An Investigation of Spatial Hearing in Children with Normal Hearing and with Cochlear Implants and the Impact of Executive Function

By

Sara M. Misurelli

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

(Communication Sciences and Disorders)

at the

UNIVERSITY OF WISCONSIN-MADISON

2014

Date of final oral examination: 4/28/2014

This dissertation is approved by the following members of the Final Oral Committee:
Ruth Litovsky, Professor, Communication Sciences and Disorders
Jan Edwards, Professor, Communication Sciences and Disorders
Margaritis Fourakis, Professor, Communication Sciences and Disorders
Margarita Kaushanskaya, Associate Professor, Communication Sciences and Disorders
Vanessa Simmering, Assistant Professor, Psychology

AN INVESTIGATION OF SPATIAL HEARING IN CHILDREN WITH NORMAL HEARING AND WITH COCHLEAR IMPLANTS AND THE IMPACT OF EXECUTIVE FUNCTION

Sara M. Misurelli

Under the supervision of Professor Ruth Y. Litovsky

At the University of Wisconsin-Madison

The ability to analyze an “auditory scene”—that is, to selectively attend to a target source while simultaneously segregating and ignoring distracting information—is one of the most important and complex skills utilized by normal hearing (NH) adults. The NH adult auditory system and brain work rather well to segregate auditory sources in adverse environments. However, for some children and individuals with hearing loss, selectively attending to one source in noisy environments can be extremely challenging. In a normal auditory system, information arriving at each ear is integrated, and thus these binaural cues aid in speech understanding in noise.

A growing number of individuals who are deaf now receive cochlear implants (CIs), which supply hearing through electrical stimulation to the auditory nerve. In particular, bilateral cochlear implants (BICIs) are now becoming more prevalent, especially in children. However, because CI sound processing lacks both fine structure cues and coordination between stimulation at the two ears, binaural cues may either be absent or inconsistent. For children with NH and with BiCIs, this difficulty in segregating sources is of particular concern because their learning and development commonly occurs within the context of complex auditory environments.

This dissertation intends to explore and understand the ability of children with NH and with BiCIs to function in everyday noisy environments. The goals of this work are to (1)

Investigate source segregation abilities in children with NH and with BiCIs; (2) Examine the effect of target-interferer similarity and the benefits of source segregation for children with NH and with BiCIs; (3) Investigate measures of executive function that may predict performance in complex and realistic auditory tasks of source segregation for listeners with NH; and (4) Examine source segregation abilities in NH listeners, from school-age to adults.

DEDICATION

"Character cannot be developed in ease and quiet. Only through experience of trial and suffering can the soul be strengthened, ambition inspired, and success achieved." Helen Keller

To my family, friends, and mentors who have continually encouraged me in times when I have doubted myself, inspired me to follow my dreams when they have seemed out of reach, and loved me unconditionally throughout it all.

Special thanks:

To my husband, Shaun Hernandez, who has supported me each step of the way, sacrificed to allow me to pursue this work, continually discussed these projects, and read countless drafts of this document. Thank you for always being my "rock," through the good and bad, without complaint, hesitation, or judgment. You challenge me to think critically and strive for success. This dissertation belongs to you as much as it does to me.

To my parents, Eugene and Bonnie Misurelli, who taught me the value of hard work, gave me confidence to believe in myself, and have instilled in me the importance of being kind above all else. Without your encouragement and support I would not be who I am today.

To my father-in-law and mother-in-law, Carlos and Leslie Hernandez, who have graciously shared their son with me and continually supported both of us along the way.

To my grandparents, here and in heaven, who have taught me the importance of family and appreciation of non-material things.

To my many close and dear friends, who have tolerated my obsessive personality and persuaded me to relax, understood when I needed to cancel plans at the last minute, and who have provided me with encouragement on days that I thought would never end.

ACKNOWLEDGMENTS

"If I have seen further, it is by standing upon the shoulders of giants." Sir Issac Newton

A special thanks to my advisor and mentor, Ruth Litovsky. You have pushed me to strive beyond what I thought was possible, and served as a role model—both personally and professionally. Thank you for having confidence in me and providing me with means to grow as a researcher, as well as lessons that allowed me to grow as a person.

Thank you to Jan Edwards, Marios Fourakis, Rita Kaushanskaya, and Vanessa Simmering, the members of my dissertation committee. You have all given me endless support, encouragement, and provided me with invaluable suggestions that are included in this work.

Thank you to the Department of Communication Sciences and Disorders for giving me the opportunity to learn from intelligent models of success within the department faculty. It is because of your welcoming and kindness that I decided to transfer to UW-Madison for my undergraduate career, and ultimately decided to stay to pursue a graduate degree. You make it feel like home. A special thanks to John Westbury for riding his bike to the Red Gym on a hot summer day to reassure my family and me that transferring to UW-Madison and entering the department was a good fit for me.

Thank you to the members of the Binaural Hearing and Speech Laboratory who have challenged me to excel to my fullest potential, tolerated many long conversations about these projects, helped program experiments and collect data, and been models of blossoming researchers. I am filled with gratitude to say that I can call my co-workers my friends.

Lastly, thank you to the many families that have dedicated their time to be a part of this research. Without you this work would not exist.

This research was supported by grants from the NIH-NIDCD (R01DC008365, Litovsky), in part by a core grant to the Waisman Center from the NIDCD (P30 HD03352), and in part by the Emma Allen Fellowship Award (awarded by the UW-Madison Department of Communication Sciences and Disorders).

TABLE OF CONTENTS

Chapter 1: Background and Introduction

Auditory scene analysis is the process by which the human auditory system segregates and groups sounds into various auditory streams based on the relationship of timing and frequency of the incoming signals [1]. The listener must be able to differentiate between multiple incoming signals in this complex auditory scene and ultimately determine elements to which to attend. In addition to selecting the elements of interest and streaming them, the listener must maintain attention as background sources change.

Selectively attending to a target source while simultaneously segregating and ignoring distracting information is one of the most important and complex skills utilized in everyday noisy situations; the ability to segregate sources is essential to successful communication [2-4]. The normal adult auditory system works rather well to segregate target and interfering auditory sources; however, for some children and people with hearing loss, selectively attending to target sound sources in noisy environments can be extremely challenging given that these populations require a larger signal-to-noise ratio (SNR, the level of the target signal relative to the level of the interferer) than normal hearing (NH) adults [5-9]. For children especially, source segregation is vital in facilitating the ability of children to learn and communicate in typical environments. They spend a majority of their time in these spaces, including classrooms, cafeterias, and playgrounds, and in the studies described here, some aspects of these auditory environments were simulated.

The goals of the current paper are to: (1) Investigate source segregation abilities in children with normal hearing (NH) and with bilateral cochlear implants (BiCIs); (2) Examine the effect of target-interferer similarity and the benefits of source segregation for children with NH and with BiCIs; (3) Investigate measures of executive function that may predict performance in complex and realistic auditory tasks of source segregation for listeners with NH; and (4) Examine source segregation abilities in NH listeners, from school-age children to adults.

Spatial Release from Masking

Adults and (more often) children work, learn, and interact in auditory environments in which there is one target talker and numerous distracting auditory sources, a situation that is commonly referred to as the "cocktail party environment" [10]. If interferers decrease the audibility of the target, perceptual auditory "masking" of the target source has occurred. When a target signal is embedded in background speech and/or noise, two types of auditory masking may occur: "energetic" and/or "informational."

Energetic masking is thought to occur at the level of the peripheral auditory system when auditory stimuli contain energy that overlaps, spectrally and temporally, within identical critical bands at the same time [11, 12]. Auditory filter models, which describe masking at the peripheral level, typically account for the amount of energetic masking that is observed experimentally.

Informational masking is a term that refers to multiple phenomena, and is less strictly defined than energetic masking. It is thought to be mediated by auditory and non-auditory mechanisms beyond the auditory periphery. Given that central auditory mechanisms are involved, information masking contributes to a perceptual challenge when listeners attempt to segregate target and interfering sources. For example, while both the target and interferer are audible to the listener, they cannot be easily segregated. Informational masking is characterized by confusability between the target and interfering stimuli, and occurs due to uncertainty in the target stimulus and/or similarity between the target and interfering stimuli [13-17]. It has been shown that more informational masking is observed when the target and interferer are spectrally similar [17-19]. For example, in 2001, Brungart and colleagues conducted a study with NH adults investigating the effect of target-interferer similarity. In his study, participants listened to speech stimuli over headphones and had to identify the target speech amongst various interferers (modulated noise, Gaussian noise, different sex-talker, same-sex talker, and same-talker). They found that with the noise interferers, intelligibility of the target did not decrease until the noise became loud enough to mask the energy in the target signal. Once the target signal was masked, performance systematically declined as the SNR became less favorable. This masking could be accounted for primarily by energetic masking. A different pattern of performance was demonstrated with the speech interferers, where it was shown that performance gradually decreased until a certain SNR and stayed stable thereafter. In addition, performance for the same SNR with different interferers was shown to be increasingly worse as the interferer became more similar to the target (i.e. same-talker interferer), suggesting that more informational masking is exhibited with more similar target and interferer stimuli. Chapter three of this dissertation explores the effect of target-interferer similarity and the benefits of source segregation in children.

There are various cues that are useful to help alleviate masking, such as location cues (e.g. spatial separation of the target and interfering sources), visual cues (e.g. lip-reading and facial expressions of the target talker), individual talker characteristics (e.g. pitch of voice and rate of speech), and transitions (e.g. knowledge of the subject matter and voice dynamics) [10]. Location cues will be the main focus of this paper and discussed in all three studies.

The amount of masking, in part, depends on the perceived spatial separation of the target and the interferer [20, 21]. That is, in a realistic environment, the closer in proximity the interferer is to the target source, the more difficult the target is to segregate [22]. Performance in noisy environments often improves when sources are spatially separated. Specifically, the target source can be correctly identified at a higher rate of percent correct, or lower SNR, when the target and interferers are spatially separated relative to when they are co-located. This improvement in performance due to spatial separation of sources is known as spatial release from masking (SRM).

SRM is usually the greatest for interfering sources that have a large physical separation from the target in a room. This dissertation describes studies that use asymmetrical placement of target and interferers (target at 0° and interferers directed to only one ear at 90°), and the less commonly used, symmetrically distributed interferers (target at 0° and interferers directed to both the right and left ears at ±90°). With asymmetrical placement of interferers there is an improved SNR at the ear without interferers, and therefore listeners are able to use the monaural head shadow cue to help segregate sources [23]. The symmetrical distribution of interferers is meant to reduce the benefit from the monaural head shadow and challenge listeners to rely mainly on binaural cues for source segregation. Although monaural cues are reduced in the symmetrical condition, there may be transient glimpses in which one ear will receive a better SNR than the other, resulting in some better ear effects [24, 25]. More SRM is demonstrated when interferers are distributed asymmetrically rather than symmetrically due to the availability of a more prominent head shadow cue [26, 27].

