acoustic condition revealed whether the total proportion of variance in speech recognition was greater when age and aided PTA were added to the model (Model 2) as compared to ACD response sensitivity alone (Model 1; **Table 3.6**).

As expected based on the relations among age, aided PTA, and ACD response sensitivity described in *Chapter 2*, the addition of age and aided PTA to the individual models accounted for similar variance in speech recognition as ACD response sensitivity for the majority of acoustic conditions (i.e., 14 out of 21). This was indicated by ACD response sensitivity losing its significance as a predictor in Model 2 and was observed for all acoustic conditions containing no masker, the majority (i.e., 7 out of 9) of conditions containing a speech-shaped noise masker, and a minority (i.e., 4 out of 9) of conditions containing a two-talker speech masker (**Table 3.5; M2)**. These findings suggest that the predictability of ACD response sensitivity for speech recognition observed in Model 1 generally reflected the influence of children's age and/or hearing status.

For the remaining subset of acoustic conditions (i.e., 7 out of 21), ACD response sensitivity remained a significant predictor of speech recognition when age and aided PTA were added to Model 2. These results suggest that individual differences in children's ability to selectively attend to a target stream could not be fully accounted for by age and aided PTA. As described above, these conditions are of particular interest as they indicate situations during which children's ability to selectively attend to a target stream and understand speech was better/poorer than expected based on their age and/or aided PTA. Notably, these conditions all contained a speech-shaped noise or two-talker speech masker. **Figure 3.7** displays children's speech recognition in these acoustic conditions as a function of ACD response sensitivity. While the strength of this relation differed across conditions, these results suggest that children's ability to selectively attend to a target speech stream was significantly predictive of speech recognition over and above age and aided PTA.

In addition to considering whether age and aided PTA accounted for similar variance in speech recognition as ACD response sensitivity, it was necessary to consider whether the model containing all three factors (Model 2) accounted for a significantly greater proportion of variance in speech recognition than ACD response sensitivity alone (Model 1). Significant differences between Model 1 and Model 2 were observed for all conditions containing no masker as well as conditions containing a two-talker speech masker (**Table 3.6**). These findings suggest that greater variance in children's speech recognition was accounted for when other perceptual and cognitive factors related to their age and hearing status were considered in addition to selective attention. For the majority of conditions containing a speech-shaped noise masker (i.e., 8 out of 9), however, there were no differences in the proportion of variance accounted for by Model 1 and Model 2. In conjunction with the observation that the amount of variance accounted for in Models 1 and 2 was lower for these conditions than the no-masker or two-talker masker conditions overall (**Table 3.6**; **Figure 3.6**), these findings suggest that factors other than ACD response sensitivity, age, and aided PTA may have primarily contributed to children's ability to understand speech amidst a speech-shaped noise masker.

Consistent with the primary hypothesis of this study, the results from the hierarchical linear regression analyses demonstrated that children's ability to selectively attend to a target speech stream significantly related to speech recognition regardless of the acoustic condition. Additionally, as predicted, the extent to which individual differences in ACD response sensitivity accounted for variance in speech recognition differed based on reverberation time, masker type, and the spatial location of the masker. Lastly, while age and aided PTA accounted for similar variance in speech recognition as ACD response sensitivity in the majority of acoustic conditions, individual differences in children's ability to selectively attend to a target speech stream were shown to be uniquely predictive of speech recognition for a subset of conditions containing a competing masker.

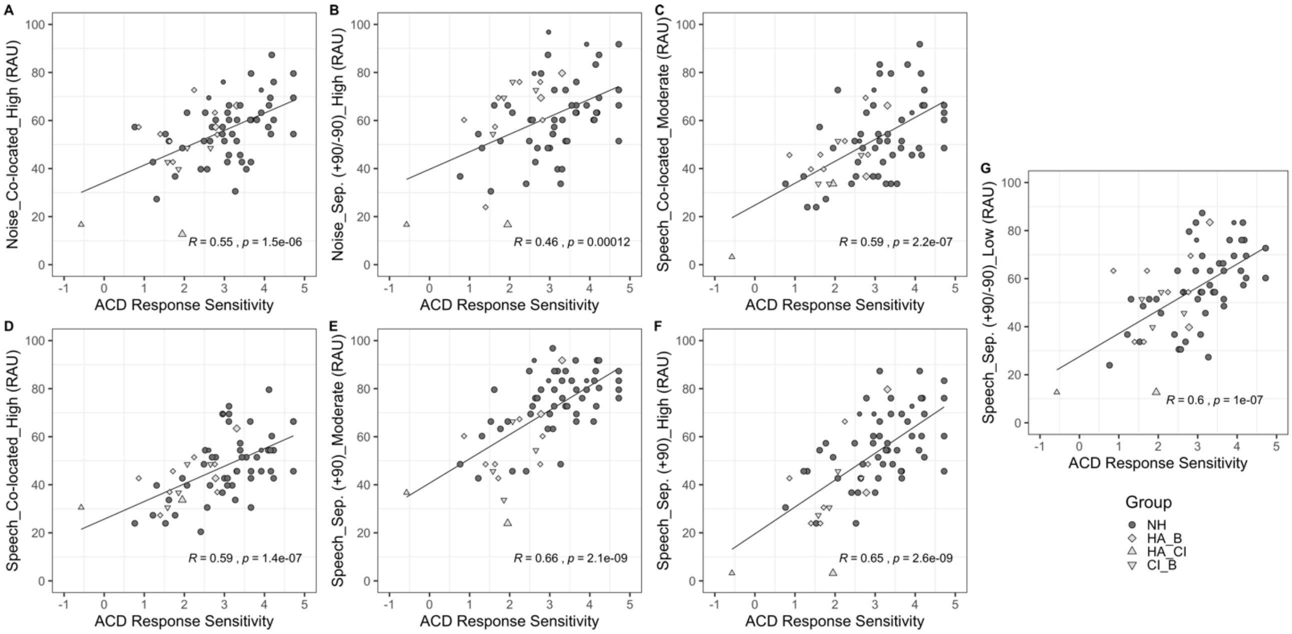


Figure 3.7: Relations between speech recognition and ACD response sensitivity for all children. These acoustic conditions reflect those for which ACD response sensitivity remained a significant predictor of speech recognition despite the addition of age and aided PTA to the model. Smaller points indicate children who completed the speech recognition task at +5 dB SNR.

3.5 DISCUSSION

The primary objective of the present study was to test the hypothesis that children's ability to selectively attend to a target speech stream contributes to their ability to understand speech in complex acoustic environments. Additionally, this study aimed to investigate whether the extent to which selective attention related to speech recognition differed based on the acoustic characteristics of the environment. To test these relations, children who were 5 to 12 years of age with NH and HL performed a speech recognition task under various acoustic conditions that differed based on reverberation time (i.e., low; moderate; high), the presence and type of a competing masker (i.e., quiet; speech-shaped noise; two-talker speech), and the spatial location of the masker, if present (i.e., co-located; separated [+90 degrees]; separated [+90/-90 degrees]). Children's performance on the speech recognition task was quantified as the proportion of key words correctly repeated from sentences presented in each acoustic condition. The separate and interacting effects of reverberation time, masker type, and spatial location as well as children's age and hearing status on speech recognition were considered. To investigate the extent to which individual differences in selective attention contributed to speech recognition, the relations between children's ability to selectively attend to a target speech stream as measured by ACD response sensitivity (*Chapter 2*) and speech recognition across acoustic conditions were assessed.

As expected, reverberation time, masker type, and the spatial location of the masker affected speech recognition in children. Children's age and hearing status were also shown to influence speech recognition as well as modulate the extent to which these acoustic conditions affected speech recognition. Additionally, children's ability to selectively attend to a target speech stream was predictive of their speech recognition regardless of reverberation time, masker type, and the spatial location of the masker. Additionally, as predicted, the extent of these relations differed across acoustic conditions. In a subset of acoustic conditions, ACD response sensitivity was uniquely predictive of speech recognition after accounting for children's age and aided hearing sensitivity. Together, these findings suggest that children's ability to selectively attend to a target speech stream relates to their ability to understand speech, and that the extent of this relation differs based on the acoustic characteristics of the environment. Furthermore, children's ability to *hear* speech is not necessarily predictive of their ability to *understand* speech in complex acoustic environments; therefore, individual differences in children's ability to selectively attend to a target speech stream are important to consider.

3.5.1 Factors that Influence Speech Recognition

Results revealed that child-specific factors (i.e., age and hearing status) as well as acoustic factors (i.e., reverberation time, masker type, and spatial location) significantly affected children's speech recognition. Additionally, significant interactions were observed between these factors, suggesting that children's ability to understand speech is dependent upon dynamic relations between the age and hearing status of the child as well as the acoustic characteristics of the environment. Of most relevance to the present study is the consideration of these factors and the interactions among them within the context of the expected range of attentional demands imposed by different acoustic conditions as well as children's ability to selectively attend to the target speech stream consistent with these demands.

Reverberation Time, Masker Type, and Spatial Location

Reverberation time, masker type, and the spatial location of the masker were manipulated in the present study as their separate and combined effects are known to affect speech recognition, especially in children (e.g., Johnstone & Litovsky, 2006; Leibold & Buss, 2013; Neuman et al., 2010; Yacullo & Hawkins, 1987). Additionally, these acoustic conditions were chosen in an attempt to modulate the extent to which children's ability to understand speech depended on the selective allocation of attentional resources to aid in segregating, tracking, and extracting information from the target speech stream. A final motivation for including these acoustic conditions was that it provided a more comprehensive representation of the variety of acoustic environments within which children must understand speech on a daily basis.

Overall, the observed main effects of reverberation time, masker type, and the spatial location of the masker on speech recognition aligned with the expected effects based on previous literature. Specifically, children's ability to understand speech significantly declined as reverberation time increased both in quiet and in the presence of a competing masker. These findings align with those of previous studies, which have shown that children are poorer at understanding speech as reverberation time increases, especially in the presence of concomitant background noise (Neuman et al., 2010; Neuman & Hochberg, 1983; Yacullo & Hawkins, 1987). Neuman et al. (2010) demonstrated that children between 6 to 12 years of age required more favorable SNRs to understand speech in the presence of a four-talker speech masker as reverberation time increased from low (i.e., T60 = 0.3 seconds) to moderate (i.e., T60 = 0.6 seconds) and high (i.e., T60 = 0.8 seconds). In both Neuman et al. (2010) and the present study, the effects of increasing reverberation time were fairly consistent across the tested age range. Together, these findings suggest that children are at greater risk for poor speech recognition in the presence of reverberation, and that even reverberation times that fall within the recommended ANSI standards (i.e., ≤ 0.6 seconds) disrupt speech recognition in children.

In addition, children demonstrated significantly poorer speech recognition in the presence of a two-talker masker than a speech-shaped noise masker, though the extent of this difference changed based on children's age and hearing status, as discussed further in the subsequent section. This finding is consistent with previous literature, which has demonstrated that maskers that impose informational masking in addition to energetic masking by placing greater demands on top-down processes, such as attention, are more detrimental to speech recognition than maskers that contain primarily energetic masking (e.g., Kidd et al., 2008; Pollack, 1975).

Lastly, consistent with previous research, children were able to benefit from the spatial separation of the target speech and masker when the masker was located at +90 degrees compared to the co-located condition (Johnstone & Litovsky, 2006; Litovsky, 2005). While children benefitted from the spatial separation of the target speech and masker at +90 degrees, the presence of two decorrelated single-stream maskers at +90/-90 degrees resulted in a small but significant reduction in speech recognition (-3 RAU on average) compared to the co-located condition, especially when the masker consisted of two-talker speech (*Appendix A.3.2*). The observation that children's speech recognition was poorer in the separated (+90/-90) conditions than the co-located conditions was inconsistent with the predicted outcomes, as the spatial separation of target and masker streams is typically advantageous for speech recognition in children (Johnstone & Litovsky, 2006; Litovsky, 2005; Litovsky, 2012). Additionally, previous research has shown that children and adults with NH experience spatial release from masking with speech maskers positioned at +90/-90 degrees relative to target speech due to an increase in the opportunities for momentary glimpsing of the target speech at each ear (Bronkhorst, 2000; Jones & Litovsky, 2011; Marrone et al., 2008b; Misurelli & Litovsky, 2012). In considering that the content and level of the maskers were consistent across the co-located, separated (+90), and separated (+90/-90) conditions (i.e., two female talkers at +90 or one female talker each at +90/-90), one possible reason for the reduction in speech recognition observed in the latter condition is the addition of a perceptually distinct second masker stream that competed with the target speech stream for attentional resources in children. Additionally, the content of the two-talker speech masker used in the present study (i.e., excerpts from child-directed science articles) may have made the masker streams more attentionally engaging than the two-talker speech maskers used in the above studies, which consisted of concatenated sentences from standardized corpora (i.e., Coordinate Response Measure [Bolia et al., 2000]; Harvard IEEE [Rothauser, 1969]). However, as reverberation time increased, the greater "smearing" of the energy within each masker stream may have become less attentionally engaging but additively more diffuse, imposing greater energetic masking and resulting in a less favorable listening condition than the co-located or separated (+90) conditions.