Spatial separation of sources has been shown to be advantageous, particularly in environments in which listeners demonstrate large amounts of informational masking [21, 28]. Previous work has shown that when the target and interferers are both speech-based stimuli, rather than when the target is speech and the interferers are noise, spatial cues are more robust [16, 18-21, 29-32]. Furthermore, when the target and interferers are both speech-based, the amount of SRM depends on the number of interferers [18, 29].

Arbogast *et al.* [16] examined the effects of energetic versus informational masking on SRM for NH listeners, using stimuli designed to better control the type of masking. The stimuli were divided into a limited number of frequency bands, similar to that of a cochlear implant (CI). The target sentences were generated by using eight of the fifteen frequency bands, and the interferer was either (1) a sentence derived from 6 different frequency bands than the target sentence, intended to create primarily informational masking due to the minimal amount of spectral interaction with the target; (2) noise derived from different frequency bands than the target sentence, intended as an energetic masking control for the different-band sentences; or (3) noise derived from the same eight frequency bands as the target sentence, intended to create primarily energetic masking. The authors concluded that the most SRM was demonstrated with the different-band sentence interferer, suggesting that the listeners demonstrate the most advantage of spatially separated sources for maskers that are primarily informational in nature than energetic. These same results were found, although to a lesser degree, for listeners with sensori-neural hearing loss [32]. This result suggests that listeners who are hearing-impaired are able to use cues of spatial separation to benefit speech understanding when listening in noise.

Source Segregation in Listeners with Normal Hearing

People with NH have two ears that work together to integrate auditory input from each ear; this is referred to as binaural hearing. Using cues derived from binaural processing, in complex auditory environments adult NH listeners are able to segregate target and interfering sources and function relatively well. When the target and interferers are spatially separated, NH listeners are able to benefit from both monaural head shadow and binaural squelch. As described above, the monaural head shadow effect, also known as the "better ear effect," occurs when the target and interfering stimuli are spatially separated and the ear without the interferers has a more favorable SNR due to the interference of the listener's head. "Binaural squelch" occurs when there is an improvement in performance when the ear with the poorer SNR is added because the auditory system uses the differences in the timing and level of the target and interferers between the two ears [33]. The third effect that NH listeners are able to use to enhance speech perception in noise is "binaural summation," an effective binaural cue when a target signal is presented from directly in front of a listener. When the target is presented at 0° azimuth, the signal reaches both ears at the same time and intensity and the overall signal is summed between the two ears; and therefore amplified. A normal healthy auditory system encompasses binaural processing which enables NH listeners to function in noisy environments, in part, due to the above-mentioned cues. Both squelch and summation require binaural hearing. Because the monaural head shadow effect is not a true binaural cue, it has also been shown to be beneficial to some extent to those without binaural hearing (e.g., BiCI users) [34-37]. This will be discussed in the "Source Segregation in Listeners with Bilateral Cochlear Implants" section of this introduction.

NH listeners readily show improvement in speech understanding when auditory sources are spatially separated. SRM can be as large as 12-15 dB in NH adults [38]. NH children also demonstrate SRM but to a smaller and more variable extent than adults [5, 6, 8, 27, 28]. There is evidence to suggest that when parts of the acoustic signal are manipulated over headphones, children as young as 4 years of age are able to discriminate timing differences between the ears to the same extent as adults [39], as well as differentiate out-of-phase tones more accurately than in-phase tones (binaural masking level differences) [40]. In free-field spatial unmasking tasks, NH children have also shown evidence of the ability to take advantage of spatial cues. SRM has been observed in children ages 3-4 [41], and more recently in toddlers 2-3 years of age [42]. These new findings confirm that in general NH children as young as 2 years of age show evidence of binaural hearing and are able to benefit from segregation of auditory sources, although the degree of individual benefit remains variable.

Source Segregation in Listeners with Bilateral Cochlear Implants (BICIs)

Listeners with CIs often have a great deal of difficulty understanding speech in noisy environments. It has been shown that both adults and children with BICIs often outperform those with only one CI on tasks of sound localization and speech understanding in noise [9, 43-48]. Yet, there remains a gap in performance between listeners with BiCIs and with NH, even when matched for the amount of auditory experience [27]. This gap may be due to a number of factors, including limitations of the device (e.g., lack of binaural cues, degraded spectral cues, etc.), demographic factors (e.g. age at implantation, age at deafness, amount of auditory experience, etc.), and anatomical factors (e.g. intact cochlear mechanisms, auditory nerve survival, active auditory brainstem synapses). The factors are further described in the following paragraphs.

The currently available clinical CI devices are not coordinated, and therefore BiCI users lack the integration of cues that arrive at each ear. The lack of binaural cues available to the CI user likely limits their ability to perform in multi-source environments. Previous research has confirmed that CI users rely most heavily on the monaural head shadow cue, rather than binaural cues (i.e., summation or squelch), when listening in noisy environments [34, 37, 40, 49, 50]. The speech processing strategies are limited in that the device extracts the envelope of the signal, ignoring fine structure, and delivers the signal to a limited number of spectral channels. The signal that the user receives is degraded both spectrally and temporally [29]. Also, microphone placement above the pinna, compression functions, and directional settings alter the monaural and interaural level directional cues and therefore users cannot access the full signal. Variable surgical placement of the electrode array may also affect performance outcomes. The ideal placement of the array is in the scala tympani closest to the modiolus, which is proximal to the auditory nerve dendrites. Any variability in placement may cause a mismatch of frequencies being delivered to various frequency ranges on the basilar membrane of each ear and therefore may limit binaural sensitivity [51]. More work is needed to verify how much variability in performance between individual users is due to differences in electrode placement. Electrodes may be re-inserted into the cochlea if necessary, although it is neither ideal nor standard practice to undergo another surgery.

Peripheral and central auditory factors also likely play a role in user outcome. It has been shown that spiral ganglion cells within the auditory system begin to degrade after a period of auditory deprivation [52]. Previous work with animal models has also shown that the auditory cortex begins to reorganize when no there is no auditory stimulation [53, 54].

Electrophysiological studies in humans suggest that binaural auditory pathways may degrade and not be able to repair themselves if bilateral stimulation is not received within a certain period of time [55].

In recent years, researchers have begun to investigate the effects of higher-level cognition and linguistic processing on speech understanding. There have been numerous studies investigating auditory and cognitive processes that may have an effect on performance speech in noise tasks for older adults with and without hearing impairments. As a whole, the studies have concluded that there is an interaction between cognitive and peripheral mechanisms. They have shown that the predictive factors of ability to function in complex auditory environments are foremost accounted for by degree of hearing loss, but can secondarily be accounted for by various measures of cognitive function [56-58]. It has also been shown that children with CIs demonstrate a deficit in some areas of executive function (working memory, aspects of attention, processing speed) compared to their NH peers [59-62]. Although working memory is not a direct measure of the ability to hear speech in noise, in order to function in noisy environments listeners must demonstrate capacity to attend to, retain, and manipulate incoming information in challenging situations. Numerous studies have investigated variables, such as age at implantation, age at deafness, device type, and etiologies to help explain the variability in speech and language performance for children with CIs. Even when all of the above mentioned demographic factors are accounted for, a significant amount of variability in performance outcomes remains [63, 64]. A study by Figueras *et al.* [65] compared 8-12 year old children with NH and with CIs on measures of working memory, attention, and language to determine if differences existed between the two groups. Difference in performance on some measures of attention and working memory were found, although the differences disappeared when language ability was accounted for. Other studies have shown that tests of working and short-term memory are highly correlated with language outcomes [61, 64, 66]. Deficits in executive functioning may impact the performance on real-world auditory tasks requiring organization, inhibition, and planning.

In summary, children with BiCIs seem to perform better in noisy environments than those who use only one CI and a hearing aid in the other ear [34, 67], but BiCI users lack comparable performance to NH listeners. Chapters 2 and 3 of this paper compare children with NH and with BiCIs, matching according to the amount of auditory experience, in attempt to further explore the differences between hearing types. A great deal of work is being done to improve bottom-up signal processing as well as isolate top-down processes that are involved in speech understanding in noise, to ultimately increase success and ease of listening for people with hearing loss.

Effect of Executive Function on Variability in Performance on Speech-in-Noise Tasks

After conducting the studies described in the second and third chapter of this dissertation, as well as comparing with previous research in the lab, we discovered a large amount of variability in the benefit from spatial cues within each subject group, for both children with NH and with BiCIs [9, 27, 41, 68]. Despite this large within-group variability, individual participant performance was confirmed to be stable throughout repeated trials [27]. This suggests that there are integral differences *between* subjects, rather than *within*. The variability shown in our BiCI child groups was consistent with previous research [68], and although interesting, this was not surprising. In our work in the lab we not only discovered variability within the BiCI groups, but our data also revealed that children with NH showed variable performance on tasks of source segregation. This is interesting because the peripheral system of NH listeners is intact and functional at birth, and therefore variability cannot be attributed to any of the previously discussed peripheral factors. Infants and children have been described as inefficient listeners compared to adults, even though the peripheral auditory system of children is thought to accurately represent sound at a very young age [69, 70]. Peripheral processing of intensity, frequency and time are adult-like by six months of age, and therefore inefficiency in listening is likely influenced by maturation of perceptual organization, as well as a variety of central factors (e.g. memory and attention) [3, 4]. It remains unknown as to why children of different age groups vary, and at what age performance becomes adult-like.

Auditory filter models have commonly focused on how the peripheral auditory system organizes and filters peripheral sensory information. Models have explored explanations for differences in the ability of children and of adults to successfully identify target stimuli in complex listening tasks. Results have shown that the peripheral auditory system, intact at birth [71], organizes sensory information in similar ways for both NH children and adults. Yet, it is known why adults outperform children on identical speech-in-noise tasks. These results suggest that some difference between child and adult listener groups originates in mechanisms beyond the peripheral structure [7]. Previous work fosters the idea that once the sensory input is filtered, children integrate information over a larger number of auditory filters than adults, often including filters that do not contain the target stimuli (i.e., children include unnecessary auditory information when extracting a target signal from background noise) [13]. Because, in general, children include unnecessary filters when processing speech in noise, they have more difficultly than adults extracting the target signal [5, 6, 8, 9, 27]. When listening to speech in noise, it remains unclear why some children include unnecessary filters and others do not.

While a number of studies have focused on peripheral factors that influence performance outcomes, they have failed to identify any specific processes that explain why some NH children are better able to segregate auditory sources than others, and why there is variability between different age groups. It may be that processes beyond the periphery can help to better explain some variability when listening to speech in noise. To better understand the underlying variability in performance on speech-in-noise tasks, the fourth chapter in this dissertation will investigate measures of executive function that are hypothesized to correlate with auditory processing and source segregation. “Executive function” is an umbrella term that can refer to a multitude of abilities, including attention, working memory, planning, and inhibition [72]. Overall, executive function encompasses the organizational abilities that are required for goal-directed, purposeful, and non-automatic behavior. Specific components of executive function such as working memory, short-term memory, and attention are hypothesized to be involved in the processing of complex auditory stimuli. Some of the above-mentioned factors may help to better understand the variability in NH listeners on tasks of source segregation, and may serve as predictors of individual performance.