Age and Hearing Status

Regarding the effect of age on speech recognition, younger children demonstrated greater difficulty understanding speech than older children across acoustic conditions. Overall, age-related improvements in speech recognition were consistent for children with NH and HL regardless of reverberation time and the spatial location of the masker. However, age-related differences in the effect of masker type on speech recognition were observed. Specifically, younger children experienced greater detriments in speech recognition in the presence of a two-talker speech masker than a speech-shaped noise masker compared to older children. These findings align with those of previous studies, which have demonstrated that younger children are more susceptible to the deleterious effects of informational masking than older children and adults – an expected consequence of their cognitive and linguistic immaturity (e.g., Leibold & Buss, 2013; Wightman & Kistler, 2005). For children with NH in the present study, the differences in speech recognition between conditions containing a speech-shaped noise masker and two-talker speech masker converged with increasing age. An age-by-masker type interaction was not observed for children with HL, which may be due to the variability in speech recognition across ages and masker types observed in this limited sample. However, both groups of children demonstrated robust age-related improvements in speech recognition regardless of masker type. This finding is consistent with those of previous studies demonstrating that younger children are more prone to distractibility due to the presence of a competing masker or task-irrelevant sound than older children (Elliott, 2002; Meinhardt-Injac et al., 2015).

In addition to the observed effects of age, children's hearing status significantly affected their ability to understand speech both in quiet and in the presence of a competing masker. Specifically, children with HL demonstrated poorer speech recognition than children with NH, overall. However, the extent of this difference changed based on the specific acoustic conditions under which speech recognition was measured as demonstrated by the presence of a significant four-way interaction between group, reverberation time, masker type, and spatial location. Post hoc Bonferroni-corrected paired comparisons revealed that the most substantial differences in speech recognition between children with NH and HL were observed in the acoustic conditions that contained a spatially-separated masker (*Appendix A.3.2*; **Figure 3.4)**. These findings may reflect group differences in children's ability to selectively attend to a target speech stream while inhibiting attention to spatially separated masker streams. However, these findings may also reflect differences in access to monaural and binaural spatial cues between children with NH and HL. Specifically, children with HL may have received distorted interaural timing and level difference cues due to their use of hearing aids and cochlear implants, which may have limited their ability to benefit from the spatial separation of the target speech from the masker. Consistent with this, previous studies have shown variability in spatial release from masking in children who use hearing aids and cochlear implants, with some children with HL showing no benefit of spatial separation for speech recognition (Ching et al., 2011; Litovsky et al., 2006; Misurelli & Litovsky, 2015). In the present study, the speech recognition of children with HL mirrored the patterns of children with NH across spatial locations such that significant benefits of spatial separation were observed. However, these benefits were reduced for children with HL compared to children with NH, on average.

While multiple bottom-up and top-down processes are expected to have contributed to the observed differences in children's performance across age, groups, and acoustic conditions, the primary objective of the present study was to assess the role of selective attention in speech recognition. Specifically, a potential factor contributing to the observed age- and hearing status-related differences in speech recognition across acoustic conditions is that younger children and children with HL had greater difficulty selectively attending to the target speech stream and inhibiting attention to competing maskers. The study described in *Chapter 2* provided evidence of differences in children's ability to selectively attend to a target speech stream based on their age and hearing status. If children's ability to selectively attend to a target speech stream contributes to their ability to understand speech in complex acoustic environments, children who have poorer selectively attention should concomitantly demonstrate poorer speech recognition. The present study directly tested this relation across acoustic conditions, as discussed below.

3.5.2 Relation between Selective Attention and Speech Recognition

Consistent with the primary hypothesis of this study, children's ability to selectively attend to a target speech stream was found to significantly relate to their speech recognition across acoustic conditions. Furthermore, the proportion of variance in speech recognition accounted for by selective attention differed across acoustic conditions. A possible explanation is that children's ability (or inability) to selectively attend to a target speech stream was more predictive of speech recognition in acoustic conditions that imposed greater attentional demands. On one hand, this may reflect acoustic conditions during which it is more difficult to segregate and track target and masker streams, such as those that involve longer reverberation times or a co-located target and masker. This interpretation is consistent with the ELU model of speech understanding, which suggests that dependence on top-down processes, such as attention, is greater when needed to aid in resolving degraded auditory input (Rönnberg et al., 2013). In addition, acoustic conditions that impose greater informational masking or more opportunities for distractibility by the competing masker, such as when the masker contains meaningful speech content or when there is more than one masker stream to ignore, may result in greater attentional demands for speech recognition.

In the present study, the extent to which ACD response sensitivity accounted for variance in speech recognition differed across conditions, which may reflect how the combined effects of reverberation time, masker type, and the spatial location of the masker interacted to alter attentional demands. Notably, the proportion of variance in speech recognition accounted for by ACD response sensitivity (**Figure 3.6**) appeared to inversely reflect the observed patterns of speech recognition across acoustic conditions (**Figure 3.5**). As described in *Results, Section 3.4.2*, this may indicate that children's ability to selectively attend to a target speech stream may be more predictive of speech recognition under acoustic conditions that impose greater attentional demands.

For instance, the proportion of variance in speech recognition accounted for by ACD response sensitivity tended to be greater for the acoustic conditions containing a two-talker speech masker as compared to a speech-shaped noise masker. Furthermore, in five of the acoustic conditions containing a two-talker speech masker, the significant relation between ACD response sensitivity and speech recognition remained despite the inclusion of age and aided PTA in the model. This is consistent with the observation that children, especially children with HL, had greater difficulty understanding speech amidst a two-talker masker than a speech-shaped noise masker. Together, these findings suggest that children's ability to selectively attend to a target speech stream contributes to their speech recognition to a greater extent in environments that contain an informational masker, perhaps due to the increased attentional demands associated with inhibiting attention to the masker. In line with this, results from previous studies suggest that selective attention is especially important for speech recognition in the presence of background noise that contains informational masking (Corbin et al., 2016; Leibold & Buss, 2013; Wightman & Kistler, 2005; see Leibold, 2017 for a review).

Another observed trend was that, in conditions that contained low reverberation, ACD response sensitivity appeared to account for the most variance in speech recognition across masker types when decorrelated single-stream maskers were presented from +90/-90 degrees. This observation may reflect the greater attentional demands associated with children's ability to selectively attend to the target speech stream while inhibiting attention to two competing masker streams that were perceptually distinct. This competition for attentional resources and individual differences in children's ability to selectively attend to the target speech stream are consistent with the observation of poorer speech recognition in the separated (+90/-90) conditions, especially in the presence of a two-talker masker. These findings align with those of previous studies, which have shown that children have greater difficulty understanding speech when they need to perceptually isolate the target speech from more than one competing masker (Bronkhorst & Plomp, 1992; Buss et al., 2017; Carhart et al., 1975; Litovsky, 2005). While these effects have typically been observed between maskers containing one and two talkers, it is possible that children may have perceived the masker as a single stream when co-located or presented at +90 degrees and two distinct streams when spatially separated at +90/-90 degrees due to the addition of a robust spatial cue (i.e., 180 degrees) between masker streams in this condition. If this was the case, it is sensible that children's ability to selectively attend to the target speech stream was more predictive of speech recognition in the conditions containing two separate masker streams to perceptually isolate and ignore (i.e., at +90/-90) as opposed to a single masker stream (i.e., at 0 degrees or +90 degrees) due to increased attentional demands.

Lastly, the proportion of variance in speech recognition accounted for by ACD response sensitivity in the conditions containing a two-talker speech masker separated at +90 degrees appeared to be greater for the moderate and high reverberation conditions than the low reverberation condition. This observation may suggest that children needed to expend additional attentional resources to benefit from the spatial separation of the masker relative to the target speech as reverberation time increased, and that children who had greater difficulty selectively attending to the target stream experienced more substantial reductions in speech recognition. This is consistent with the finding that children demonstrated less spatial release from masking as reverberation time increased from low to high. Specifically, smaller differences in speech recognition scores were observed between the co-located and separated (+90) conditions in the high reverberation conditions relative to the low reverberation conditions (*Appendix A.3.2*). Previous studies in adults investigating the interactions between reverberation time and spatial separation have produced mixed results. For instance, a study by Marrone et al., 2008a demonstrated that reverberation adversely affected the ability of adults with NH and HL to benefit from spatial separation between the target speech and masker for speech recognition. Conversely, Kidd et al., 2005 revealed no difference in spatial release from masking with increased reverberation for an informational masker; however, this lack of a difference was attributed to an approximately equal increase in speech recognition thresholds for the co-located and spatially-separated conditions in the presence of higher reverberation times.

While the relation between ACD response sensitivity and speech recognition was observed across acoustic conditions, examining how these relations change after including age and aided hearing sensitivity (i.e., aided PTA) in the model provides useful insight regarding the extent to which these factors accounted for overlapping variance in speech recognition. For the majority of acoustic conditions (i.e., 14 out of 21), age, aided PTA, and ACD response sensitivity accounted for similar variance in children's speech recognition. In other words, ACD response sensitivity was not uniquely predictive of speech recognition after accounting for age and aided PTA. These findings suggest that younger children and children with HL – the children who had greater difficulty selectively attending to a target speech stream in *Chapter 2* – were also the children who demonstrated poorer speech recognition in these acoustic conditions. This included all acoustic conditions containing reverberation but no masker (i.e., quiet) as well as the majority of conditions containing a speech-shaped noise masker (i.e., 7 out of 9). ACD response sensitivity was found to account for significant variance in speech recognition over and above age and aided PTA in the remaining seven acoustic conditions, the majority of which (i.e., 5 out of 7) contained a two-talker speech masker. In these conditions, individual differences in children's ability to selectively attend to a target speech stream that could not be explained by age and aided PTA were significantly predictive of their speech recognition (**Figure 3.7**). These findings complement those of previous studies, which have demonstrated that differences in children's cognitive and linguistic processing contribute to their ability to understand speech in complex acoustic environments after accounting for factors such as age and hearing sensitivity (e.g., Blamey et al., 2001; Klein et al., 2017; McCreery et al., 2019; Thompson et al., 2019).

Taken together, the findings from the present study provide additional empirical evidence regarding the role of selective attention in speech recognition. In summary, children's ability to selectively attend to a target speech stream related to their speech recognition, especially under acoustic conditions that may have imposed greater attentional demands. Additionally, in a subset of acoustic conditions, children's ability to selectively attend to a target speech stream accounted for unique variance in speech recognition over and above age and aided hearing sensitivity.

3.5.3 Limitations

While the findings from the present study provide novel insight regarding the relation between selective attention and children's ability to understand speech in complex acoustic environments, there are a few limitations to consider. Firstly, the speech recognition task involved children listening to an isolated sentence, which is not necessarily reflective of how communication unfolds in real-world environments. Typically, children are able to capitalize on concomitant information, such as that which arises from semantic context and visual speech cues, to inform their speech recognition. Requiring children to depend solely on the acoustic characteristics of the target speech signal provided a more well-controlled measure of the influence of reverberation time, masker type, the spatial location of the masker, and selective attention on speech recognition. However, the speech recognition performance observed here may differ from that observed in real-world environments under similar acoustic conditions, though the general relations are expected to persist.

Additionally, although the present study focused on selective attention as a top-down process involved in speech recognition, it is fully expected that other cognitive and linguistic factors, such as working memory and language ability, contributed to children's performance on the speech recognition task. The role of such processes in speech recognition is well documented in the scientific literature, and individual differences in these processes likely contributed to children's performance on the speech recognition task. The inclusion of age and aided PTA in Model 2 of the hierarchical linear regression analyses may have accounted for some of the individual differences in these processes, which could have contributed to the significant differences observed between Model 1 and Model 2 for a subset of the acoustic conditions. While the cognitive and linguistic processes expected to contribute to children's ability to understand speech in complex acoustic environments are often considered individually, the underlying interrelations among these factors must be acknowledged.

Lastly, as discussed in *Chapter 2, Section 2.5.5 Limitations*, an inherent limitation of the present study was the size and heterogeneity of the group of children with HL. While group differences were assessed during the statistical analyses when possible, the difference in statistical power between the groups precluded the examination of specific relations for children with NH and HL separately. In considering the heterogeneity within the group of children with HL, there are various factors that may have influenced their performance during the speech recognition task. Specifically, the range of the degrees of hearing loss, hearing devices used, and duration of device usage in the group of children with HL likely contributed to the observed variability in their performance on the speech recognition task across acoustic conditions. In addition, the extent of the attentional demands imposed by various acoustic conditions may have been modulated by the specific programmatic settings of children's hearing aids or cochlear implants. For instance, children whose devices provided greater directionality may have had less difficulty inhibiting attention to spatially separated maskers due to the reduction in gain of this peripheral input. Individual differences in aided hearing sensitivity were also observed within the group of children with HL, and were statistically accounted for during the hierarchical linear regression analyses by the inclusion of aided PTA in Model 2.

3.5.4 Implications and Future Directions

The findings from the present study extend those of previous studies by investigating how children's ability to selectively attend to a target speech stream influences speech recognition under acoustic conditions that more closely reflect those of real-world environments. As such, the implications of the present study include a more in-depth understanding of how individual differences in selective attention may contribute to observed variability in children's speech recognition. Specifically, understanding speech in complex acoustic environments is attentionally demanding, and children who have greater difficulty meeting these demands are expected to be less able to extract, encode, and process target speech amidst competing input. As attention is considered to be a domain-general cognitive process, it is possible that any source of auditory or visual competition may result in the allocation of attentional resources away from the target speech, especially in children who have poorer selective attention. Within the context of a classroom, for instance, these findings suggest that minimizing distractions around the room is paramount for children's speech recognition. However, additional research is needed to further elucidate the impact of auditory and visual distractors on children's ability to selectively attend to a target speech stream.