In order to process and make sense of complex and realistic auditory stimuli when listening to speech in noise, the listener must continually monitor incoming sources and integrate a string of individual words with prior knowledge. This process requires both working memory and short-term memory [61, 73]. While short-term memory is an active memory for storing incoming information, working memory encompasses the processes that are used to temporarily store, manipulate, and integrate incoming information [74, 75]. Working memory has been shown to have a central role in information processing, in that it is vital in maintaining information for further processing [76], and is critical for the development and improvement of speech and language skills [77]. Processing of complex auditory stimuli puts high demands on working memory capacity. When working memory capacity is exceeded, auditory processing is slowed or errors in accurate comprehension are made [66, 73, 78].

Attention also likely plays a role in ability to function in complex auditory environments. Selective attention is a component of executive function that helps to determine important sounds and bring attentiveness to them. How well one is able to segregate and attend to an auditory target will in turn determine how well distracting auditory information can be ignored [79-81]. This skill becomes vital when trying to communicate in noisy environments. Recent neuroimaging studies have confirmed that attention to an auditory stream changes patterns of cortical activity (for review see Lee, Larson [81]).

Working memory and attention may work together in adverse listening conditions in order for listeners to be able to accurately and efficiently process verbal information. These processes allow the listener to maintain, process, and suppress information, such that target stimuli can be attended to while simultaneously ignoring distracters [82, 83]. There have been numerous studies investigating audibility (i.e. hearing thresholds) and performance on tasks of executive function that may have an effect on performance speech-in-noise tasks for older adults with and without hearing impairments. As a whole, the studies have concluded that there is an interaction between central and peripheral processes. They have shown that the predictive factors of ability to function in complex auditory environments are foremost accounted for by degree of hearing loss, but can secondarily be accounted for by various measures of executive function, such as processing speed, reasoning, memory, intelligence, attention and fluency [56-58]. It remains unclear how spatial hearing relates to executive function for a larger age-range of children with normal hearing (NH). This topic is explored in chapter four.

In summary, this dissertation intends to explore and understand the ability of children with NH and who are deaf and use BiCIs to function in everyday noisy environments. The second and third chapters aim to evaluate performance of children with NH and with BiCIs on speech-in-noise tasks when altering factors of target-interferer similarity and interferer location. The fourth chapter aims to isolate the variability in the group of children with NH, and investigate how aspects of executive function may serve as predictors of performance on tasks of source segregation. Also in Chapter 4, we will examine source segregation abilities in children with NH that can then be used to compare to children with hearing impairments or developmental disabilities. Predictors of performance in spatial hearing tasks may, in the future, help to inform teachers of ideal classroom arrangements in order to enhance SNR, assist in clinical decisions as to what devices to fit, and to help predict success and intervention needs for children with hearing impairments.

I. REFERENCES

1. Bregman, A.S., *Auditory scene analysis the perceptual organization of sound*. 1990, MIT Press,: Cambridge, Mass. p. xiii, 773 p.

2. Shinn-Cunningham, B.G. and V. Best, *Selective attention in normal and impaired hearing.* Trends Amplif, 2008. **12**(4): p. 283-99.

3. Leibold, L.J., *Development of Auditory Scene Analysis and Auditory Attention.* Human Auditory Development, 2012. **42**: p. 137-161.

4. Werner, L., R.R. Fay, and A.N. Popper, *Human Auditory Development*, in *Springer Handbook of Auditory Research, 42*. 2012, Springer,: Dordrecht. p. 1 online resource (293 p.).

5. Wightman, F.L., et al., *Children's detection of pure-tone signals: informational masking with contralateral maskers.* J Acoust Soc Am, 2003. **113**(6): p. 3297-305.

6. Wightman, F.L. and D.J. Kistler, *Informational masking of speech in children: effects of ipsilateral and contralateral distracters.* J Acoust Soc Am, 2005. **118**(5): p. 3164-76.

7. Wightman, F., D. Kistler, and D. Brungart, *Informational masking of speech in children: auditory-visual integration.* J Acoust Soc Am, 2006. **119**(6): p. 3940-9.

8. Hall, J.W., 3rd, et al., *Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children.* Ear Hear, 2002. **23**(2): p. 159-65.

9. Litovsky, R.Y., *Speech intelligibility and spatial release from masking in young children.* J Acoust Soc Am, 2005. **117**(5): p. 3091-9.

10. E.C., C., *Some experiments on the recognition of speech, with one and with two ears.* J. Acoust. Soc. Am., 1953. **25**: p. 975-979.

11. Watson, C.S., W.J. Kelly, and H.W. Wroton, *Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty.* J Acoust Soc Am, 1976. **60**(5): p. 1176-86.

12. Neff, D.L. and T.M. Dethlefs, *Individual differences in simultaneous masking with random-frequency, multicomponent maskers.* J Acoust Soc Am, 1995. **98**(1): p. 125-34.

13. Lutfi, R.A., et al., *One factor underlies individual differences in auditory informational masking within and across age groups.* Percept Psychophys, 2003. **65**(3): p. 396-406.

14. Watson, C.S., *Some comments on informational masking.* Acta Acustica United with Acustica, 2005. **91**(3): p. 502-512.

15. Durlach, N.I., et al., *Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity.* J Acoust Soc Am, 2003. **114**(1): p. 368-79.

16. Arbogast, T.L., C.R. Mason, and G. Kidd, Jr., *The effect of spatial separation on informational and energetic masking of speech.* J Acoust Soc Am, 2002. **112**(5 Pt 1): p. 2086-98.

17. Brungart, D.S., et al., *Informational and energetic masking effects in the perception of multiple simultaneous talkers.* J Acoust Soc Am, 2001. **110**(5 Pt 1): p. 2527-38.

18. Hawley, M.L., R.Y. Litovsky, and J.F. Culling, *The benefit of binaural hearing in a cocktail party: effect of location and type of interferer.* J Acoust Soc Am, 2004. **115**(2): p. 833-43.

19. Cullington, H.E. and F.G. Zeng, *Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects.* J Acoust Soc Am, 2008. **123**(1): p. 450-61.

20. Freyman, R.L., U. Balakrishnan, and K.S. Helfer, *Spatial release from informational masking in speech recognition.* J Acoust Soc Am, 2001. **109**(5 Pt 1): p. 2112-22.

21. Freyman, R.L., et al., *The role of perceived spatial separation in the unmasking of speech.* J Acoust Soc Am, 1999. **106**(6): p. 3578-88.

22. Marrone, N., C.R. Mason, and G. Kidd, *Tuning in the spatial dimension: evidence from a masked speech identification task.* J Acoust Soc Am, 2008. **124**(2): p. 1146-58.

23. Zurek, P.M., *A note on onset effects in binaural hearing.* J Acoust Soc Am, 1993. **93**(2): p. 1200-1.

24. Lavandier, M., et al., *Binaural prediction of speech intelligibility in reverberant rooms with multiple noise sources.* J Acoust Soc Am, 2012. **131**(1): p. 218-31.

25. Jelfs, S., J.F. Culling, and M. Lavandier, *Revision and validation of a binaural model for speech intelligibility in noise.* Hear Res, 2011. **275**(1-2): p. 96-104.

26. Jones and R.Y. Litovsky, *A cocktail party model of spatial release from masking by both noise and speech interferers.* J Acoust Soc Am, 2011. **130**(3): p. 1463-74.

27. Misurelli, S.M. and R.Y. Litovsky, *Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry.* J Acoust Soc Am, 2012. **132**(1): p. 380-91.

28. Kidd, G., Jr., et al., *Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns.* J Acoust Soc Am, 1998. **104**(1): p. 422-31.

29. Culling, J.F., M.L. Hawley, and R.Y. Litovsky, *The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources.* J Acoust Soc Am, 2004. **116**(2): p. 1057-65.

30. Brungart, D.S., *Informational and energetic masking effects in the perception of two simultaneous talkers.* J Acoust Soc Am, 2001. **109**(3): p. 1101-9.

31. Jones, J.A. and R.L. Freyman, *Effect of priming on energetic and informational masking in a same-different task.* Ear Hear, 2012. **33**(1): p. 124-33.

32. Arbogast, T.L., C.R. Mason, and G. Kidd, Jr., *The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners.* J Acoust Soc Am, 2005. **117**(4 Pt 1): p. 2169-80.

33. Drullman, R. and A.W. Bronkhorst, *Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation.* J Acoust Soc Am, 2000. **107**(4): p. 2224-35.

34. Litovsky, R., et al., *Simultaneous bilateral cochlear implantation in adults: a multicenter clinical study.* Ear Hear, 2006. **27**(6): p. 714-31.

35. Litovsky, R.Y., *Review of recent work on spatial hearing skills in children with bilateral cochlear implants.* Cochlear Implants Int, 2011. **12 Suppl 1**: p. S30-4.

36. van Hoesel, R.J. and R.Y. Litovsky, *Statistical bias in the assessment of binaural benefit relative to the better ear.* J Acoust Soc Am, 2011. **130**(6): p. 4082-8.

37. Loizou, P.C., et al., *Speech recognition by bilateral cochlear implant users in a cocktail-party setting.* J Acoust Soc Am, 2009. **125**(1): p. 372-83.

38. Jones, G.L. and R.Y. Litovsky, *Role of masker predictability in the cocktail party problem.* J Acoust Soc Am, 2008. **124**(6): p. 3818-30.

39. Ashmead, D.H., et al., *Sound localization and sensitivity to interaural time differences in human infants.* Child Dev, 1991. **62**(6): p. 1211-26.

40. Van Deun, L., et al., *Sound localization, sound lateralization, and binaural masking level differences in young children with normal hearing.* Ear Hear, 2009. **30**(2): p. 178-90.

41. Garadat, S.N. and R.Y. Litovsky, *Speech intelligibility in free field: spatial unmasking in preschool children.* J Acoust Soc Am, 2007. **121**(2): p. 1047-55.

42. Hess, C.L., *Speech discrimination and spatial hearing in toddlers with bilateral cochlar implants*. 1 online resource (iv, 110 pages).

43. Godar, S.P. and R.Y. Litovsky, *Experience with bilateral cochlear implants improves sound localization acuity in children.* Otol Neurotol, 2010. **31**(8): p. 1287-92.

44. Grieco-Calub, T.M. and R.Y. Litovsky, *Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing.* Ear Hear, 2010. **31**(5): p. 645-56.

45. Litovsky, R.Y., et al., *Bilateral cochlear implants in adults and children.* Arch Otolaryngol Head Neck Surg, 2004. **130**(5): p. 648-55.

46. Gantz, B.J., et al., *Binaural cochlear implants placed during the same operation.* Otol Neurotol, 2002. **23**(2): p. 169-80.

47. Tyler, R.S., et al., *Patients utilizing a hearing aid and a cochlear implant: speech perception and localization.* Ear Hear, 2002. **23**(2): p. 98-105.

48. Muller, J., F. Schon, and J. Helms, *Speech understanding in quiet and noise in bilateral users of the MED-EL COMBI 40/40+ cochlear implant system.* Ear Hear, 2002. **23**(3): p. 198-206.