Furthermore, while the findings from the present study provide evidence of a relation between children's selective attention and their speech recognition in complex acoustic environments, selective attention is a multifaceted and dynamic cognitive process that is not expected to be fully characterized by a single task. ACD response sensitivity reflects children's ability to selectively attend to a target speech stream and inhibit attention to a competing speech stream, which parallels the attentional demands of understanding speech in complex acoustic environments. This finding is of clinical relevance as it underscores the importance of training the development of “listening skills”, such as directing attention toward the person talking, as part of aural habilitation programs. However, it is possible that training selective attention more generally may not have concomitant benefit for speech recognition. Future research should assess the most valid method of quantifying selective attention as it relates to speech recognition in order to provide an outcomes-based metric for interventions targeting selective attention.

Lastly, in considering the size and heterogeneity of the group of children with HL included in the present study, as described in *Section 3.5.3 Limitations*, future studies should aim to replicate the findings described here with a larger cohort of children with HL. In addition to providing additional evidence regarding the influence of selective attention on speech recognition in children with HL, considering how factors related to children’s auditory experience influence the relation between selective attention and speech recognition would provide additional insight into the mechanisms underlying these processes.

3.6 CONCLUSION

Consistent with previous studies, children’s age and hearing status affected their ability to understand speech in acoustic conditions that differed based on reverberation time, masker type, and the spatial location of the masker. Furthermore, as hypothesized, children’s ability to selectively attend to a target speech stream related to their speech recognition across acoustic conditions, though the amount of variance in speech recognition accounted for by selective attention varied. This observation may reflect differences in the attentional demands associated with selectively attending to the target speech stream and inhibiting attention to competing input across acoustic conditions. Together, the findings from the present study contribute to the growing body of research aimed at understanding the cognitive and linguistic processes that contribute to children's ability to understand speech in complex acoustic environments.

CHAPTER 4 | CONCLUSION

The primary objectives of this dissertation were twofold: 1) to determine how age and hearing loss alter selective attention during childhood; and 2) to quantify the extent to which individual differences in selective attention account for the observed variability in children's speech recognition. A multi-experiment within-subjects design was utilized to investigate these processes in children between 5 and 12 years of age with normal hearing and hearing loss.

4.1 SUMMARY OF FINDINGS

The findings from this dissertation support the central hypothesis that poor selective attention during childhood, due to immaturity and disrupted auditory experience, contributes to the difficulties younger children and children with hearing loss have understanding speech in complex acoustic environments. Specifically, *Chapter 2* described how the ability to selectively attend to a target auditory or visual stream is influenced by age and hearing status during childhood, while *Chapter 3* discussed the role of selective attention in children's ability to understand speech in complex acoustic environments. The key findings described in these chapters along with their scientific contributions, implications for speech recognition in real-world environments, and applications to clinical practice are discussed below.

4.1.1 Selective Attention in Children

The results of the study described in *Chapter 2* support the hypothesis that immaturity and disrupted auditory experience impede children's ability to selectively attend to a target stream and inhibit attention to competing input. Children performed behavioral change detection tasks in the auditory and visual domains during which they were instructed to attend and respond to deviant stimuli in a designated target stream and inhibit attention to a distractor stream. As predicted, younger children and children with hearing loss responded less frequently to deviants in the target stream and more frequently to deviants in the distractor stream than older children and children with normal hearing.

Notably, the differences in performance related to age and hearing status were observed across the auditory *and* visual domains. For children with hearing loss, these findings suggest that disrupted auditory experience during childhood has implications for selective attention beyond the sensory domain of impairment. Consistent with this observation, children's hearing sensitivity with the use of their clinical hearing devices (i.e., aided) and without the use of these devices (i.e., unaided) was found to relate to their performance on the auditory and visual change detection tasks. These relations changed as a function of age such that younger children with hearing loss demonstrated greater difficulty selectively attending to a target auditory or visual stream than older children with similar degrees of hearing loss. Together, these findings indicate that children's ability to selectively attend to a target stream and inhibit attention to competing input is influenced by their age and hearing status, and that improvements in executive function later in childhood may compensate for these differences.

4.1.2 The Role of Selective Attention in Speech Recognition

The results from the study described in *Chapter 3* support the hypothesis that children's ability to selectively attend to a target speech stream and inhibit attention to competing auditory input contributes to their speech recognition. Children performed a speech recognition task with sentences in various acoustic conditions comprised of different reverberation times (i.e., low; moderate; high), masker types (i.e., no masker; speech-shaped noise; two-talker speech), and spatial locations of the masker (i.e., separated [+90]; separated [+90/-90]; co-located). Children's speech recognition was best in the presence of low reverberation with no masker and poorest in the presence of high reverberation with a two-talker speech masker separated at +90/-90 degrees azimuth. Similar to the results of the study described in *Chapter 2*, the extent to which these acoustic conditions affected speech recognition was modulated by the age and hearing status of the child. Overall, younger children and children with hearing loss demonstrated poorer speech recognition than older children and children with normal hearing.

Of particular interest was the observation that children's ability to selectively attend to a target speech stream, as quantified in *Chapter 2*, was significantly predictive of their speech recognition across acoustic conditions. However, the extent to which selective attention contributed to speech recognition appeared to differ based on the acoustic characteristics of the environment. For instance, selective attention accounted for greater variance in speech recognition when the masker consisted of two-talker speech as compared to speech-shaped noise. These findings are thought to partially reflect differences in the attentional demands imposed by the acoustic conditions under which children performed the speech recognition task. Additionally, as expected, age and aided hearing sensitivity accounted for similar variance in speech recognition as selective attention in a majority of acoustic conditions, though children's ability to selectively attend to a target speech stream was uniquely predictive of speech recognition in a subset of acoustic conditions containing a masker. In conjunction with the findings from *Chapter 2*, these results suggest that individual differences in children's ability to selectively attend to a target speech stream significantly influence their speech recognition in complex acoustic environments.

4.2 SCIENTIFIC CONTRIBUTIONS

The primary scientific contributions of the research described in this dissertation include an enhanced understanding of the influence of age and hearing loss on selective attention during childhood (*Chapter 2*) and novel insight regarding the contribution of selective attention to speech recognition in children (*Chapter 3*). In considering the interdisciplinary nature of this work, these advancements in knowledge may be of relevance to multiple fields and academic disciplines, including hearing science, cognitive hearing science, psychology, education, and pediatric audiology.

4.2.1 Effects of Age and Hearing Loss on Selective Attention During Childhood

A rich body of literature has documented improvements in selective attention with increasing age during childhood using a variety of behavioral, electrophysiologic, and imaging-based methodologies. While these age-related changes in selective attention have been observed across the auditory and visual domains, few studies have examined selective attention using analogous tasks across domains within the same cohort of children. The findings from *Chapter 2* provide converging evidence of age-related improvements in children's ability to selectively attend to a target stream using novel behavioral change detection tasks across the auditory and visual domains. In addition, the methodological design of these tasks enabled the observation of disparate patterns of responses between younger and older children. Specifically, while older children selectively allocated attention to the target stream, as instructed, younger children deployed attention more diffusely across the target and distractor streams. These findings corroborate those of previous studies by providing additional documentation of differences in children's selective attention across the elementary school age range.

Additionally, previous studies have demonstrated differences in selective attention between children with normal hearing and children with hearing loss; however, much of this evidence has stemmed from tasks in the visual domain (Dye & Bavelier, 2010; Dye & Hauser, 2014; Quittner et al., 1994). In considering the prevalence of auditory-domain tasks requiring selective attention – such as speech recognition – in everyday life, it is important to additionally consider the effects of hearing loss on children's selective attention under conditions that reflect these real-world demands. While only a few studies to date have examined the effects of hearing loss on selective attention in the auditory domain (Holmes et al., 2017; McCreery et al., 2019), the findings from *Chapter 2* contribute to this current state of knowledge by revealing differences in the abilities of children with normal hearing and children with hearing loss to selectively attend to a target speech stream within the context of a behavioral change detection task. A considerable challenge of quantifying selective attention in the auditory domain in children with hearing loss, however, is ruling out the possible influence of impaired access to auditory stimuli on performance. The study described in *Chapter 2* addressed this potential confound by having the same cohorts of children complete an analogous change detection task in the visual domain. Together, these findings provide novel insight regarding the effects of hearing loss on selective attention across the auditory and visual domains during childhood. While the results from this study may not reflect global differences in selective attention between children with normal hearing and children with hearing loss, per se, they demonstrate for the first time that children's ability to selectively attend to a specific auditory or visual input in line with behavioral goals is influenced by their hearing status. Furthermore, they provide preliminary evidence that the continued development of executive functions into adolescence may partially compensate for the observed deleterious effects of hearing loss on selective attention during childhood.

4.2.2 Contribution of Selective Attention to Speech Recognition in Children

The role of selective attention in analyzing complex auditory scenes – a prerequisite for speech recognition – is well-established from a theoretical perspective. Specifically, attention is a capacity-limited resource, which impedes a listener's ability to effectively process multiple streams of auditory input at any given time. Therefore, auditory input is subjected to an attentional "filter" that determines the components of the acoustic signal to be engaged in higher-level processing (Bregman, 1990; Broadbent, 1958; Kahneman, 1973). Empirically, previous studies have utilized dichotic listening paradigms during which separate auditory streams are presented to contralateral ears via headphones to assess the relations between selective attention and speech recognition. Results have demonstrated that a listener's ability to recall the content and features of an auditory stream presented to either ear is dependent upon their selective allocation of attention toward that stream and away from the competing stream – an ability that has been shown to be poorer in children (Cherry, 1953; Wightman & Kistler, 2005). While these studies provide valuable insight regarding how attention alters the perception of simultaneous streams of auditory input, the isolated presentation of individual streams to each ear represents a favorable listening scenario that underestimates the attentional demands of real-world speech recognition. Thus, the generalizability of these findings to speech recognition in complex acoustic environments, such as a classroom, within which children must detect, segregate, track, and selectively attend to target speech amidst competing auditory input may be limited. The results from *Chapter 3* of this dissertation addressed these limitations by elucidating the relation between selective attention and speech recognition under acoustic conditions that more closely mimic those of real-world complex acoustic environments.

Specifically, a particularly unique contribution of the study described in *Chapter 3* was the variety of acoustic conditions under which speech recognition was quantified for all children. Real-world acoustic environments are heterogeneous, yet studies aimed at investigating speech recognition and its underlying processes most commonly adhere to a single acoustic condition or a limited set of acoustic conditions. While this approach has practical benefits, the dynamicity and multiplicity of real-world acoustic environments limit the generalizability of the findings from such studies. This is an especially important consideration when investigating processes such as selective attention, as the dependence on this process for speech recognition is modulated by the attentional demands imposed by the acoustic characteristics of the environment.

4.3 IMPLICATIONS

4.3.1 Speech Recognition in Complex Acoustic Environments

In considering the ubiquity of background noise and reverberation in everyday environments, an in-depth understanding of the top-down processes that contribute to speech recognition in children is necessary in order to bolster their academic success and social well-being. While other cognitive and linguistic factors, such as working memory and language ability, are known to contribute to speech recognition in children, only a few studies have directly tested the relation between selective attention and children's ability to understand speech in complex acoustic environments.

Primarily, the findings from this dissertation support the notion that access to the acoustic characteristics of speech is necessary, but not sufficient, for speech recognition, and that children's ability to selectively attend to the target speech stream is a significant contributing factor. This motivates the need to consider, in any given environment, what actions can be taken to facilitate children's access to the target speech as well as limit opportunities for distractibility from other sources of input. For instance, even if children are preferentially seated in a classroom such that they are within close proximity to the teacher, it is possible that other sources of competing input, such as an HVAC system or other children talking, may reduce speech recognition by competing for attentional resources even if not impeding children's ability to *hear* the teacher's voice. While an important consideration for all children, the results presented here suggest that younger children and children with hearing loss are especially at-risk for poor speech recognition in such scenarios due to a reduced ability to selectively attend to the target speech and inhibit attention to competing input.

A second implication from this dissertation is that the attentional demands associated with understanding speech are expected to differ based on the acoustic characteristics of the environment. Therefore, children who have poorer selective attention may be more or less able to recognize speech based on the reverberation time as well as the presence, content, and location of background noise in the environment. For instance, in considering the example described above, children may have less difficulty attending to the teacher's voice in the presence of noise from the HVAC system and greater difficulty doing so when other children are talking due to the greater attentional demands associated with inhibiting attention to background noise that contains speech and imposes informational masking.

Lastly, the results from this dissertation suggest that, although age and hearing status influence children's ability to selectively attend to a target speech stream, individual differences in selective attention that are not accounted for by these factors may also contribute to children's speech recognition in complex acoustic environments. While younger children and children with hearing loss are at greatest risk for deficits in speech recognition related to selective attention, it is necessary to consider these abilities on a child-by-child basis when addressing real-world functional communication concerns.