49. Litovsky, R.Y., A. Parkinson, and J. Arcaroli, *Spatial hearing and speech intelligibility in bilateral cochlear implant users.* Ear Hear, 2009. **30**(4): p. 419-31.

50. van Hoesel, R., et al., *Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies.* J Acoust Soc Am, 2008. **123**(4): p. 2249-63.

51. van Hoesel, R.J., *Exploring the benefits of bilateral cochlear implants.* Audiol Neurootol, 2004. **9**(4): p. 234-46.

52. Leake, P.A., G.T. Hradek, and R.L. Snyder, *Chronic electrical stimulation by a cochlear implant promotes survival of spiral ganglion neurons after neonatal deafness.* J Comp Neurol, 1999. **412**(4): p. 543-62.

53. Kral, A., et al., *Spatiotemporal patterns of cortical activity with bilateral cochlear implants in congenital deafness.* J Neurosci, 2009. **29**(3): p. 811-27.

54. Kral, A., *Unimodal and cross-modal plasticity in the 'deaf' auditory cortex.* Int J Audiol, 2007. **46**(9): p. 479-93.

55. Gordon, K.A., S. Jiwani, and B.C. Papsin, *Benefits and detriments of unilateral cochlear implant use on bilateral auditory development in children who are deaf.* Front Psychol, 2013. **4**: p. 719.

56. Lindenberger, U. and P.B. Baltes, *Sensory functioning and intelligence in old age: a strong connection.* Psychol Aging, 1994. **9**(3): p. 339-55.

57. Humes, L.E., *Factors underlying the speech-recognition performance of elderly hearing-aid wearers.* J Acoust Soc Am, 2002. **112**(3 Pt 1): p. 1112-32.

58. Pichora-Fuller, M.K., *Cognitive aging and auditory information processing.* Int J Audiol, 2003. **42 Suppl 2**: p. 2S26-32.

59. Burkholder, R.A. and D.B. Pisoni, *Speech timing and working memory in profoundly deaf children after cochlear implantation.* J Exp Child Psychol, 2003. **85**(1): p. 63-88.

60. Cleary, M., D.B. Pisoni, and A.E. Geers, *Some measures of verbal and spatial working memory in eight- and nine-year-old hearing-impaired children with cochlear implants.* Ear Hear, 2001. **22**(5): p. 395-411.

61. Pisoni, D.B., et al., *Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation.* Ear Hear, 2011. **32**(1 Suppl): p. 60S-74S.

62. Beer, J., D.B. Pisoni, and W. Kronenberger, *Executive Function in Children with Cochlear Implants: The Role of Organizational-Integrative Processes.* Volta Voices, 2009. **16**(3): p. 18-21.

63. Geers, A., C. Brenner, and L. Davidson, *Factors associated with development of speech perception skills in children implanted by age five.* Ear Hear, 2003. **24**(1 Suppl): p. 24S-35S.

64. Pisoni, D.B. and M. Cleary, *Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation.* Ear Hear, 2003. **24**(1 Suppl): p. 106S-20S.

65. Figueras, B., L. Edwards, and D. Langdon, *Executive function and language in deaf children.* J Deaf Stud Deaf Educ, 2008. **13**(3): p. 362-77.

66. Conway, C.M., D.B. Pisoni, and W.G. Kronenberger, *The Importance of Sound for Cognitive Sequencing Abilities: The Auditory Scaffolding Hypothesis.* Curr Dir Psychol Sci, 2009. **18**(5): p. 275-279.

67. Mok, M., et al., *Speech perception benefit for children with a cochlear implant and a hearing aid in opposite ears and children with bilateral cochlear implants.* Audiol Neurootol, 2010. **15**(1): p. 44-56.

68. Litovsky, R.Y., P.M. Johnstone, and S.P. Godar, *Benefits of bilateral cochlear implants and/or hearing aids in children.* Int J Audiol, 2006. **45 Suppl 1**: p. S78-91.

69. Werner, L.A. and K. Boike, *Infants' sensitivity to broadband noise.* J Acoust Soc Am, 2001. **109**(5 Pt 1): p. 2103-11.

70. Hartley, D.E., et al., *Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children.* J Speech Lang Hear Res, 2000. **43**(6): p. 1402-15.

71. Boothroyd, A., *Auditory development of the hearing child.* Scand Audiol Suppl, 1997. **46**: p. 9-16.

72. Lyon, G.R. and N.A. Krasnegor, *Attention, memory, and executive function*. 1996, Baltimore: P.H. Brookes Pub. Co. xvii, 424 p.

73. Pichora-Fuller, M.K. and P.E. Souza, *Effects of aging on auditory processing of speech.* Int J Audiol, 2003. **42 Suppl 2**: p. 2S11-6.

74. Baddeley, A., *Modularity, mass-action and memory.* Q J Exp Psychol A, 1986. **38**(4): p. 527-33.

75. Baddeley, A.D. and J.D. Larsen, *The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative.* Q J Exp Psychol (Hove), 2007. **60**(4): p. 497-504.

76. Cowan, N., et al., *List composition and the word length effect in immediate recall: a comparison of localist and globalist assumptions.* Psychon Bull Rev, 2003. **10**(1): p. 74-9.

77. Peterson, N.R., D.B. Pisoni, and R.T. Miyamoto, *Cochlear implants and spoken language processing abilities: review and assessment of the literature.* Restorative neurology and neuroscience, 2010. **28**(2): p. 237-250.

78. Pisoni, D.B. and A.E. Geers, *Working memory in deaf children with cochlear implants: correlations between digit span and measures of spoken language processing.* Ann Otol Rhinol Laryngol Suppl, 2000. **185**: p. 92-3.

79. Knudsen, E.I., *Fundamental components of attention.* Annu Rev Neurosci, 2007. **30**: p. 57-78.

80. Shinn-Cunningham, B.G., *Object-based auditory and visual attention.* Trends Cogn Sci, 2008. **12**(5): p. 182-6.

81. Lee, A.K., et al., *Using neuroimaging to understand the cortical mechanisms of auditory selective attention.* Hear Res, 2014. **307**: p. 111-20.

82. Baddeley, A.D. and G.J. Hitch, *Working memory.* The psychology of learning and motivation, 1974. **8**: p. 47-89.

83. Engle, R.W., *Working memory capacity as executive attention.* Current directions in psychological science, 2002. **11**(1): p. 19-23.

Chapter 2: Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry (*published in J. Acoust. Soc. Am. 132(1): 380-391)

ABSTRACT

Spatial release from making (SRM) was measured in groups of children with bilateral cochlear implants (BiCIs), average ages 6.0 and 7.9 yrs, and with normal hearing (NH), average ages 5.0 and 7.8 yrs. Speech reception thresholds (SRTs) were measured for target speech in front (0°), and interferers either in front, distributed asymmetrically around the head (+90°/+90°) or distributed symmetrically around the head (+90°/-90°). In the asymmetrical condition both monaural "better ear" and binaural cues are available. In the symmetrical condition, listeners rely more on binaural cues to segregate sources. SRM was computed as the difference between SRTs in the front condition and SRTs in either the asymmetrical or symmetrical conditions. Results showed that asymmetrical SRM was smaller in BiCI users than NH children. Furthermore, NH children showed symmetrical SRM, suggesting they are able to use binaural cues for source segregation, whereas children with BiCIs had minimal symmetrical SRM. These findings suggest that children who receive bilateral cochlear implants can segregate speech from noise under conditions that maximize monaural "better ear" cues. It is likely device limitations play an important role in limiting SRM. Thus, improvement in spatial hearing abilities in children with BiCIs may require binaural processing strategies.

I. INTRODUCTION

Cochlear implants (CIs) provide hearing to people who are deaf through electrical stimulation to the auditory nerve. A growing number of recipients are now receiving CIs in both ears. The use of bilateral CIs (BiCIs) was clinically motivated, in part, by the fact that in normal-hearing (NH) listeners, binaural hearing plays an important role in facilitating sound localization and speech understanding in noise. In the BiCI literature, the emphasis to date has been primarily on the benefit arising when two CIs are used vs. when a single CI is used. Numerous studies to date have shown that adult BiCI users have significantly better sound localization abilities with BiCIs vs. a single CI (e.g. [1-3]). In addition, these patients generally show better speech recognition in noise when using BiCIs vs. a single CI [1, 3, 4]. The largest benefits are not due to coordinated stimulation in both ears; rather, they can typically be attributed to monaural "head shadow" (better-ear effect) cues. Binaural "summation" (redundancy in information with two ears) and the "squelch" effect, which require binaural integration of inputs to the two ears, are generally small and inconsistent in adult BiCI users [3, 5].

Adult BiCI users vary in their etiology, hearing experience and age at onset of deafness. Many patients studied to date will have spent years, if not decades, deprived of hearing between the time of hearing loss onset and implantation. Unlike adult BiCI users, the majority of children receiving CIs are being implanted at a young age, before the development of speech and language. Most of these children are born deaf and receive no benefit from auditory input until their CIs are activated. Although these children spend very little time deprived of hearing, most have very little exposure to sound prior to implantation. The present study thus focused on spatial unmasking of speech in children who use BiCIs. This population, like adults, generally appears to show significant benefits from two vs. one CI. For example, they perform better on measures of sound localization [6, 7] and on measures of spatial discrimination of target stimuli, such as minimum audible angle (MAA) [8-11]. Interestingly, however, resolution of spatial hearing in children with BiCIs is significantly worse than that of children with NH, as they show higher root-mean-square error on localization and higher MAAs on discrimination. The source of this gap in performance is not well understood.

The present study was concerned with measuring spatial unmasking in children who use BiCIs and in NH children, to determine whether a gap exists on this measure as well. All children who use BiCIs were recruited prospectively for a larger study, and had at least 1 year of experience with BiCIs. We focused on their bilateral listening mode, which is the natural everyday listening mode for BiCI users, rather than comparing with an unnatural (and usually un-preferred) unilateral listening mode. An area of focused interest in the NH literature is that of speech understanding when target speech is either spatially co-located with interfering sources (maskers/interferers), or spatially separated from interferers. In NH listeners, the improvement, or benefit, due to spatial separation, is known as spatial release from masking (SRM). SRM can be as large as 12-15 dB in NH adults and depends on numerous factors, including the: number of sources [12, 13], type of interfering sources [13, 14], perceived spatial separation of the sources [15, 16], and room acoustics [17, 18]. In the existing literature, one way to quantify SRM is to add effects that are due to (1) monaural head shadow which results in a more favorable signal-to-noise ratio (SNR) at the “better ear” and (2) binaural unmasking [13, 19]. This approach may be reasonably suitable for environments that contain only one interfering source from a single location. A different approach to quantifying SRM, which has been proven to be reliable for environments containing one interferer type that is simultaneously distributed in multiple locations, involves additive components that arise from (1) spatial separation of the target and interferer(s) and (2) spatial asymmetry of the interferers. This model was originally described for noise interferers and more recently modified to successfully account for effects arising in the presence of speech interferers similar to those used in the present study [20].