4.3.2 Considerations for Clinical Practice

In addition to the implications of these findings for children's ability to understand speech in complex acoustic environments, there are applications to clinical practice to consider. Firstly, the observation of age- and hearing-status related differences in children's ability to selectively attend to a target speech stream reinforces the importance of incorporating specific training of "listening skills" into aural habilitation programs for children with hearing loss. Highlighting the importance of directing attention toward the person speaking – both auditorily and visually – is expected to facilitate children's ability to understand speech, especially in complex acoustic environments. Furthermore, educating parents and caregivers about the importance of selective attention for speech recognition may empower them to foster the development of these skills in their children.

Additionally, these findings underscore the need to include tests of functional auditory abilities, such as selectively attending to target speech, into the pediatric audiologic test battery. Standard clinical assessments include the detection of non-speech stimuli (e.g., tones, noise bursts) as well as the recognition of speech stimuli (e.g., words) in quiet. While these tests provide useful information regarding hearing sensitivity, children's performance on these measures is oftentimes not commensurate with their ability to understand speech in real-world environments. A potential reason for this is a reduced ability to selectively attend to a target speech stream in complex acoustic environments that impose greater attentional demands. Therefore, incorporating measures that impose greater attentional demands to further probe selective attention into the current test battery, such as a word or sentence recognition task in the presence of a spatially separated two-talker speech masker, may provide additional information to inform clinical recommendations and treatment plans for children with listening difficulties.

Finally, while the assessment and diagnosis of central auditory processing disorders is a well-established subspecialty within the field of audiology, the results of this dissertation support the need to recognize subclinical differences in auditory processing abilities, such as selective attention, that do not meet these diagnostic criteria. This consideration is especially relevant for young children and children with hearing loss whose listening difficulties may be able to be managed by providing additional training and support to bolster selective attention, which has positive downstream effects for speech recognition.

4.4 FINAL CONCLUSION

Together, the findings from this dissertation contribute to the growing body of research aimed at elucidating the top-down processes that contribute to children's ability to understand speech in complex acoustic environments. By providing innovative evidence that age- and hearing status-related differences in selective attention account for observed variability in children's speech recognition, this work has the potential to inform clinical practice and classroom-based interventions. Finally, these results provide a foundation for future research to investigate the specific mechanisms underlying the effects of age and hearing loss on the development of selective attention during childhood in order to better understand the neurobiological and experiential factors that ultimately contribute to children's speech recognition in complex acoustic environments.

REFERENCES

Abundis-Gutiérrez, A., Checa, P., Castellanos, C., & Rueda, M. R. (2014). Electrophysiological correlates of attention networks in childhood and early adulthood. *Neuropsychologia*, *57*, 78-92.

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*(6), 716-723.

Alain, C., & Arnott, S. R. (2000). Selectively attending to auditory objects. *Frontiers in Bioscience*, *5*(10), D202-D212.

Alain, C., Arnott, S. R., & Picton, T. W. (2001). Bottom-up and top-down influences on auditory scene analysis: evidence from event-related brain potentials. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(5), 1072–1089.

Alain, C., & Bernstein, L. J. (2008). From sounds to meaning: the role of attention during auditory scene analysis. *Current Opinion in Otolaryngology & Head and Neck Surgery, 16*(5), 485-489.

Anderson, K. (2004). The problems of classroom acoustics: The typical classroom soundscape is a barrier to learning. *Seminars in Hearing*, *25*(2), 117-129.

ANSI S12.60-2010. (2010). *Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools*.

Arbogast, T. L., Mason, C. R., & Kidd, G. (2002). The effect of spatial separation on informational and energetic masking of speech. *The Journal of the Acoustical Society of America, 112*(5), 2086-2098.

Asbjørnsen, A., & Ma, A. H. (2000). Lasting auditory attention impairment after persistent middle ear infections: a dichotic listening study. *Developmental Medicine & Child Neurology, 42*(7), 481-486.

Assmann, P., & Summerfield, Q. (2004). The perception of speech under adverse conditions. In *Speech Processing in the Auditory System* (pp. 231-308). Springer, New York, NY.

Baddeley, A. (2003). Working memory and language: an overview. *Journal of Communication Disorders*, *36*(3), 189-208.

Baddeley, A. (2010). Working memory. *Current Biology*, *20*(4), 136-140.

Baddeley, A., & Hitch, G. J. (1974). *Working Memory*. In *Psychology of Learning and Motivation* (Vol. 8, pp. 47-89). Academic Press.

Bamford, J., & Bench, R. J. (1979). *Speech-hearing tests and the spoken language of hearing-impaired children*. Academic Press.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1).

Beer, J., Kronenberger, W. G., Castellanos, I., Colson, B. G., Henning, S. C., & Pisoni, D. B. (2014). Executive functioning skills in preschool-age children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1521-1534.

Bench, J., Kowal, A., & Bamford, J. (1979). The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children. *British Journal of Audiology*, *13*(3), 108-112.

Bernstein, J. G. W., & Oxenham, A. J. (2006). The relationship between frequency selectivity and pitch discrimination: Sensorineural hearing loss. *The Journal of the Acoustical Society of America*, *120*(6), 3929-3945.

Bess, F. H., Dodd-Murphy, J., & Parker, R. (1998). Children with minimal sensorineural hearing loss: prevalence, educational performance, and functional status. *Ear and Hearing*, *19*(5), 339-354.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, *44*(2), 264-285.

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *The Journal of the Acoustical Society of America*, *107*(2), 1065-1066.

Booth, J. R., Burman, D. D., Meyer, J. R., Lei, Z., Trommer, B. L., Davenport, N. D., Li, W., Parrish, T. B., Gitelman, D. R., & Mesulam, M. M. (2003). Neural development of selective attention and response inhibition. *Neuroimage*, *20*(2), 737-751.

Boothalingam, S., Purcell, D. W., Allan, C., Allen, P., & Macpherson, E. (2019). Auditory localization and spatial release from masking in children with suspected auditory processing disorder. *Ear and Hearing*, *40*(5), 1187-1196.

Boothroyd, A. (2004). Room Acoustics and Speech Perception. *Seminars in Hearing*, *25*(02), 155-166.

Botting, N., Jones, A., Marshall, C., Denmark, T., Atkinson, J., & Morgan, G. (2017). Nonverbal executive function is mediated by language: A study of deaf and hearing children. *Child Development*, *88*(5), 1689-1700.

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *The Journal of the Acoustical Society of America*, *123*(4), 2078-2086.

Bregman, A. S. (1990). Auditory scene analysis. *International Encyclopedia of the Social Behavioral Sciences*, *Elsevier*, 940-942.

Bregman, A. S. (1993). Auditory scene analysis: Hearing in complex environments. In *Thinking in Sound: The Cognitive Psychology of Human Audition* (pp. 10-36).

Broadbent, D. E. (1958). *Perception and communication*. 1-367.

Brocki, K. C., & Bohlin, G. (2004). Executive functions in children aged 6 to 13: A dimensional and developmental study. *Developmental Neuropsychology*, *26*(2), 571-593.

Bronkhorst, A. W., & Plomp, R. (1992). Effect of multiple speechlike maskers on binaural speech recognition in normal and impaired hearing. *The Journal of the Acoustical Society of America*, *92*(6), 3132-3139.

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*(1), 117-128.

Brooks, B. L., Sherman, E. M. S., & Strauss, E. (2009). NEPSY-II: A developmental neuropsychological assessment. *Child Neuropsychology*, *16*(1), 80-101.

Brungart, D. S. (2005). Informational and energetic masking effects in multitalker speech perception. In *Speech Separation by Humans and Machines* (pp. 261-267). Springer, Boston, MA.

Brungart, D. S., Simpson, B. D., Ericson, M. A., & Scott, K. R. (2001). Informational and energetic masking effects in the perception of multiple simultaneous talkers. *The Journal of the Acoustical Society of America*, *110*(5), 2527-2538.

Buss, E., Leibold, L. J., Porter, H. L., & Grose, J. H. (2017). Speech recognition in one- and two-talker maskers in school-age children and adults: Development of perceptual masking and glimpsing. *The Journal of the Acoustical Society of America*, *141*(4), 2650–2660.

Caldwell, A., & Nittrouer, S. (2013). Speech perception in noise by children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *56*(1), 13-30.

Camos, V., & Barrouillet, P. (2011). Developmental change in working memory strategies: From passive maintenance to active refreshing. *Developmental Psychology*, *47*(3), 898-904.

Camos, V., & Barrouillet, P. (2014). Attentional and non-attentional systems in the maintenance of verbal information in working memory: The executive and phonological loops. *Frontiers in Human Neuroscience*, *8*, 1-11.

Carhart, R., Johnson, C., & Goodman, J. (1975). Perceptual masking of spondees by combinations of talkers. *The Journal of the Acoustical Society of America*, *58*(S1), S35-S35.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 115.

Carlyon, R. P., Plack, C. J., Fantini, D. A., & Cusack, R. (2003). Cross-modal and non-sensory influences on auditory streaming. *Perception*, *32*(11), 1393-1402.

Carroll, J., & Zeng, F. G. (2007). Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations. *Hearing Research*, *231*(1-2), 42-53.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, *25*(5), 975-979.

Chevalier, N., Kurth, S., Doucette, M. R., Wiseheart, M., Deoni, S. C. L., Dean, D. C., O'Muircheartaigh, J., Blackwell, K. A., Munakata, Y., & LeBourgeois, M. K. (2015). Myelination is associated with processing speed in early childhood: Preliminary insights. *PLoS ONE*, *10*(10), 1-14.

Ching, T. Y. C., Van Wanrooy, E., Dillon, H., & Carter, L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *The Journal of the Acoustical Society of America*, *129*(1), 368-375.

Cleary, M., Pisoni, D. B., & Kirk, K. I. (2005). Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *48*(1), 204-223.

Coch, D., Sanders, L. D., & Neville, H. J. (2005). An event-related potential study of selective auditory attention in children and adults. *Journal of Cognitive Neuroscience*, *17*(4), 605-622.

Conway, C. M., Pisoni, D. B., & Kronenberger, W. G. (2009). The importance of sound for cognitive sequencing abilities: The auditory scaffolding hypothesis. *Current Directions in Psychological Science*, *18*(5), 275-279.

Corbin, N. E., Bonino, A. Y., Buss, E., & Leibold, L. J. (2016). Development of open-set word recognition in children: Speech-shaped noise and two-talker speech maskers. *Ear and Hearing*, *37*(1), 55-63.

Cortese, M. J., & Khanna, M. M. (2008). Age of acquisition ratings for 3,000 monosyllabic words. *Behavior Research Methods*, *40*(3), 791-794.

Crandell, C. C. (1993). Speech recognition in noise by children with minimal degrees of sensorineural hearing loss. *Ear and Hearing*, *14*(3), 210–216.

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools*, *31*(4), 362-370.

Crandell, C. C., & Smaldino, J. J. (2001). Classroom acoustics: Understanding barriers to learning. *Volta Review*, *101*(5), n5.

Crukley, J., Scollie, S., & Parsa, V. (2011). An exploration of non-quiet listening at school. *Journal of Educational Audiology*, *17*(1), 23-35.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037-2078.

De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., & Pantelis, C. (2003). Normative data from the Cantab. I: Development of executive function over the lifespan. *Journal of Clinical and Experimental Neuropsychology*, *25*(2), 242-254.

Dockrell, J. E., & Shield, B. M. (2006). Acoustical barriers in classrooms: The impact of noise on performance in the classroom. *British Educational Research Journal*, *32*(3), 509-525.

Doelle, L. L. (1972). *Environmental acoustics*.

Doyle, A. B. (1973). Listening to distraction: A developmental study of selective attention. *Journal of Experimental Child Psychology*, *15*(1), 100-115.

Dye, M. W. G., & Bavelier, D. (2010). Attentional enhancements and deficits in deaf populations: An integrative review. *Restorative Neurology and Neuroscience*, *28*(2), 181-192.

Dye, M. W. G., & Hauser, P. C. (2014). Sustained attention, selective attention and cognitive control in deaf and hearing children. *Hearing Research*, *309*, 94-102.

Eggermont, J. J., & Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In *Human Auditory Development* (pp. 61-105). Springer, New York, NY.

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory & Cognition*, *30*(3), 478-487.

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. *The Journal of the Acoustical Society of America*, *66*(3), 651-653.

Enns, J. T., & Cameron, S. (1987). Selective attention in young children: The relations between visual search, filtering, and priming. *Journal of Experimental Child Psychology*, *44*(1), 38-63.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2000). Children's perception of speech in multitalker babble. *The Journal of the Acoustical Society of America*, *108*(6), 3023-3029.

Fallon, M., Trehub, S. E., & Schneider, B. A. (2002). Children's use of semantic cues in degraded listening environments. *The Journal of the Acoustical Society of America*, *111*(5), 2242-2249.

Figueras, B., Edwards, L., & Langdon, D. (2008). Executive function and language in deaf children. *Journal of Deaf Studies and Deaf Education*, *13*(3), 362-377.

Finitzo-Hieber, T., & Tillman, T. W. (1978). Room acoustics effects on monosyllabic word discrimination ability for normal and hearing-impaired children. *Journal of Speech and Hearing Research*, *21*(3), 440-458.

Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, *12*(1), 47-65.

Forte, A. E., Etard, O., & Reichenbach, T. (2017). The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *ELife*, *6*, 1-12.

Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *109*(5), 2112-2122.

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America*, *106*(6), 3578-3588.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—focusing the searchlight on sound. *Current Opinion in Neurobiology*, *17*(4), 437-455.

Fuchigami, T., Okubo, O., Fujita, Y., Okuni, M., Noguchi, Y., & Yamada, T. (1993). Auditory event-related potentials and reaction time in children evaluation of cognitive development. *Developmental Medicine & Child Neurology*, *35*(3), 230-237.

Gałecki, A., & Burzykowski, T. (2013). Linear mixed-effects model. In *Linear Mixed-Effects*

Models Using R (pp. 245-273). Springer, New York, NY.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevelhierarchical models (Vol. 1)*. Cambridge University Press; New York, NY.

Geraci, M. (2014). Linear quantile mixed models: The lqmm package for Laplace quantile regression. *Journal of Statistical Software*, *57*(13), 1-29.

Geraci, M., & Bottai, M. (2014). Linear quantile mixed models. *Statistics and Computing*, *24*(3), 461–479.

Gomes, H., Molholm, S., Christodoulou, C., Ritter, W., & Cowan, N. (2000). The development of auditory attention in children. *Frontiers in Bioscience*, *5*(1), D108-D120.

Goodenough, F. L. (1935). The development of the reactive process from early childhood to maturity. *Journal of Experimental Psychology*, *18*(4), 431-450.

Gordon-Salant, S., & Fitzgibbons, P. J. (1993). Temporal factors and speech recognition performance in young and elderly listeners. *Journal of Speech, Language, and Hearing Research*, *36*(6), 1276-1285.

Gravel, J. S., Fausel, N., Liskow, C., & Chobot, J. (1999). Children's speech recognition in noise using omni-directional and dual- microphone hearing aid technology. *Ear and Hearing*, *20*(1), 1-11.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics* (Vol. 1). New York: Wiley.

Grieco-Calub, T. M., Collins, M. S., Snyder, H. E., & Ward, K. M. (2019). Background speech disrupts working memory span in 5-year-old children. *Ear and Hearing*, *40*(3), 437-446.

Grieco-Calub, T. M., Ward, K. M., & Brehm, L. (2017). Multitasking during degraded speech recognition in school-age children. *Trends in Hearing*, *21*, 1-14.

Hall, J. W., Grose, J. H., Buss, E., & Dev, M. B. (2002). Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children. *Ear and Hearing*, *23*(2), 159-165.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science*, *13*(4), 622-635.

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of the Acoustical Society of America*, *115*(2), 833-843.

Heim, S., Benasich, A. A., Wirth, N., & Keil, A. (2015). Tracking the attentional blink profile: A cross-sectional study from childhood to adolescence. *Psychological Research*, *79*(1), 19-27.

Hicks, C., & Tharpe, A. (2002). Listening effort and fatigue in school-age children with and without hearing loss. *Journal of Speech, Language and Hearing Research*, 573-584.

Hinkley, D. V., & Revankar, N. S. (1977). Estimation of the Pareto law from underreported data: A further analysis. *Journal of Econometrics*, *5*(1), 1-11.

Hirsh, I. J. (1950). The relation between localization and intelligibility. *The Journal of the Acoustical Society of America*, *22*(2), 196-200.

Holmes, E., Kitterick, P. T., & Summerfield, A. Q. (2017). Peripheral hearing loss reduces the ability of children to direct selective attention during multi-talker listening. *Hearing Research*, *350*, 160-172.

Houston, D. M., & Bergeson, T. R. (2014). Hearing versus listening: Attention to speech and its role in language acquisition in deaf infants with cochlear implants. *Lingua*, *139*, 10-25.

Iglehart, F. (2016). Speech perception in classroom acoustics by children with cochlear implants and with typical hearing. *American Journal of Audiology*, *25*(2), 100-109.

Irwin-Chase, H., & Burns, B. (2000). Developmental changes in children's abilities to share and allocate attention in a dual task. *Journal of Experimental Child Psychology*, *77*(1), 61-85.

Jamieson, D. G., Kranjc, G., Yu, K., & Hodgetts, W. E. (2004). Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *Journal of the American Academy of Audiology*, *15*(7), 508-517.

Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America*, *120*(4), 2177-2189.

Jones, G. L., & Litovsky, R. Y. (2011). A cocktail party model of spatial release from masking by both noise and speech interferers. *The Journal of the Acoustical Society of America*, *130*(3), 1463-1474.

Jones, P. R., Moore, D. R., & Amitay, S. A. (2015). Development of auditory selective attention: Why children struggle to hear in noisy environments. *Developmental Psychology*, *51*(3), 353-369.

Kahneman, D. (1973). *Attention and effort* (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kail, R. V. (1991). Development of processing speed in childhood and adolescence. In *Advances in Child Development and Behavior* (Vol. 23, pp. 151-185). JAI.

Kail, R. V., & Salthouse, T. A. (1994). Processing speed as a mental capacity. *Acta Psychologica*, *86*(2-3), 199-225.

Kail, R. V., & Ferrer, E. (2007). Processing speed in childhood and adolescence: Longitudinal models for examining developmental change. *Child Development*, *78*(6), 1760-1770.

Karatekin, C. (2004). Development of attentional allocation in the dual task paradigm. *International Journal of Psychophysiology*, *52*(1), 7-21.

Karns, C. M., Isbell, E., Giuliano, R. J., & 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. *Developmental Cognitive Neuroscience*, *13*, 53-67.

Kidd, G., Arbogast, T. L., Brughera, A., Mason, C. R., & Hartmann, W. M. (2002). The role of reverberation in release from masking due to spatial separation of sources in speech recognition. *The Journal of the Acoustical Society of America*, *111*(5), 2427-2427.

Kidd, G., Mason, C. R., Richards, V. M., Gallun, F. J., & Durlach, N. I. (2008). Informational masking. In *Auditory Perception of Sound Sources* (pp. 143-189). Springer, Boston, MA.

Klatte, M., Hellbrück, J., Seidel, J., & Leistner, P. (2010). Effects of classroom acoustics on performance and well-being in elementary school children: A field study. *Environment and Behavior*, *42*(5), 659-692.

Klatte, M., Lachmann, T., & 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*(49), 270-282.

Klatte, M., Lachmann, T., Schlittmeier, S., & Hellbrück, J. (2010). The irrelevant sound effect in short-term memory: Is there developmental change? *European Journal of Cognitive Psychology*, *22*(8), 1168-1191.

Klein, K. E., Walker, E. A., Kirby, B., & McCreery, R. W. (2017). Vocabulary facilitates speech perception in children with hearing aids. *Journal of Speech, Language, and Hearing Research*, *60*(8), 2281-2296.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3-to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*(1), 407-428.

Knecht, H. A., Nelson, P. B., Whitelaw, G. M., & Feth, L. L. (2002). Background noise levels and reverberation times in unoccupied classrooms. *American Journal of Audiology*, 65-71.

Koehnke, J., & Besing, J. M. (1996). A procedure for testing speech intelligibility in a virtual listening environment. *Ear and Hearing*, *17*(3), 211-217.

Kral, A., & Eggermont, J. J. (2007). What's to lose and what's to learn: Development under auditory deprivation, cochlear implants and limits of cortical plasticity. *Brain Research Reviews*, *56*(1), 259-269.

Kronenberger, W. G., Beer, J., Castellanos, I., Pisoni, D. B., & Miyamoto, R. T. (2014). Neurocognitive risk in children with cochlear implants. *JAMA Otolaryngology – Head & Neck Surgery*, *140*(7), 608-615.

Kronenberger, W. G., Xu, H., & Pisoni, D. B. (2020). Longitudinal Development of Executive Functioning and Spoken Language Skills in Preschool-Aged Children With Cochlear Implants. *Journal of Speech, Language, and Hearing Research*, 1-20.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,0000 English words. *Behavior Research Methods*, *44*(4), 978-990.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: tests in linear mixed effects models . *Journal of Statistical Software*, *82*(13).

Lehmann, A., & Schönwiesner, M. (2014). Selective attention modulates human auditory brainstem responses: relative contributions of frequency and spatial cues. *PloS ONE*, *9*(1).

Leibold, L. J. (2012). Development of auditory scene analysis and auditory attention. In *Human Auditory Development* (pp. 137-161). Springer, New York, NY.

Leibold, L. J. (2017). Speech perception in complex acoustic environments: Developmental effects. *Journal of Speech, Language, and Hearing Research*, *60*(10), 3001-3008.

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech shaped noise and two-talker masker. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1144-1155.

Leibold, L. J., & Buss, E. (2019). Masked speech recognition in school-age children. *Frontiers in Psychology*, *10*, 1981.

Leibold, L. J., Hillock-Dunn, A., Duncan, N., Roush, P. A., & 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 and Hearing*, *34*(5), 575-584.

Leon-Carrion, J., García-Orza, J., & Pérez-Santamaría, F. J. (2004). Development of the inhibitory component of the executive functions in children and adolescents. *International Journal of Neuroscience*, *114*(10), 1291-1311.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the*

Acoustical Society of America, 49(2B), 467-477.

Lewis, D., Hoover, B., Choi, S., & Stelmachowicz, P. (2010). The relationship between speech perception in noise and phonological awareness skills for children with normal hearing. *Ear and Hearing, 31*(6), 761.

Lewis, D. E., Mannimen, C. M., Valente, D. L., & Smith, N. A. (2014). Children's understanding of instructions presented in noise and reverberation. *American Journal of Audiology, 23*(3), 326-336.

Lewis, D., Schmid, K., O'Leary, S., Spalding, J., Heinrichs-Graham, E., & 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. *Journal of Speech, Language, and Hearing Research*, *59*(5), 1218-1232.

Litovsky, R. Y. (2015). Development of the auditory system. In *Handbook of Clinical Neurology* (Vol. 129, pp 55-72). Elsevier.

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*, *117*(5), 3091-3099.

Litovsky, R. Y. (2012). Spatial release from masking. *Acoustics Today*, *8*(2), 18–25.

Litovsky, R. Y., Johnstone, P. M., & Godar, S. P. (2006). Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology*, *45*(sup1), 78-91.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*(5), 1357-1372.

MacCutcheon, D., Pausch, F., Füllgrabe, C., Eccles, R., van der Linde, J., Panebianco, C., Fels, J., & Ljung, R. (2019). The contribution of individual differences in memory span and language ability to spatial release from masking in young children. *Journal of Speech, Language, and Hearing Research*, *62*(10), 3741-3751.

Machinskaya, R. I. (2006). Functional maturation of the brain and formation of the neurophysiological mechanisms of selective voluntary attention in young schoolchildren. *Human Physiology*, *32*(1), 20-29.

MacKeith, N. W., & Coles, R. R. A. (1971). Binaural advantages in hearing of speech. *The Journal of Laryngology & Otology*, *85*(3), 213-232.

Mackersie, C. L., & Cones, H. (2011). Subjective and psychophysiological indexes of listening effort in a competing-talker task. *Journal of the American Academy of Audiology*, *22*(2), 113-122.

Magimairaj, B. M., & Montgomery, J. W. (2012a). Children's verbal working memory: Role of processing complexity in predicting spoken sentence comprehension. *Journal of Speech, Language, and Hearing Research*, *55*(3), 669-682.

Magimairaj, B. M., & Montgomery, J. W. (2012b). Children's verbal working memory: Relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychologica*, *140*(3), 196-207.

Magimairaj, B. M., Nagaraj, N. K., & Benafield, N. (2018). Children's speech perception in noise: Evidence for dissociation from language and working memory. *Journal of Speech, Language, and Hearing Research, 61*(5), 1294-1305.

Manly, T., Anderson, V., Crawford, J., George, M., Underbjerg, M., & Robertson, I. (2016). Test of everyday attention for children (TEA–Ch2). London. *UK: Harcourt Assessment.*

Marrone, N., Mason, C. R., & Kidd Jr, G. (2008a). The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms. *The Journal of the Acoustical Society of America*, *124*(5), 3064-3075.

Marrone, N., Mason, C. R., & Kidd Jr, G. (2008b). Tuning in the spatial dimension: Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146-1158.

Martens, S., & Wyble, B. (2010). The attentional blink: Past, present, and future of a blind spot in perceptual awareness. *Neuroscience and Biobehavioral Reviews*, *34*(6), 947-957.

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*(7-8), 953-978.

McCreery, R. W., Spratford, M., Kirby, B., & Brennan, M. (2017). Individual differences in language and working memory affect children's speech recognition in noise. *International Journal of Audiology*, *56*(5), 306-315.

McCreery, R. W., Walker, E. A., Spratford, M., Lewis, D., & Brennan, M. (2019). Auditory, cognitive, and linguistic factors predict speech recognition in adverse listening conditions for children with hearing loss. *Frontiers in Neuroscience, 13,* 1093.

McFadden, B., & Pittman, A. (2008). Effect of minimal hearing loss on children's ability to multitask in quiet and in noise. *Language, Speech, and Hearing Services in Schools*, *39*(3), 342-351.

McGarrigle, R., Gustafson, S. J., Hornsby, B. W., & 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 and Hearing*, *40*(2), 381-392.

Meinhardt-Injac, B., Schlittmeier, S., Klatte, M., Otto, A., Persike, M., & Imhof, M. (2015). Auditory distraction by meaningless irrelevant speech: A developmental study. *Applied Cognitive Psychology*, *29*(2), 217-225.