SRM is generally largest for conditions in which there are large angular separation between target and interferers. In adults with NH, Jones and Litovsky [20] measured speech reception thresholds (SRTs) for conditions with target at 0° (front) and two-talker interferers at either +90/-90° (two interferers symmetrically distributed, one talker on the right and one on the left) or +90/+90° (two interferers asymmetrically placed so that two talkers are co-located and both are on one side of the head). On average, the total amount of SRM (difference in SRT between spatially co-located and separated conditions) was 12 dB improvement in SRT. Of that 12 dB improvement, 8.7 dB was accounted for by spatial separation and 3.3 dB was accounted for by the effect of asymmetry, i.e., there was greater SRM in the +/+90° than in the +/-90° condition. The model does not provide definitive conclusions about whether the division of SRM into separation and asymmetry components has significant parallels with the division of SRM into better ear and binaural unmasking components. In particular, in the symmetrical condition, having independent speech interferers at +/-90° means that there are potential opportunities for momentary ‘glimpsing’ of information through one ear at a time so that some better ear cues might be available in the symmetrical condition. Thus, we consider the symmetrical condition as having reduced, but not absent better ear effects. In this condition, when better ear cues are minimized, listeners must rely on binaural cues or other cues such as pitch voice, for source segregation. In the binaural domain, the role of interaural time difference (ITD) has been shown to be more robust than the role of interaural level difference (ILD) cues in SRM [12].

Studies in children have been conducted with either one or two interferers, but with a standard spatial configuration of target and interferers whereby the latter are positioned asymmetrically relative to the head, i.e., interferers are located on one side of the head. SRM in NH children can range from 5-10 dB [21-24], and at least one report has shown that SRM in children ages 11-14 is around 13 dB [25]. Regarding children who use BiCIs, relatively little is known about spatial unmasking. Mok *et al.* [26] showed that children with BiCIs exhibit better speech perception in noise in spatially separated conditions than children who use a CI in one ear and a hearing aid in the opposite ear, suggesting advantages for BiCIs compared with unilateral CIs. Litovsky *et al.* [10] and Chadha *et al.* [27] reported that children who received BiCIs sequentially showed an improvement in speech understanding when interferers were shifted towards the side of their second CI, suggesting that children have a preference towards their first CI. Van Deun *et al.* [6] reported that it was advantageous for all of the children with BiCIs to listen with both CIs rather than one when listening to speech in noise. In the latter, as shown in adult studies (e.g., [1, 5]), it was suggested that the majority of the benefits for children with BiCIs are perceived due to better ear cues, but there was no evidence for spatial effects arising from binaural unmasking.

In the present study we were concerned with SRM in children with NH and in children who are deaf and fitted with bilateral cochlear implants (BiCIs). Regarding NH children, in Litovsky [23] SRM was studied in children grouped across the range of 4-7 years, and there was considerable variability, thus here we examined age effects by comparing younger and older children within a similar range. Children in the BiCI groups may vary by certain characteristics, including age at the time of activation of the first CI, age at the time of activation of the second CI, inter-CI gap, and in previous studies they also varied greatly by chronological age (CA). The present study was designed with the goal of narrowing some of the participant characteristics in these groups. Children in an older BiCI group were matched for CA with children in the older NH group. Their average hearing-age (HA) (defined as the amount of time the child has been exposed to sound) was 2 years younger than the average HA of the children in the younger NH group. The younger BiCI group was matched for HA with the younger NH group, and the CA difference was relatively small, because these children received their first CI during infancy. The HA and CA values for each group are shown in Table I. Comparisons between children who have NH (binaural hearing) and children who use BiCIs (bilateral hearing) offers the opportunity to ask whether SRM can be achieved to the same extent with these vastly different stimulation and listening modes.

This study was designed to test the hypothesis that (1) SRM would be reduced with symmetrical interferers (each at 90° but one on the right and one on the left) compared with asymmetrical interferers (both at 90° on the same side), due to the reduction of available monaural cues in the symmetrical distribution. In addition, we aimed to test the hypothesis that (2) children with greater amount of listening experience (older HA) would demonstrate greater SRM. In the NH population, for whom binaural cues are delivered to the auditory system with fidelity, demonstration of SRM in the symmetrical condition is indicative of the ability to segregate targets from interferers under conditions of reduced monaural cues, possibly indicative of the ability to utilize binaural cues for source segregation. There is evidence to suggest that by age 4 children have similar thresholds to adults for ITD discrimination [28] and binaural masking level differences [29]. However, it is unclear as to whether children can utilize these cues on spatial unmasking tasks. There are some conditions under which children perform worse than adults, possibly due to immature binaural temporal processing (e.g.,[11, 30]). Thus, the symmetrical SRM condition serves as a good tool for evaluating the development of functional binaural hearing abilities in children. By comparison, the asymmetrical condition, which is more typically used in SRM studies, is useful for determining whether children benefit from target-interferer spatial separation when multiple spatial cues are present.

II. METHODS

A. Listeners

Participants were 38 native English-speaking children (n=30) and adults (n=8), all of whom received payment for their participation. There were two groups of children with BiCIs, two groups of children with NH, and one group of adults with NH. Hearing age (HA) in children with BiCIs was defined according to parental report and audiologist records as the amount of time the listener had been exposed to sound (time of activation of first CI in addition to any prior acoustic experience (i.e. progressive hearing loss)). HA of the NH groups were equivalent to (CA). Children with BiCIs fell into one of two groups (demographic information is described in Table I): BiCI-A (younger) with HA of 4.0-5.8 years (N=8, HA mean and standard deviation= 5.0 ± 0.6 yrs, CA mean and standard deviation= 6.0 ± 0.8 yrs, bilateral experience mean and standard deviation= 2.1 ± 1.1 yrs) and BiCI-B (older) with HA of 6.5-7.4 years (N=6, HA mean and standard deviation= 6.9 ± 0.3 yrs, CA mean and standard deviation= 7.9 ± 0.8 yrs, bilateral experience mean and standard deviation=4.0 ± 1.6 yrs). These children were recruited from cochlear implant centers throughout the United States and traveled to Madison, WI for participation in the research. All but one subject received their CIs sequentially (CIBB was simultaneously implanted at 7 months of age). The first CI was activated prior to 18 months of age in all but one subject (CIAY), who had post-lingual onset of deafness. The main mode of communication was oral for all BiCI users, as noted by parents. None of the BiCI users had known co-morbidities due to other identified disabilities. During all testing sessions children's CI programs were set to those used most often in daily listening, based on parental and audiologist reports. A loudness balancing procedure, using subjective responses from the participant, was conducted at the first testing session and volume control and/or sensitivity were adjusted to equalize the loudness between the two CI devices as best as possible.

NH children were recruited from the Madison, WI area to match BiCI groups by HA, and were divided into two groups: NH-A (younger) with HA of 4.1-6.3 years (N=8, HA mean and standard deviation= 5.0 ± 0.8 yrs), and NH-B (older) with HA of 6.9-8.8 years (N=8, HA mean and standard deviation= 7.8 ± 0.6 yrs). In addition, 8 adults (ages 18-25 years) were recruited from the student population at the University of Wisconsin-Madison to serve as a comparison group. All NH listeners had hearing sensitivity within normal limits, as indicated by pure-tone air conduction thresholds of 20 dB HL or less at octave frequencies between 250 and 8000 Hz. Because middle-ear problems are common in young children, tympanometry was performed prior to the commencement of testing. According to tympanometric results, all NH subjects exhibited normal peak-compensated static admittance and no children were disqualified due to middle ear anomalies.

This research was approved by, and carried out in accordance with the University of Wisconsin-Madison's Human Subjects IRB regulations. Children ages 7+ yrs signed an assent form, and parents or caregivers of children signed consent forms. NH children were paid an hourly rate for their participation. BiCI users received per diem payments, plus the cost of travel to Madison and all costs associated with the travel were reimbursed to the family.

B. Environment

Testing was conducted in a carpeted standard IAC sound booth (2.8x3.25 m) with reverberation time (RT_{60}) of 250ms. Listeners sat at a foam-covered desk in the center of a loudspeaker array, facing a computer monitor placed at 0° azimuth. Listeners were approximately 1.5m from surrounding loudspeakers. Stimuli were presented from three loudspeakers (Cambridge Soundworks, Center/Surround IV) calibrated prior to each testing session using a sound level meter mounted at the approximate position of the listener's head. The experimental test session lasted approximately one hour for adults and three hours for children. Frequent breaks were given in order to maximize participant attentiveness.

C. Stimuli

Stimuli consisted both of target words and interfering sentences. Target stimuli were a closed set of twenty-five bisyllabic spondees pre-recorded using a male talker. Target spondees were within the vocabulary of the children and were represented by easily identifiable icons (e.g. baseball, cupcake, etc.) [23]. The root-mean-square levels of all target words were equalized. The interfering speech consisted of sentences from the Harvard IEEE corpus [31], which were pre-recorded using a female talker. Sentences were filtered in order to match the long-term average speech spectrum of the target. Two-talker interferers were created by overlaying two recordings of the same voice. The target stimuli described here are identical to those used in a prior experiment with children 4-7 years old [23].

D. Design and procedure

Speech reception thresholds (SRTs) for known target words were measured in quiet and in the presence of interfering speech fixed at 55 dB SPL. For each listener SRTs were measured under 4 conditions: SRT_{Quiet} (target 0° front, no interferers); SRT_{Front} (target and 2-talker interferers both 0° front); $SRT_{+90/-90}$ also known as the Symmetrical condition (target 0° front, 2 interferers, one at +90° right and one at -90° left); and $SRT_{+90/+90}$ also known as the Asymmetrical condition (target 0° front, 2 interferers, at 90° right or left). Note that in the latter condition interferers were placed on the right side for the NH children and near the side of the first CI for the children with BiCIs. When interferers occurred, the trial began with the interferer, the target was subsequently added to the running speech, and after target offset the interferer continued for approximately 1s. All listeners were instructed to attend to the target (male) talker and to ignore the female voices.

Prior to testing, each listener participated in a brief familiarization session in order to determine that he or she could accurately identify all target spondees and corresponding pictures or icons. The task was designed to test speech intelligibility rather than vocabulary, and therefore any of the 25 target spondees that a child was not able to recognize were not used. Out of the 38 listeners, only 8 used spondee lists of fewer than the original 25 words, with no subject using a list of fewer than 20 words (BiCI-A: 2 listeners, BiCI-B: 2 listeners, NH-A: 3 listeners, NH-B: 1 listener). Listeners practiced on the Quiet and Front conditions in order to become familiar with the stimuli, as well as various aspects of the experimental task (e.g. computer controls, listener position). Data collected during the familiarization session was not used in the analysis.