Misurelli, S. M., & Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry. *The Journal of the Acoustical Society of America*, *132*(1), 380-391.

Misurelli, S. M., & Litovsky, R. Y. (2015). Spatial release from masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity. *The Journal of the Acoustical Society of America*, *138*(1), 319-331.

Mitchell, T. V., & Quittner, A. L. (1996). Multimethod study of attention and behavior problems in hearing-impaired children. *Journal of Clinical Child Psychology*, *25*(1), 83-96.

Monroy, C., Shafto, C., Castellanos, I., Bergeson, T., & Houston, D. (2019). Visual habituation in deaf and hearing infants. *PloS ONE*, *14*(2), 1-15.

Moore, B. C., & Gockel, H. E. (2012). Properties of auditory stream formation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1591), 919-931.

Moore, D. R. (2012). Listening difficulties in children: Bottom-up and top-down contributions. *Journal of Communication Disorders*, *45*(6), 411-418.

Moore, D. R., Ferguson, M. A., Halliday, L. F., & Riley, A. (2008). Frequency discrimination in children: Perception, learning and attention. *Hearing Research*, *238*(1-2), 147-154.

Moore, J. K., & Linthicum Jr, F. H. (2007). The human auditory system: a timeline of development. *International Journal of Audiology*, *46*(9), 460-478.

Muth, C., Bales, K. L., Hinde, K., Maninger, N., Mendoza, S. P., & Ferrer, E. (2016). Alternative models for small samples in psychological research: applying linear mixed effects models and generalized estimating equations to repeated measures data. *Educational and Psychological Measurement*, *76*(1), 64-87.

Naylor, G. M. (1993). ODEON – Another hybrid room acoustical model. *Applied Acoustics*, *38*(2-4), 131-143.

Nelson, E. L., Smaldino, J., Erler, S., & Garstecki, D. (2008). Background noise levels and reverberation times in old and new elementary school classrooms. *Journal of Educational Audiology*, *14*, 16-22.

Nelson, P., & Soli, S. (2000). Acoustical barriers in learning: Children at risk in every classroom. *Language, Speech and Hearing Services in Schools*, *31*(4), 356-361.

Neuman, A. C., & Hochberg, I. (1983). Children's perception of speech in reverberation. *The Journal of the Acoustical Society of America*, *73*(6), 2145-2149.

Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear and Hearing*, *31*(3), 336-344.

Nippold, M. A. (1998). *Later language development: The school-age and adolescent years*. Pro-Ed, 8700 Shoal Creek Blvd., Austin, TX 78757-6897.

Nittrouer, S., & Boothroyd, A. (1990). Context effects in phoneme and word recognition by young children and older adults. *The Journal of the Acoustical Society of America*, *87*(6), 2705-2715.

Oberg, E., & Lukomski, J. (2011). Executive functioning and the impact of a hearing loss: Performance-based measures and the Behavior Rating Inventory of Executive Function (BRIEF). *Child Neuropsychology*, *17*(6), 521-545.

Oleson, J. J., Brown, G. D., & McCreery, R. (2019). The evolution of statistical methods in speech, language, and hearing sciences. *Journal of Speech, Language, and Hearing Research*, *62*(3), 498-506.

Oxenham, A. J. (2008). Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends in Amplification*, *12*(4), 316-331.

Pearson, D. A., & Lane, D. M. (1991). Auditory attention switching: A developmental study. *Journal of Experimental Child Psychology*, *51*(2), 320-334.

Picard, M., & Bradley, J. S. (2001). Revisiting speech interference in classrooms. *Audiology*, *40*(5), 221-244.

Pinheiro, J., & Bates, D. (2000). Linear mixed-effects models: basic concepts and examples. In *Mixed-effects Models in S and S-Plus* (pp. 3-56). Springer-Verlag.

Pisoni, D. B., & Cleary, M. (2003). Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation. *Ear and Hearing*, *24*(1 Suppl), 106S.

Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, *28*(6), 723-732.

Plomp, R. (1976). *Aspect of Tone Sensation* (pp. 85-110).

Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica*, *86*(2–3), 227-272.

Pollack, I. (1975). Auditory informational masking. *The Journal of the Acoustical Society of*

America, *57*(S1), S5-S5.

Porter, H., Sladen, D. P., Ampah, S. B., Rothpletz, A., & Bess, F. H. (2013). Developmental outcomes in early school-age children with minimal hearing loss. *American Journal of Audiology*, *22*(2), 263-270.

Quittner, A. L., Smith, L. B., Osberger, M. J., Mitchell, T. V, & Katz, D. B. (1994). The impact of audition on the development of visual attention: Analyses of deaf children using cochlear implants. *Psychological Science*, *5*(6), 347–353.

R Core Team (2019). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*. https://www.r-project.org/

Ricketts, T., Galster, J., & Tharpe, A. M. (2007). Directional benefit in simulated classroom environments. *American Journal of Audiology*, *16*(2), 130-144.

Rimmele, J. M., Zion Golumbic, E., Schröger, E., & Poeppel, D. (2015). The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex*, *68*, 144-154.

Robinson, C., & Schumacker, R. (2009). Interaction effects: centering, variance inflation factor, and interpretation issues. *Multiple Linear Regression Viewpoints*, *35*(1), 6-11.

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., ... & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, 31.

Rothauser, E. H. (1969). IEEE recommended practice for speech quality measurements. *IEEE Trans. on Audio and Electroacoustics*, *17*, 225-246.

Sanders, L. D., Stevens, C., Coch, D., & Neville, H. J. (2006). Selective auditory attention in 3- to 5-year-old children: An event-related potential study. *Neuropsychologia*, *44*(11), 2126-2138.

Sanes, D. H., & Woolley, S. M. N. (2011). A behavioral framework to guide research on central auditory development and plasticity. *Neuron*, *72*(6), 912-929.

Sarant, J. Z., Harris, D. C., & Bennet, L. A. (2015). Academic outcomes for school-aged children with severe–profound hearing loss and early unilateral and bilateral cochlear implants. *Journal of Speech, Language, and Hearing Research*, *58*(3), 1017-1032.

Shapiro, K. L., Raymond, J. E., & Arnell, K. M. (1997). The attentional blink. *Trends in Cognitive Sciences*, *1*(8), 291-296.

Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *The Journal of the Acoustical Society of America*, *123*(1), 133-144.

Shield, B. M., & Dockrell, J. E. (2003). The effect of noise on children at school: A review. *Building Acoustics*, *10*(2), 97-106.

Shinn-Cunningham, B. G. (2008). Object-based auditory and visual attention. *Trends in Cognitive Sciences*, *12*(5), 182-186.

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283-299.

Smith, L. B., Quittner, A. L., Osberger, M. J., & Miyamoto, R. (1998). Audition and visual attention: the developmental trajectory in deaf and hearing populations. *Developmental Psychology*, *34*(5), 840.

Smoski, W. J., Brunt, M. A., & Tannahill, C. J. (1998). Children's Auditory Performance Scale (CHAPS). In *Phonak* (Vol. 2). Tampa, FL: Educational Audiology Association.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(2), 174-215.

Snyder, J. S., Gregg, M. K., Weintraub, D. M., & Alain, C. (2012). Attention, awareness, and the perception of auditory scenes. *Frontiers in Psychology*, *3*, 15.

Stenfelt, S., & Rönnberg, J. (2009). The signal-cognition interface: Interactions between degraded auditory signals and cognitive processes. *Scandinavian Journal of Psychology*, *50*(5), 385-393.

Studebaker, G. A., Gray, G. A., & Branch, W. E. (1999). Prediction and statistical evaluation of speech recognition test scores. *Journal of the American Academy of Audiology*, *10*(7), 355-370.

Studebaker, G. A. (1985). A" rationalized" arcsine transform. *Journal of Speech, Language, and Hearing Research*, *28*(3), 455-462.

Sullivan, J. R., Osman, H., & Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *Journal of Speech, Language, and Hearing Research*, *58*(3), 1043-1051.

Sussman, E. S., Ceponiene, R., Shestakova, A., Näätänen, R., & Winkler, I. (2001). Auditory stream segregation processes operate similarly in school-aged children and adults. *Hearing Research*, *153*(1–2), 108-114.

Sussman, E. S., Horváth, J., Winkler, I., & Orr, M. (2007). The role of attention in the formation of auditory streams. *Perception & Psychophysics*, *69*(1), 136-152.

Sussman, E. S. (2017). Auditory scene analysis: An attention perspective. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2989-3000.

Sussman, E. S., & Steinschneider, M. (2009). Attention effects on auditory scene analysis in children. *Neuropsychologia*, *47*(3), 771-785.

Tharpe, A. M., Ashmead, D. H., & Rothpletz, A. M. (2002). Visual attention in children with normal hearing, children with hearing aids, and children with cochlear implants. *Journal of Speech, Language, and Hearing Research*, *45*(2), 403-413.

Thomas, J. R., Gallagher, J. D., & Purvis, G. J. (1981). Reaction time and anticipation time: Effects of development. *Research Quarterly for exercise and sport*, *52*(3), 359-367.

Thompson, E. C., Krizman, J., White-Schwoch, T., Nicol, T., Estabrook, R., & Kraus, N. (2019). Neurophysiological, linguistic, and cognitive predictors of children's ability to perceive speech in noise. *Developmental Cognitive Neuroscience*, *39*, 100672.

Torkildsen, J. V. K., Hitchins, A. R. C., Myhrum, M., & Wie, O. B. (2019). Speech-in-noise perception in children with cochlear implants, hearing aids, developmental language disorder and typical development: The effects of linguistic and cognitive abilities. *Frontiers in Psychology*, *10*, 2530.

Walker, E. A., Sapp, C., Dallapiazza, M., Spratford, M., McCreery, R. W., & Oleson, J. J. (2020). Language and reading outcomes in fourth-grade children with mild hearing loss compared to age-matched hearing peers. *Language, Speech, and Hearing Services in Schools*, *51*(1), 17-28.

Walker, E., Sapp, C., Oleson, J., & McCreery, R. W. (2019). Longitudinal speech recognition in noise in children: Effects of hearing status and vocabulary. *Frontiers in Psychology*, *10*, 2421.

Weintraub, S., Anderson, J. E., & Manly, J. J. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*.

Wightman, F., & Allen, P. (2004). Individual differences in auditory capability among preschool children. *Developmental Psychoacoustics*, 113–133.

Wightman, F. L., & Kistler, D. J. (2005). Informational masking of speech in children: Effects of ipsilateral and contralateral distracters. *Journal of the Acoustical Society of America*, *118*(5), 3164-3176.

Wild, C. J., Yusuf, A., Wilson, D. E., Peelle, J. E., Davis, M. H., & Johnsrude, I. S. (2012). Effortful listening: the processing of degraded speech depends critically on attention. *The*

Journal of Neuroscience: The Official Journal of the Society for Neuroscience, *32*(40), 14010-14021.

Wu, L. (2009). *Mixed Effects Models for Complex Data*. CRC Press.

Yacullo, W. S., & Hawkins, D. B. (1987). Speech Recognition in Noise and Reverberation by School-Age Children. *Audiology*, *26*(4), 235-246.

Yoncheva, Y. N., Zevin, J. D., Maurer, U., & McCandliss, B. D. (2010). Auditory selective attention to speech modulates activity in the visual word form area. *Cerebral Cortex*, *20*(3), 622-632.

Yoo, W., Mayberry, R., Bae, S., Singh, K., He, Q. P., & Lillard Jr, J. W. (2014). A study of effects of multicollinearity in the multivariable analysis. *International Journal of Applied Science and Technology*, *4*(5), 9.

Zurek, P. M. (1993). A note on onset effects in binaural hearing. *Journal of the Acoustical Society of America*, *93*(2), 1200-1201.

APPENDIX

A.2.1: Linear quantile mixed-effects model output with main effects, two-way interactions, and three-way interactions for each dependent variable of interest.