Listeners performed a 4-alternative-forced-choice task [21-24]. Each trial began with the carrier phrase "Ready" followed by a target spondee randomly chosen from the closed set of 25. Four of the previously familiarized visual pictures (3 randomly selected pictures and one picture that matched the target spondee) were displayed on the computer monitor, each matching one of the spondees. The listener was instructed to identify the picture that matched the target word. Responses were provided by pointing to the object with a computer mouse; some children were aided by the experimenter in manipulating the mouse after the subject verbally indicated a response. Incorrect responses were followed by "negative" feedback consisting of a pre-recorded phrase (e.g. "let's try another one" or "that must have been difficult"). There was no auditory or visual feedback following correct responses. To engage and motivate the children during testing, a digitized "puzzle" of a child-friendly picture was revealed one piece at a time following each response, regardless of whether the child was correct or incorrect. Additional reinforcement via stickers and prizes was provided between testing conditions.

E. Speech reception threshold estimation

SRTs were estimated using an adaptive tracking algorithm and calculated using Maximum Likelihood Estimation (MLE) methods for finding the point on the psychometric function at which performance yielded 79.4% correct as described previously [21, 23, 24]. The level of the target words varied according to a hybrid set of rules. Initially, target levels were 60 dB SPL. Levels decreased by 8 dB following each correct response. After the first incorrect response, levels were changed using a 3-down/1-up rule such that three consecutive correct responses resulted in a lower target level and single incorrect responses resulted in a higher target level. After each reversal the step size was halved, with the minimum step size limited to 2 dB. Each adaptive track ended after four reversals.

Unlike previous studies in which a single SRT measurement was obtained (e.g., [10, 21-23]), the goal here was to estimate the repeatability of SRTs in each condition for individual children. We thus sought to obtain three SRT measurements from each child on each of the four conditions. In order to achieve this, the order of testing was randomized over all conditions. An average of the SRTs obtained for each condition for every child was then inserted in the group analysis. Due to time constraints, 7 NH listeners and 7 BiCI users were tested with fewer than three SRTs of each condition. Conditions in which this occurred varied between subjects. Subject CIDP (in group BiCI-A) had one SRT measurement in three of the four conditions, and therefore was excluded from the repeated measures analysis. When time constraints occurred, an emphasis was placed on obtaining more than one SRT measurement in the Front condition, which was used to evaluate release from masking.

F. Spatial release from masking calculation

Spatial release from masking (SRM) was computed for each listener from the mean SRTs in the Front and either Symmetrical or Asymmetrical conditions, such that:

(1) $\mathrm{SRM}_{\mathrm{Symmetrical}} = \mathrm{SRT}_{\mathrm{Front}}\text{-}\mathrm{SRT}_{+90/-90}$

(2) $\mathrm{SRM}_{\mathrm{Asymmetrical}} = \mathrm{SRT}_{\mathrm{Front}}\text{-}\mathrm{SRT}_{+90/+90}$.

In addition, we computed the SRM due to asymmetry, independent of the spatial separation, as follows:

(3) $SRM_{Asymmetry} = SRM_{Asymmetrical} - SRM_{Symmetrical}$

These are based on equations (3), (4) and (5) of Jones and Litovsky [20]. It is important to note here that these values reflect the contributions of separation and asymmetry only at the locations tested here, i.e., at 90.

III. RESULTS

A. Speech reception threshold

In Figure 1, panels A-D show each of the individual SRT values for each of the listeners, in all four conditions (A: Quiet, B: Front, C: 90 Asymmetrical, D: 90 Symmetrical). This plot provides a view of the individual variability observed across participants and within participant. Within each panel, for each group the means (±SD) are also shown (filled circles in the right-most portion of the panels). To test whether repeated measures of SRTs resulted in similar findings for subjects at the conditions tested, SRTs were analyzed for listeners who completed all three repetitions for each condition (BiCI-A, N=2; BiCI-B, N=4; NH-A, N=3; NH-B, N=6; Adult, N=7). Repeated measures ANOVAs were conducted on each group with condition (Quiet, Front, Asymmetrical, Symmetrical) by trial number (1, 2, 3) as the within-subjects factors. There were no significant interactions in any of the groups, and there were no statistically significant differences between the three SRTs per condition (i.e. trials) for any of the groups, suggesting stable SRTs. Thus, averaging SRTs within conditions appears to be a reasonable approach.

Figure 2 shows group means (±SD) for all conditions, and results of statistical comparisons are highlighted (*). SRTs were analyzed using planned t-test comparisons of either within- or between-group differences. A Bonferroni correction for 4 comparisons was applied to each group. Results of the planned comparisons were as follows: (1) Younger NH children (NH-A) had poorer SRTs than older NH children (NH-B) in the Symmetrical [t(14)=3.18, p=0.007, two-tailed] condition, and there were no group differences in the Quiet, Front, or Asymmetrical conditions. (2) Older NH (NH-B) children had poorer SRTs than Adults in all conditions tested: Quiet [t(14)=3.24, p=0.006, two-tailed], Front [t(14)=4.86, p<0.001, two-tailed], Asymmetrical [t(14)=6.31, p<0.001, two-tailed] and Symmetrical [t(14)=4.0, p<0.001, two-tailed]. (3) There were no group differences in younger children with BiCIs (BiCI-A) and older children with BiCIs (BiCI-B), where, for most, SRTs were very close to a SNR of 0. (4) Comparing children in the two older groups (where CA is equivalent and HA is greater in NH than BiCI by a year), BiCI-B had poorer SRTs than NH-B on all four conditions tested: Quiet [t(12)=7.40, p<0.001, two-tailed], Front [t(12)=4.97, p<0.001, two-tailed], Asymmetrical [t(12)=7.42, p<0.001, two-tailed] and Symmetrical [t(12)=10.26, p<0.001, two-tailed]. (5) Comparing children in the two younger groups (where HA is equivalent and CA is greater in BiCI than NH by ~12 months), BiCI users had poorer SRTs than NH children on all four conditions tested: Quiet [t(14)=9.93, p<0.001, two-tailed], Front [t(14)=5.25, p<0.001, two-tailed], Asymmetrical [t(14)=6.58, p<0.001, two-tailed] and Symmetrical [t(14)=4.59, p<0.001, two-tailed]. (6) To further test HA as a critical factor, we compared the older children with BiCIs (BiCI-B) with the younger NH children (NH-A), and here the BiCI users had a HA that was greater by an average of 2 years compared to the NH children. Results showed BiCI users had poorer SRTs than NH children on all four conditions tested: Quiet [t(12)=6.64, p<0.001, two-tailed], Front [t(12)=4.72, p<0.001, two-tailed], Asymmetrical [t(12)=4.92, p<0.001, two-tailed] and Symmetrical [t(12)=5.25, p<0.001, two-tailed].

B. Spatial release from masking

Figure 3 shows individual data points for $\mathrm{SRM}_{\mathrm{Symmmetrical}}$ as a function of $\mathrm{SRM}_{\mathrm{Asymmmetrical}}$ for each listening group (panels A-E). Each data point represents SRM for an individual listener, computed from their average SRTs regardless of how many SRTs they had per condition. Positive SRM indicates that a listener performed better when the interferers were spatially separated from the target than when they were co-located with the target. Negative SRM values indicate that performance was worse with spatial separation of target and interferers. The diagonal line represents equal SRM for the Asymmetrical and Symmetrical conditions. Data points below the diagonal suggest more SRM in the Asymmetrical than the Symmetrical condition. Data points towards the right indicate more SRM, hence a greater benefit with spatial separation of target and interferers. Planned paired sample t-tests were thus conducted for each group (BiCI-A, BiCI-B, NH-A, NH-B, Adult), showing that $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ was greater than $\mathrm{SRM}_{\mathrm{Symmetrical}}$ for two groups. In the BiCI–B group the mean SRM values in the Asymmetrical and Symmetrical conditions were 2.25 dB and -1.07 dB, respectively [$t(5)=2.75$, $p<0.05$, two-tailed]. In the adult group, SRM values in the Asymmetrical and Symmetrical conditions were 7.43 dB and 2.81 dB, respectively [$t(7)=4.78$, $p<0.05$, two-tailed]. While there is a hint of these differences in the other three groups, there is lack of significance, which may be due to large individual variability.

SRM data are also shown in Figure 4, grouped by condition for each subject population. The individual differences can be seen along the y-axis and group mean (±SD) for SRM values are plotted alongside the individual data, for Asymmetrical (Fig.4A) and Symmetrical (Fig.4B) conditions. Figure 4C summarizes the group means and highlights statistically significant comparisons (*).

Group differences were analyzed separately for $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ and $\mathrm{SRM}_{\mathrm{Symmetrical}}$ using planned t-test comparisons, and the Bonferroni correction for 2 comparisons was applied. Results of the planned comparisons were as follows: (1) Comparing children in the two older groups, BiCI-B had less SRM than NH-B in both conditions: $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ [t(12)=-2.62, p=0.022, two-tailed] and $\mathrm{SRM}_{\mathrm{Symmetrical}}$ [t(12)=1.29, p=0.004, two-tailed]. (2) Comparing children in the two younger groups, BiCI-A had less SRM than NH-A in only the Asymmetrical condition: $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ [t(14)=-4.02, p=0.001, two-tailed]. (3) Comparing the NH younger group with the BiCI older group that had 2 years more in HA, results showed BiCI-B had less SRM than NH-A in both the Asymmetrical and Symmetrical conditions: $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ [t(12)=-2.69, p=0.020, two-tailed] and $\mathrm{SRM}_{\mathrm{Symmetrical}}$ [t(12)=-3.079, p=0.010, two-tailed] (4) Comparisons of the younger and older NH child groups with the Adult group showed no difference for either group in either condition.

As per Jones and Litovsky [20], the amount of SRM due to asymmetry ($\mathrm{SRM}_{\mathrm{Asymmetry}}$) was computed for each subject as ($\mathrm{SRM}_{\mathrm{Asymmetrical}}$ - $\mathrm{SRM}_{\mathrm{Symmetrical}}$) and these values are shown for the five participant groups in Figure 5. Figure 5A show individual data points for each listener, and Figure 5B summarizes the group means (±SD) and highlights statistically significant comparisons (*). Visual inspection of the data suggests that adults had the largest values, on average, and that the BiCI-A group had the smallest average values. However, variability was quite high as can be seen from the error bars. Planned t-test comparisons were conducted to test for group differences of $\mathrm{SRM}_{\mathrm{Asymmetry}}$. Older NH (NH-B) children had a smaller effect of $\mathrm{SRM}_{\mathrm{Asymmetry}}$ than Adults [t(14)=-2.38, p=0.04, two-tailed], and no other comparisons were significant, which is not surprising given the large within-group variation.

IV. DISCUSSION

Children spend a vast amount of time communicating and learning in complex acoustic environments, in which multiple sound sources compete for attention and localization. The ability to segregate target speech from competing sources in the environment is likely to yield greater success in learning and attainment of critical communication skills. Studies in which spatial cues differ between target and interferers are aimed at assessing the extent to which listeners can utilize spatial cues in complex listening tasks for source segregation. Here we investigated this ability in adults and in four groups of children, two NH that varied in age, and two with BiCIs, that varied in auditory experience (hearing age) and chronological age.