	Estimate	SE	95% CI Lower Bound	95% CI Upper Bound	*p*
A. Hits					
AIC = -5.37; df = 10					
References: NH; ACD					
Intercept	0.854	0.019	0.816	0.891	**<0.001**
Age	0.033	0.005	0.022	0.044	**<0.001**
Group	-0.245	0.056	-0.357	-0.133	**<0.001**
Task	-0.054	0.022	-0.099	-0.009	**<0.05**
Age × Group	0.044	0.030	-0.017	0.105	0.154
Age × Task	0.002	0.009	-0.016	0.021	0.809
Group × Task	0.091	0.064	-0.038	0.219	0.163
Age × Group × Task	-0.013	0.037	-0.088	0.062	0.728
B. False Alarms					
AIC = -263.40; df = 10					
References: NH; ACD					
Intercept	0.045	0.006	0.031	0.058	**<0.001**
Age	-0.005	0.002	-0.010	-0.001	**<0.05**
Group	0.037	0.018	0.001	0.074	**<0.05**
Task	0.003	0.009	-0.014	0.021	0.712
Age × Group	0.0006	0.009	-0.018	0.019	0.947
Age × Task	0.006	0.003	0.0002	0.013	**<0.05**
Group × Task	-0.0003	0.041	-0.083	0.082	0.995
Age × Group × Task	-0.017	0.023	-0.063	0.028	0.439
C. Response Sensitivity					
AIC = 374.20; df = 10					
References: NH; ACD					
Intercept	3.249	0.092	3.065	3.434	**<0.001**
Age	0.194	0.041	0.112	0.276	**<0.001**
Group	-1.249	0.231	-1.714	-0.785	**<0.001**
Task	-0.484	0.132	-0.749	-0.220	**<0.001**

Age × Group	-0.011	0.135	-0.283	0.260	0.933
Age × Task	-0.063	0.067	-0.197	0.072	0.354
Group × Task	0.657	0.394	-0.135	1.449	0.102
Age × Group × Task	0.132	0.248	-0.366	0.630	0.597
D. Non-Deviant Responses					
AIC = 813.50; df = 10					
References: NH; ACD					
Intercept	3.443	0.532	2.373	4.512	**<0.001**
Age	-0.549	0.192	-0.935	-0.162	**<0.01**
Group	6.561	1.841	2.861	10.262	**<0.001**
Task	-2.443	0.479	-3.406	-1.480	**<0.001**
Age × Group	0.540	1.049	-1.568	2.649	0.609
Age × Task	0.549	0.191	0.165	0.932	**<0.01**
Group × Task	-5.305	2.519	-10.368	-0.242	**<0.05**
Age × Group × Task	-1.123	0.908	-2.947	0.701	0.222
E. Reaction Time for Hits					
AIC = 1915.00; df = 10					
References: NH; ACD					
Intercept	863.953	18.228	827.322	900.584	**<0.001**
Age	-58.975	6.976	-72.994	-44.956	**<0.001**
Group	213.626	57.916	97.240	330.011	**<0.001**
Task	-178.909	17.494	-214.065	-143.753	**<0.001**
Age × Group	-30.072	31.657	-93.689	33.544	0.347
Age × Task	1.954	7.621	-13.362	17.269	0.799
Group × Task	-175.611	50.734	-277.564	-73.658	**<0.01**
Age × Group × Task	1.125	27.863	-54.868	57.119	0.968
F. Reaction Time for False Alarms					
AIC = 1600.00; df = 10					
References: NH; ACD					
Intercept	867.175	66.296	733.948	1000.402	**<0.001**
Age	-67.658	18.488	-104.811	-30.506	**<0.001**
Group	351.350	121.978	106.226	596.473	**<0.01**
Task	-81.437	103.570	-289.568	126.694	0.435
Age × Group	-69.628	63.579	-197.394	58.138	0.279
Age × Task	47.512	30.489	-13.758	108.781	0.126
Group × Task	-349.453	670.923	-1697.724	998.818	0.605
Age × Group × Task	31.320	275.834	-522.989	585.629	0.910

A.2.2: Linear quantile mixed-effects model output with main effects and two-way interactions for each dependent variable of interest. Post hoc paired comparisons are displayed for significant or trending-toward-significant effects.

	Estimate	SE	95% CI Lower Bound	95% CI Upper Bound	*p*
A. Hits					
AIC = -6.74; df = 9					
References: NH; ACD					
Intercept	0.853	0.019	0.815	0.892	**<0.001**
Age	0.034	0.005	0.024	0.045	**<0.001**
Group	-0.242	0.050	-0.343	-0.141	**<0.001**
Task	-0.053	0.023	-0.100	-0.007	**<0.05**
Age × Group	0.037	0.021	-0.005	0.080	0.084
Age × Task	0.000	0.010	-0.019	0.020	0.975
Group × Task	0.084	0.045	-0.005	0.174	0.065
ACD: HL vs. NH	-0.242	0.050	-0.343	-0.141	**<0.001**
VCD: HL vs. NH	-0.158	0.051	-0.260	-0.056	**<0.01**
NH: VCD vs. ACD	-0.053	0.023	-0.100	-0.007	**<0.05**
HL: VCD vs. ACD	0.037	0.064	-0.091	0.165	0.566
B. False Alarms					
AIC = -268.70; df = 9					
References: NH; ACD					
Intercept	0.044	0.006	0.031	0.057	**<0.001**
Age	-0.004	0.002	-0.008	-0.0001	**<0.05**
Group	0.042	0.018	0.005	0.078	**<0.05**
Task	0.004	0.008	-0.013	0.020	0.657
Age × Group	-0.008	0.009	-0.026	0.010	0.358
Age × Task	0.004	0.003	-0.001	0.009	0.159
Group × Task	-0.009	0.035	-0.079	0.061	0.805
C. Response Sensitivity					
AIC = 372.90; df = 9					
References: NH; ACD					
Intercept	3.255	0.088	3.079	3.432	**<0.001**
Age	0.187	0.042	0.102	0.272	**<0.001**
Group	-1.258	0.212	-1.684	-0.832	**<0.001**
Task	-0.491	0.127	-0.746	-0.235	**<0.001**

Age × Group	0.062	0.087	-0.113	0.236	0.480
Age × Task	-0.056	0.069	-0.194	0.082	0.420
Group × Task	0.700	0.354	-0.012	1.413	0.054
ACD: HL vs. NH	-1.258	0.212	-1.684	-0.832	**<0.001**
VCD: HL vs. NH	-0.571	0.245	-1.063	-0.078	**<0.05**
NH: VCD vs. ACD	-0.491	0.127	-0.746	-0.235	**<0.001**
HL: VCD vs. ACD	0.174	0.375	-0.579	0.927	0.645
D. Non-Deviant Responses					
AIC = 813.50; df = 9					
References: NH; ACD					
Intercept	3.182	0.509	2.160	4.204	**<0.001**
Age	-0.427	0.209	-0.847	-0.006	**<0.05**
Group	6.515	1.722	3.053	9.976	**<0.001**
Task	-2.182	0.467	-3.120	-1.244	**<0.001**
Age × Group	-0.224	0.768	-1.766	1.319	0.772
Age × Task	0.427	0.211	0.003	0.850	**<0.05**
ACD: Age	-0.427	0.209	-0.847	-0.006	**<0.05**
VCD: Age	-0.00006	0.089	-0.178	0.178	0.999
Group × Task	-6.051	2.261	-10.595	-1.507	**<0.05**
ACD: HL vs. NH	6.515	1.722	3.053	9.976	**<0.001**
VCD: HL vs. NH	0.474	1.711	-2.963	3.912	0.783
NH: VCD vs. ACD	-2.182	0.467	-3.120	-1.244	**<0.001**
HL: VCD vs. ACD	-7.579	2.471	-12.545	-2.612	**<0.01**
E. Reaction Time for Hits					
AIC = 1914.00; df = 9					
References: NH; ACD					
Intercept	864.111	18.957	826.016	902.207	**<0.001**
Age	-59.013	6.553	-72.182	-45.845	**<0.001**
Group	211.065	53.804	102.942	319.189	**<0.001**
Task	-180.296	18.197	-216.864	-143.729	**<0.001**
Age × Group	-29.070	19.410	-68.076	9.935	0.141
Age × Task	2.349	6.843	-11.402	16.101	0.733
Group × Task	-171.034	47.119	-265.723	-76.345	**<0.001**
ACD: HL vs. NH	211.065	53.804	102.942	319.189	**<0.001**
VCD: HL vs. NH	38.803	40.875	-43.339	120.945	0.347
NH: VCD vs. ACD	-180.296	18.197	-216.864	-143.729	**<0.001**
HL: VCD vs. ACD	-337.169	49.657	-436.958	-237.380	**<0.001**
F. Reaction Time for False Alarms					

AIC = 1598.00; df = 9					
References: NH; ACD					
Intercept	863.431	67.514	727.756	999.106	**<0.001**
Age	-69.117	18.833	-106.964	-31.270	**<0.001**
Group	359.625	129.646	99.092	620.157	**<0.01**
Task	-77.957	102.776	-284.493	128.579	0.452
Age × Group	-53.846	68.131	-190.761	83.069	0.433
Age × Task	53.636	27.581	-1.789	109.061	0.058
ACD: Age	-69.117	18.833	-106.964	-31.270	**<0.001**
VCD: Age	-15.472	18.814	-53.281	22.337	0.415
Group × Task	-377.976	252.946	-886.289	130.337	0.142

A.2.3: Linear quantile mixed-effects model output displaying the contribution of aided and unaided pure-tone averages to the observed effect of group.

	Estimate	SE	95% CI		*p*
			Lower Bound	Upper Bound	
A. Base Model					
AIC = 374.70; df = 6					
References: NH; ACD					
Intercept	3.177	0.083	3.010	3.344	**<0.001**
Age	0.164	0.027	0.110	0.217	**<0.001**
Group	-0.987	0.164	-1.315	-0.658	**<0.001**
Task	-0.361	0.125	-0.612	-0.110	**<0.01**
B. Aided PTA with Group					
AIC = 372.80; df = 7					
References: NH; ACD					
Intercept	3.200	0.107	2.986	3.416	**<0.001**
Age	0.163	0.035	0.092	0.234	**<0.001**
Group	-0.695	0.247	-1.190	-0.200	**<0.01**
Task	-0.391	0.139	-0.669	-0.112	**<0.01**
Aided PTA	-0.020	0.019	-0.058	0.019	0.309
C. Unaided PTA with Group					
AIC = 376.20; df = 7					

References: NH; ACD					
Intercept	3.184	0.097	2.989	3.380	**<0.001**
Age	0.155	0.022	0.111	0.199	**<0.001**
Group	-0.583	0.425	-1.438	0.271	0.176
Task	-0.369	0.130	-0.630	-0.107	**<0.01**
Unaided PTA	-0.007	0.010	-0.027	0.012	0.440

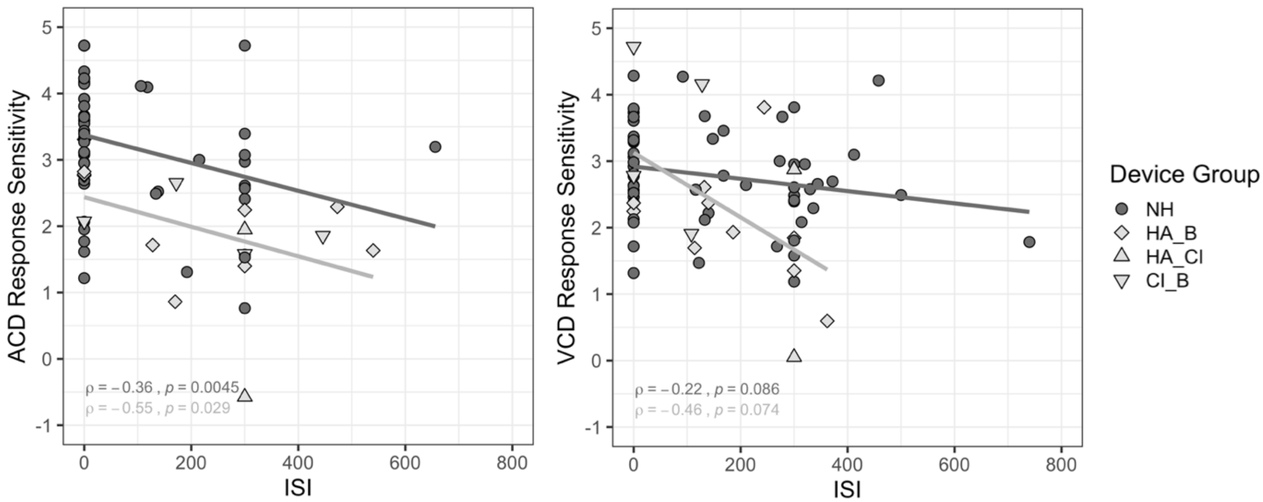
A.2.4: Linear quantile mixed-effects model output displaying the relation between aided pure-tone average, unaided pure-tone average, and response sensitivity.

			95% CI		
	Estimate	**SE**	**Lower Bound**	**Upper Bound**	***p***
A. Aided PTA without Group					
AIC = 368.90; df = 8					
References: ACD					
Intercept	3.075	0.081	2.913	3.238	**<0.001**
Age	0.147	0.025	0.097	0.198	**<0.001**
Task	-0.348	0.101	-0.552	-0.145	**<0.01**
Aided PTA	-0.061	0.010	-0.081	-0.041	**<0.001**
Aided PTA × Task	0.021	0.017	-0.014	0.055	0.236
Aided PTA × Age	0.014	0.005	0.005	0.023	**<0.01**
B. Unaided PTA without Group					
AIC = 367.30; df = 8					
Reference: ACD					
Intercept	3.201	0.078	3.045	3.357	**<0.001**
Age	0.135	0.023	0.090	0.181	**<0.001**
Task	-0.522	0.100	-0.724	-0.321	**<0.001**
Unaided PTA	-0.024	0.003	-0.031	-0.018	**<0.001**
Unaided PTA × Task	0.015	0.005	0.004	0.026	**<0.01**
ACD: Unaided PTA	-0.024	0.003	-0.031	-0.018	**<0.001**
VCD: Unaided PTA	-0.009	0.004	-0.016	-0.002	**<0.05**
Unaided PTA × Age	0.004	0.002	0.000	0.007	**<0.05**

A.2.5: Linear quantile mixed-effects model output displaying the relation between performance on standardized measures of executive function and response sensitivity.