A. Speech Reception Thresholds (SRTs)

Not surprising, results of the current study suggest that, on average, speech reception thresholds are lower for adult listeners when compared to all children groups. This may be due to a number of factors including continued maturation of central auditory mechanisms and attention throughout childhood (e.g., [32, 33]). Within the children groups, on average, NH listeners had lower SRTs than children with BiCIs, that is, the level of the target required for them to understand speech at ~80% correct was lower. This is important to note in quiet, and also in the presence of interferers, because it suggests that BiCI users need a better SNR than NH children in order to understand the same proportion of words. Similar reports have been noted in children who are hearing impaired and use hearing aids [34] and in adults who use BiCIs (e.g., [35]). This was observed regardless of the fact that the older BiCI group had been wearing two CIs for an average of 4 years. One of the hallmarks of research with cochlear implant users, which was also observed in the present study, is that of individual variability. There was overlap in the SRTs of the better-performing BiCI users and the worse-performing NH children, suggesting that some BiCI users have attained a level of performance that is within the normal range of performance, while other children with BiCIs have substantially poorer performance than their NH peers.

In this study we also examined, for the first time, the repeatability of SRT measurements. In our past work we had selected to obtain SRTs, rather than to measure percent correct, as we were ultimately interested in the effect of spatial cues on the level required for listeners to succeed in 80% speech intelligibility. Unlike previous studies in which a single SRT measurement was obtained for each listener per condition [21, 24, 34, 36], here we introduced a repeated measures factor, such that listeners were tested on each condition several times during the testing session. SRTs were averaged for each subject across the repeated runs, and a single value consisting of the average was used in the data analysis. The findings that SRTs did not differ across repeated measures leads us to conclude that this approach is reliable for children as young as 4 years of age, and for children who are fitted with CIs.

When examining age effects on SRT measurements, older NH children had lower thresholds than younger NH children, but only on the Symmetrical condition with the target and interferers spatially separated towards both ears. This age effect was not observed in the BiCI groups. This finding is consistent with other reports showing that spatial hearing in NH children undergoes considerable maturation during childhood [11], and may be mediated by the continued maturation of auditory cortical mechanisms throughout childhood [37, 38]. The lack of age effects in the BiCI groups is likely related to the fact that bilateral CI processors do not provide binaural cues with fidelity; see below for further discussion.

B. Spatial Release from Masking

Previous research in NH populations has shown that, in Asymmetrical interferer configurations, SRM occurs for children as young as 3-4 yrs [10, 21, 34, 36]. Here we add to this literature by demonstrating that, when the interferers are placed symmetrically in the two hemifields, children who are 4-9 yrs old can benefit from spatial separation of target and interferers. Because monaural head shadow cues are greatly reduced in the Symmetrical condition, the results suggest that children were able to rely on binaural cues for release from masking. Previous work with NH adult listeners suggests that SRM is largest in conditions that contain both monaural and binaural cues, and decreases when maskers are symmetrical so that binaural cues are reduced [20]. Here the difference between Asymmetrical and Symmetrical SRM values was only significant for the Adult group and one of the child groups (BiCI-B, older), who had negative SRM in the Symmetrical condition (see Figure 3). It should be noted that this range of SRM falls within the range considered to be within the margin of error for the test battery used here (±2 dB; [23]). Thus, the spatial separation did not, on average, have a meaningful effect nor did it negatively impact performance. The challenge posed by the Symmetrical condition for BiCI users can be understood by considering the variety of factors they contend with. First, in this condition they are unable to use monaural cues. Second, the microphones are most likely amplifying the two interferers, thus creating a more difficult listening situation than the condition with the co-located stimuli. Third, binaural cues are likely to be minimal or absent.

This last factor points to a well-known characteristic of BiCIs: the lack of binaural coordination between the CI speech processors in the two ears, which likely causes binaural cues to be weak, absent or inconsistent [1, 30]. BiCI users are essentially fit with two monaural systems, which are not coordinated regarding their sampling time or onset time, thus the chances that binaural cues are preserved with fidelity is minimal [39]. In addition, speech processing strategies in CIs which use pulsatile non-simultaneous multi-channel stimulation extract envelopes of signals from the output band-pass filters, discarding the fine structure. Thus, low-frequency interaural time differences, which are known to be important for source segregation (e.g., [12]), are absent to BiCI users.

BiCI users most likely do have access to high-frequency interaural level difference cues; however, the results in the present study suggest that those cues may not be optimal for SRM. In general, ITDs have been shown to play a more important role than ILDs for SRM (e.g., [12]), hence the absence of ITDs is likely to have rendered Symmetrical SRM difficult to achieve. On a related note, as discussed above in relation to SRTs, the spectral degradation of the speech signal in CIs is likely to reduce the dissimilarity between the target and masker, hence the BiCI users are likely to have experienced more informational masking than the NH listeners. In a recent study on BiCI simulation with vocoders that varied the number of spectral channels, Garadat and Litovsky [21] demonstrated that binaural cues play a strong role in enhancing SRM when target and interfering speech are more likely to be confused.

The small SRM in the Asymmetrical conditions for the BiCI users may be due to the small dynamic range, and the microphone characteristics, which, even in the directional setting have a broad range of locations that are amplified in the frontal hemifield. Thus, unlike NH listeners whose acoustic system renders front and side speech signals as having more differentiated levels, CI users may receive front/side signals that are much less differentiated from one another. The Asymmetrical and Symmetrical data are consistent with a study by Loizou *et al.* [5] in adult BiCI users, whereby SRM was studied using a method that replicated conditions for NH adult listeners in Hawley *et al.* [13]. Stimuli in the Loizou *et al.* [5] study consisted of target and interferers, convolved through head related transfer functions measured in a manikin, thus bypassing the microphone of the CIs. Stimuli were provided to listeners via direct connect input to the auxiliary port of the CI in each ear. Results showed that SRM due to binaural interaction was about 0 dB, in contrast with 6 dB in NH listeners. SRM due to monaural cues was about 4 dB in both groups of listeners, suggesting that when the CI microphones are bypassed at least the monaural head shadow cue observed in NH listeners is retained.

Recent research has suggested that children who receive BiCIs simultaneously perform better on speech in noise tasks than those who receive their CIs sequentially [27]. The current study did not address that issue since it includes only one child who received BiCIs simultaneously (CIBB), with all the rest having been implanted sequentially. Interestingly, the one subject who received implants sequentially did not demonstrate markedly notable SRM (Asymmetrical: -1.09 dB, Symmetrical: -2.56 dB).

In this study, NH children were divided into two age groups in order to consider possible effects of maturation and development of speech intelligibility and spatial hearing abilities. While there were age effects in SRTs, these effects were only in the Symmetrical condition. These were not powerful enough to affect significant SRM effects, presumably because SRM is derived from comparisons of the spatially separated and front conditions. Lack of robust age effect on SRM may be due to the dissimilarity between the stimuli used here (male talker and female interferers), whereby informational masking is limited [40]. Because informational masking in children is thought to have a strong developmental component [32, 33], the different-sex stimuli may have created a listening task that reduced the utility of spatial cues in source.

Finally, adults in this study did show SRM on the order of 7.43 dB and 2.81 dB in the Asymmetrical and Symmetrical conditions, respectively. This is within the range of SRM reported previously with the same target-interferer speech corpus and with target and interferers spoken by different-sex talkers, thus easily distinguishable from one another (e.g. [22]). However, the SRM was also smaller than that reported with the same speech corpus but for same-sex target-interferers [20]. These differences are most likely due to the relatively easy task (male-talker: female-interferers using a 4-AFC) and reduced or absent informational masking [24, 41].

V. CONCLUSIONS

(1) Speech reception thresholds (SRTs) in NH children were poorer than in adults.

(2) There was overlap in the SRTs of the better-performing BiCI users and the worse-performing NH children, suggesting that some BiCI users have attained a level of performance that is within the normal range of performance, while other children with BiCIs have substantially poorer performance than their NH peers.

(3) SRTs were poorer in younger NH children than older NH children for conditions with interferers separated, suggesting an age related effect of using spatial cues to hear speech in noise, but this effect did not hold significance when SRM values were analyzed, possibly because SRM is derived from the spatially separated and front conditions.

(4) When SRM was compared within each group between the Asymmetrical and Symmetrical conditions, only the BiCI-B and Adult groups showed a significant difference between the two, with more SRM in the Asymmetrical condition. Lack of significant difference in SRM between the two conditions in other groups may be due to the large individual variability.

(5) When matched for chronological age (older groups) children with NH showed more SRM than children with BiCI for conditions with interferers distributed both asymmetrically and symmetrically.

(6) When matched for hearing age (younger groups), children with NH showed more SRM than children with BiCIs in the Asymmetrical condition. In the Asymmetrical condition, a combination of monaural and binaural cues are available.

(7) When comparing children with BiCIs, who are chronologically older, but have 2 years less hearing experience than the younger normal hearing children, the younger children with NH showed more SRM in both the Asymmetrical and Symmetrical conditions. This suggests that access to binaural cues through acoustic stimulation, whereby the cues are preserved with fidelity, may be a key factor in determining SRM.

(8) Contribution of individual variability to performance is an important next step in these measures, which may help to understand how the ability to use spatial cues for source segregation changes with age and experience in children with BiCIs.

Table I: Subject Demographics

	Subject Code	Chronological Age (yr)	Hearing Age (yr)	Etiology	Age at first CI activation (yr)	Age at second CI activation (yr)	Time between first and second CI (yr)	First CI	Second CI
BiCI-A (younger BiCI users)	CIBU	6.3	5.1	Connexin-26	1.1	5.0	3.9	Med-El Tempo; left ear	Med-El Pulsar; right ear
	CICY	5.7	4.7	Unknown	1.0	4.7	3.7	Advanced Bionics HiRes90K; right ear	Advanced Bionics HiRes90K; left ear
	CIDN	7.0	5.8	Genetic	1.2	6.1	4.9	Med-El Combi40+; left ear	Med-El Pulsar; right ear
	CIDQ	6.6	5.8	Unknown	9 mo	4.4	3.4	N24C; right ear	Nucleus Freedom; left ear
	CIDP	4.9	4	Connexin-26	11 mo	2.8	1.8	Med-El Combi40+; left ear	Med-El Pulsar; right ear
	CICB	5.2	4.6	Connexin-26	10 mo	2.1	1.2	N24C; right ear	Nucleus Freedom; left ear
	CIDX	6.8	5.4	Connexin-26	1.5	2.7	1.1	N24C; right ear	N24C; left ear
	CIBW	5.8	4.8	Connexin-26	1.0	3.9	2.7	N24C; right ear	Nucleus Freedom; left ear
		mean=6.0	**mean=5.0**						
BiCI-B (older BiCI users)	CIAY	9.0	7.1	Progressive, unknown	5.2	6.0	10 mo	N24C; right ear	N24C; left ear
	CIAW	8.6	7.4	Prenatal CMV exposure	1.2	6.6	4.2	N24C; right ear	Nucleus Freedom; left ear
	CIBB	7.0	6.5	Meningitis	7 mo	7 mo	simultaneous	N24C; right ear	N24C; left ear
	CIBI	7.3	7.0	Mondini malformation	1.1	2.1	4 mo	N24C; right ear	N24C; left ear
	CIDQ	7.7	6.9	Unknown	9 mo	4.4	3.4	N24C; right ear	Nucleus Freedom; left ear
	CIEL	7.6	6.7	Unknown	1.2	3.4	2.2	N24C; right ear	Nucleus Freedom; left ear
		mean=7.9	**mean=6.9**						
NH-A (younger NH)	CNS	4.3							
	CNI	6.3							
	CNF	4.2							
	CNJ	5.5							
	CNB	5.3							
	CNK	4.1							
	CKX	5.7							
	CNQ	4.4							
		mean= 5.0							
NH-B (older NH)	CNH	7.6							
	CNG	8.8							
	CKB	7.8							
	CNL	8.0							
	CNU	7.4							
	CKG	8.3							
	CNV	6.9							
	CNW	7.8							
		mean= 7.8							