	Estimate	SE	95% CI Lower Bound	95% CI Upper Bound	*p*
A. Flanker Inhibitory Control and Attention Task					
AIC = 414.2; df = 5					
Reference: Not Applicable					
Intercept	2.818	0.069	2.679	2.958	**<0.001**
Age	0.151	0.038	0.074	0.228	**<0.001**
Flanker Score	0.003	0.007	-0.010	0.017	0.636
B. Simple Reaction Time					
AIC = 379.9; df = 5					
Reference: Not Applicable					
Intercept	2.866	0.103	2.660	3.072	**<0.001**
Age	0.108	0.038	0.032	0.185	**<0.01**
Simple RT	-0.0006	0.0004	-0.001	0.0003	0.199
C. Hide and Seek Auditory/Cerberus					
AIC = 385.1; df = 5					
Reference: Not Applicable					
Intercept	2.891	0.086	2.719	3.064	**<0.001**
Age	0.123	0.028	0.067	0.178	**<0.001**
Hide and Seek Auditory/Cerberus	-0.0002	0.00007	-0.0003	-0.0001	**<0.05**
D. Hide and Seek Auditory/Cerberus with Task					
AIC = 380.4; df = 7					
Reference: ACD					
Intercept	3.029	0.124	2.780	3.278	**<0.001**
Age	0.128	0.028	0.073	0.184	**<0.001**
Task	-0.361	0.144	-0.651	-0.071	**<0.05**
Hide and Seek Auditory/Cerberus	-0.0002	0.0001	-0.0004	-0.0001	**<0.05**
Task × Hide and Seek Auditory/Cerberus	0.0001	0.0001	-0.00009	0.0004	0.230

<u>A.2.6</u>: Correlations between response sensitivity and ISI (in ms) for children with NH (*dark gray*) and children with HL (*light gray)* during the auditory (*left*) and visual (*right*) change detection tasks.



A.3.1: Linear mixed-effects model output for the no-masker conditions.

	Estimate	SE	Test ($df_{Den.}$)	*p*
DV = Speech Recognition (RAU)				
df = 160				
References: HL; High				
Age	3.17	0.72	F = 19.344 (58)	**<0.001**
Group	18.27	1.46	F = 155.642 (58)	**<0.001**
Reverberation Time			F = 75.537 (116)	**<0.001**
Low vs. Moderate	14.60	1.51	t = 9.654 (22)	**<0.001**
Low vs. High	19.76	1.51	t = 13.067 (22)	**<0.001**
Moderate vs. High	5.16	1.51	t = 3.413 (22)	**<0.01**
Age × Group			F = 0.737 (58)	0.394
Age × Reverberation Time			F = 0.785 (116)	0.458
Group × Reverberation Time			F = 4.882 (116)	**<0.01**
NH (Low vs. Moderate)	9.50	1.74	t = 5.469 (62)	**<0.001**
NH (Low vs. High)	16.76	1.74	t = 9.652 (62)	**<0.001**
NH (Moderate vs. High)	7.26	1.74	t = 4.183 (62)	**<0.01**
HL (Low vs. Moderate)	19.70	1.75	t = 11.251 (7)	**<0.01**
HL (Low vs. High)	22.76	1.75	t = 12.998 (7)	**<0.001**
HL (Moderate vs. High)	3.06	1.75	t = 1.747 (7)	0.999
Age × Group × Reverberation Time			F = 2.483 (116)	0.088

A.3.2: Linear mixed-effects model output for the masker conditions.

	Estimate	SE	Test ($df_{Den.}$)	*p*
DV = Speech Recognition (RAU)				
df = 1114				
References: HL; Two-Talker Speech; High; Co-located				
Intercept	54.346	1.448	t = 37.529	**<0.001**
Age	3.048	0.714	F = 18.226 (66)	**<0.001**
Group	6.069	1.448	F = 17.562 (66)	**<0.001**
Reverberation Time	-	-	F = 103.036 (1122)	**<0.001**
Low vs. Moderate	0.980	0.848	t = 1.156 (1194)	0.744
Low vs. High	10.670	0.848	t = 12.586 (1194)	**<0.001**
Moderate vs. High	9.690	0.848	t = 11.430 (1194)	**<0.001**
Masker Type	5.070	0.336	F = 228.301 (1122)	**<0.001**
Spatial Location	-	-	F = 174.916 (1122)	**<0.001**
Separated (+90) vs. Co-located	11.600	0.848	t = 13.631 (1194)	**<0.001**
Separated (+90/-90) vs. Co-located	-3.000	0.848	t = -3.534 (1194)	**<0.01**
Separated (+90) vs. Separated (+90/-90)	14.600	0.848	t = 17.165 (1194)	**<0.001**
Age × Group	-	-	F = 0.002 (66)	0.965
Age × Reverberation Time	-	-	F = 1.657 (1122)	0.191

Age × Masker Type	-	-	F = 10.074 (1122)	**<0.01**
Age [Youngest] (Speech-Shaped vs. Two-Talker)	13.860	1.44	t = 9.596 (1194)	**<0.001**
Age [Oldest] (Speech-Shaped vs. Two-Talker)	5.620	1.57	t = 3.586 (1194)	**<0.01**
Age × Spatial Location	-	-	F = 0.351 (1122)	0.704
Group × Reverberation Time	-	-	F = 2.909 (1122)	0.055
NH (Low vs. Moderate)	-0.388	0.805	t = -0.482 (1194)	0.999
NH (Low vs. High)	8.747	0.805	t = 10.872 (1194)	**<0.001**
NH (Moderate vs. High)	9.135	0.805	t = 11.354 (1194)	**<0.001**
HL (Low vs. Moderate)	2.349	1.493	t = 1.573 (1194)	**0.695**
HL (Low vs. High)	12.600	1.493	t = 8.448 (1194)	**<0.001**
HL (Moderate vs. High)	10.251	1.493	t = 6.865 (1194)	**<0.001**
Group × Masker Type	-	-	F = 43.740 (1122)	**<0.001**
NH (Speech-Shaped vs. Two-Talker)	5.700	0.657	t = 8.680 (1194)	**<0.001**
HL (Speech-Shaped vs. Two-Talker)	14.600	1.219	t = 11.958 (1194)	**<0.001**
Group × Spatial Location	-	-	F = 14.006 (1122)	**<0.001**
NH (Separated [+90] vs. Co-located)	15.812	0.805	t = 19.653 (1194)	**<0.001**
NH (Separated [+90/-90] vs. Co-located)	-1.670	0.805	t = -2.075 (1194)	**0.229**
NH (Separated [+90] vs. Separated [+90/-90])	17.481	0.805	t = 21.728 (1194)	**<0.001**
HL (Separated [+90] vs. Co-located)	7.308	1.493	t = 4.895 (1194)	**<0.001**
HL (Separated [+90/-90] vs. Co-located)	-4.325	1.493	t = -2.897 (1194)	**0.023**
HL (Separated [+90] vs. Separated [+90/-90])	11.634	1.493	t = 7.791 (1194)	**<0.001**
Reverberation Time × Masker Type	-	-	F = 35.685 (1122)	**<0.001**
Low (Speech-Shaped vs. Two-Talker)	2.160	1.20	t = 1.804 (1194)	0.215
Moderate (Speech-Shaped vs. Two-Talker)	13.430	1.20	t = 11.195 (1194)	**<0.001**
High (Speech-Shaped vs. Two-Talker)	14.830	1.20	t = 12.367 (1194)	**<0.001**
Reverberation Time × Spatial Location	-	-	F = 15.266 (1122)	**<0.001**
Low (Separated [+90] vs. Co-located)	13.048	1.47	t = 8.883 (1194)	**<0.001**
Low (Separated [+90/-90] vs. Co-located)	3.488	1.47	t = 2.375 (1194)	0.159
Low (Separated [+90] vs. Separated [+90/-90])	9.560	1.47	t = 6.508 (1194)	**<0.001**
Moderate (Separated [+90] vs. Co-located)	16.032	1.47	t = 10.914 (1194)	**<0.001**
Moderate (Separated [+90/-90] vs. Co-located)	-3.429	1.47	t = -2.334 (1194)	0.177
Moderate (Separated [+90] vs. Separated [+90/-90])	19.461	1.47	t = 13.249 (1194)	**<0.001**
High (Separated [+90] vs. Co-located)	5.601	1.47	t = 3.813 (1194)	**<0.001**
High (Separated [+90/-90] vs. Co-located)	-9.051	1.47	t = -6.162 (1194)	**<0.001**
High (Separated [+90] vs. Separated [+90/-90])	14.652	1.47	t = 9.975 (1194)	**<0.001**
Masker Type × Spatial Location	-	-	F = 18.755 (1122)	**<0.001**
Speech-Shaped (Separated [+90] vs. Co-located)	10.198	1.20	t = 8.503 (1194)	**<0.001**
Speech-Shaped (Separated [+90/-90] vs. Co-located)	0.518	1.20	t = 0.432 (1194)	0.999
Speech-Shaped (Separated [+90] vs. Separated [+90/-90])	9.679	1.20	t = 8.070 (1194)	**<0.001**
Two-Talker (Separated [+90] vs. Co-located)	12.923	1.20	t = 10.775 (1194)	**<0.001**
Two-Talker (Separated [+90/-90] vs. Co-located)	-6.513	1.20	t = -5.431 (1194)	**<0.001**
Two-Talker (Separated [+90] vs. Separated [+90/-90])	19.436	1.20	t = 16.205 (1194)	**<0.001**
Age × Group × Reverberation Time	-	-	F = 2.600 (1122)	0.074
Age × Group × Masker Type	-	-	F = 5.329 (1122)	**<0.05**

Age × Group × Spatial Location	-	-	F = 1.096 (1122)	0.334
Age × Reverberation Time × Masker Type	-	-	F = 1.689 (1122)	0.185
Age × Reverberation Time × Spatial Location	-	-	F = 2.971 (1122)	**<0.05**
Age × Masker Type × Spatial Location	-	-	F = 1.072 (1122)	0.342
Group × Reverberation Time × Masker Type	-	-	F = 1.131 (1122)	0.323
Group × Reverberation Time × Spatial Location	-	-	F = 0.903 (1122)	0.461
Group × Masker Type × Spatial Location	-	-	F = 0.750 (1122)	0.472
Reverberation Time × Masker Type × Spatial Location	-	-	F = 11.089 (1122)	**<0.001**
Age × Reverberation Time × Masker Type × Spatial Location	-	-	F = 0.875 (1122)	0.478
Age × Group × Reverberation Time × Masker Type	-	-	F = 0.449 (1122)	0.639
Age × Group × Reverberation Time × Spatial Location	-	-	F = 1.576 (1122)	0.177
Age × Group × Masker Type × Spatial Location	-	-	F = 1.173 (1122)	0.310
Group × Reverberation Time × Masker Type × Spatial Location	-	-	F = 3.111 (1122)	**<0.05**
Speech-Shaped, Low, +90 (NH vs. HL)	8.822	4.13	t = 2.135 (245)	0.608
Speech-Shaped, Low, +90/-90 (NH vs. HL)	5.508	4.13	t = 1.333 (245)	0.999
Speech-Shaped, Low, Co-located (NH vs. HL)	-1.699	4.13	t = -0.411 (245)	0.999
Speech-Shaped, Moderate, +90 (NH vs. HL)	14.134	4.13	t = 3.42 (245)	**<0.05**
Speech-Shaped, Moderate, +90/-90 (NH vs. HL)	3.829	4.13	t = 0.926 (245)	0.999
Speech-Shaped, Moderate, Co-located (NH vs. HL)	7.109	4.13	t = 1.720 (245)	0.999
Speech-Shaped, High, +90 (NH vs. HL)	15.316	4.13	t = 3.706 (245)	**<0.01**
Speech-Shaped, High, +90/-90 (NH vs. HL)	7.220	4.13	t = 1.747 (245)	0.999
Speech-Shaped, High, Co-located (NH vs. HL)	9.047	4.13	t = 2.189 (245)	0.532
Two-Talker, Low, +90 (NH vs. HL)	18.113	4.13	t = 4.383 (245)	**<0.001**
Two-Talker, Low, +90/-90 (NH vs. HL)	12.357	4.13	t = 2.990 (245)	0.055
Two-Talker, Low, Co-located (NH vs. HL)	16.541	4.13	t = 3.00 (245)	0.054
Two-Talker, Moderate, +90 (NH vs. HL)	23.386	4.13	t = 5.659 (245)	**<0.001**
Two-Talker, Moderate, +90/-90 (NH vs. HL)	15.713	4.13	t = 3.802 (245)	**<0.01**
Two-Talker, Moderate, Co-located (NH vs. HL)	11.895	4.13	t = 2.878 (245)	0.078
Two-Talker, High, +90 (NH vs. HL)	21.754	4.13	t = 5.264 (245)	**<0.001**
Two-Talker, High, +90/-90 (NH vs. HL)	21.812	4.13	t = 5.278 (245)	**<0.001**
Two-Talker, High, Co-located (NH vs. HL)	7.612	4.13	t = 1.842 (245)	0.999
Age × Group × Reverberation Time × Masker Type × Spatial Location	-	-	F = 0.438 (1122)	0.781

A.3.3: Pearson correlation matrix for predictor variables included in the hierarchical linear regression analyses.

	Age	**Aided PTA**	**ACD Response Sensitivity**
Age	1		
Aided PTA	-0.09	1	
ACD Response Sensitivity	0.45***	-0.54***	1