FIGURES

Figure 1

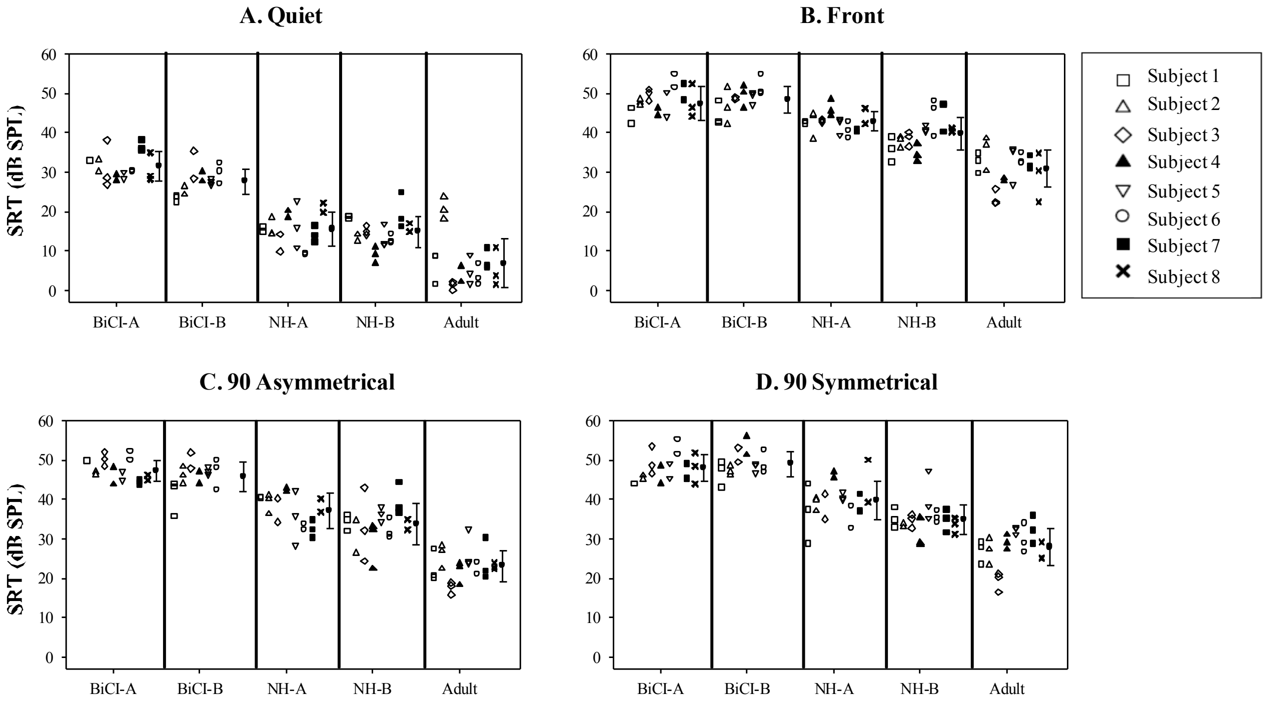

FIG. 1. Speech reception thresholds (SRTs) are plotted for each group in each condition (A: Quiet, B: Front, C: 90 Asymmetrical, D: 90 Symmetrical). Within each group individual subjects are represented by a different symbol. For each subject individual symbols represent a single SRT. Group means (±SD) are shown to the right of all subjects in each group by the filled circles.

Figure 2

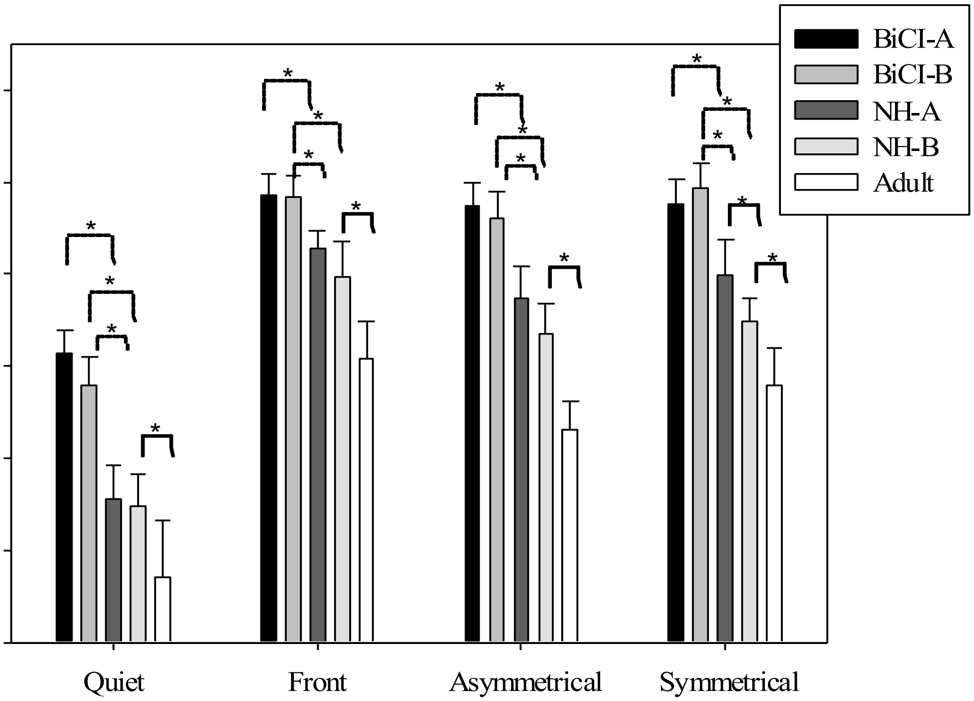

FIG. 2. Mean (±SD) SRTs are shown for each group in each condition (i.e. Quiet, Front, 90 Asymmetrical, 90 Symmetrical). Significant differences (p<0.05) are bracketed and indicated with an asterisk (*). Solid brackets indicate differences within hearing type (i.e. A vs. B). Dashed brackets indicate differences between hearing type (i.e. NH vs. BiCI).

Figure 3

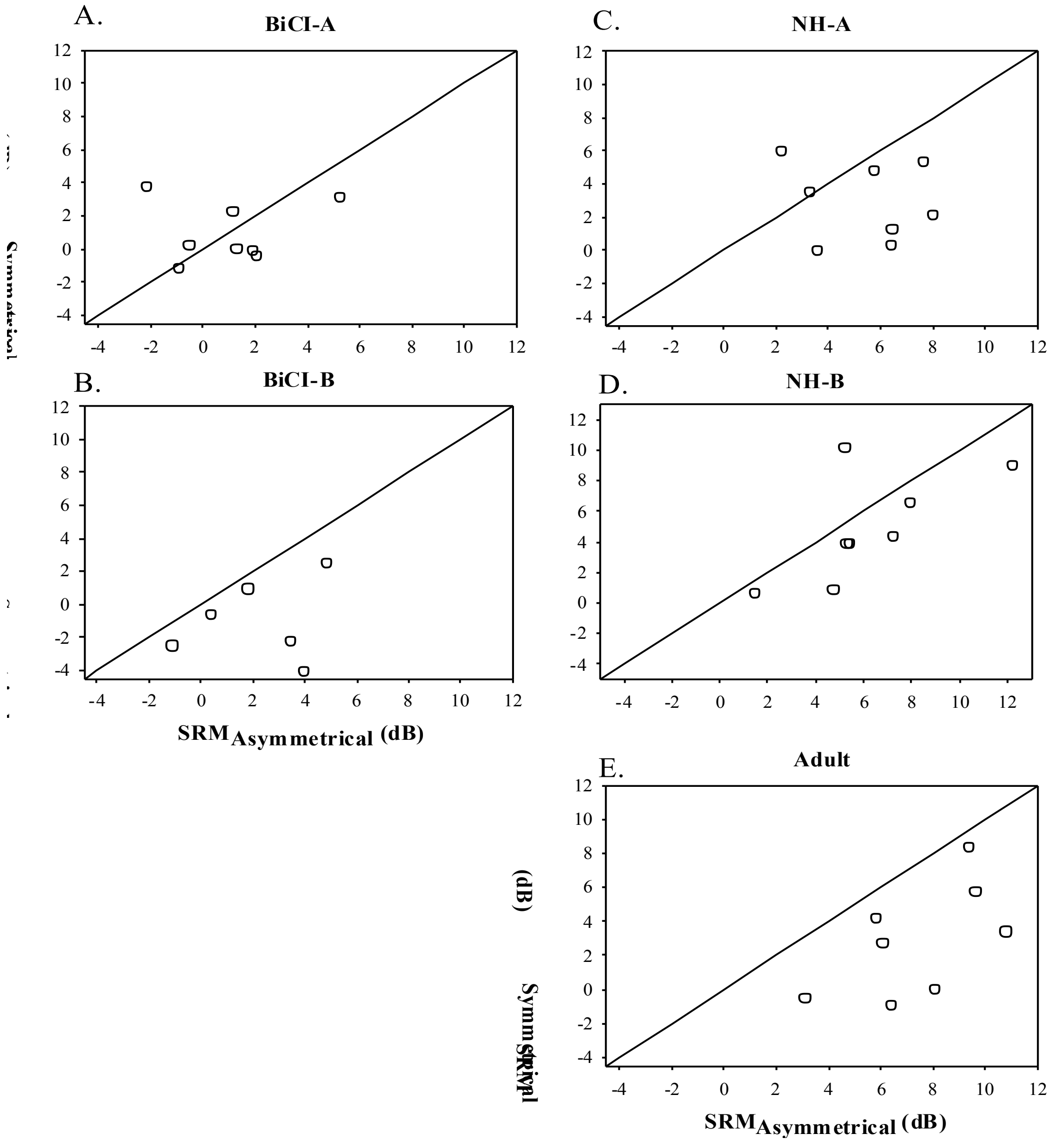

FIG. 3. Spatial release from masking (SRM) values for Asymmetrical and Symmetrical conditions are compared for each individual listener for each group. Panels A-D show groups as follows: A: BiCI-A, B: BiCI-B, C: NH-A, D: NH-B, E: Adult. Each data point represents one listener. The diagonal line corresponds to equivalent SRM in the Asymmetrical and Symmetrical conditions.

Figure 4

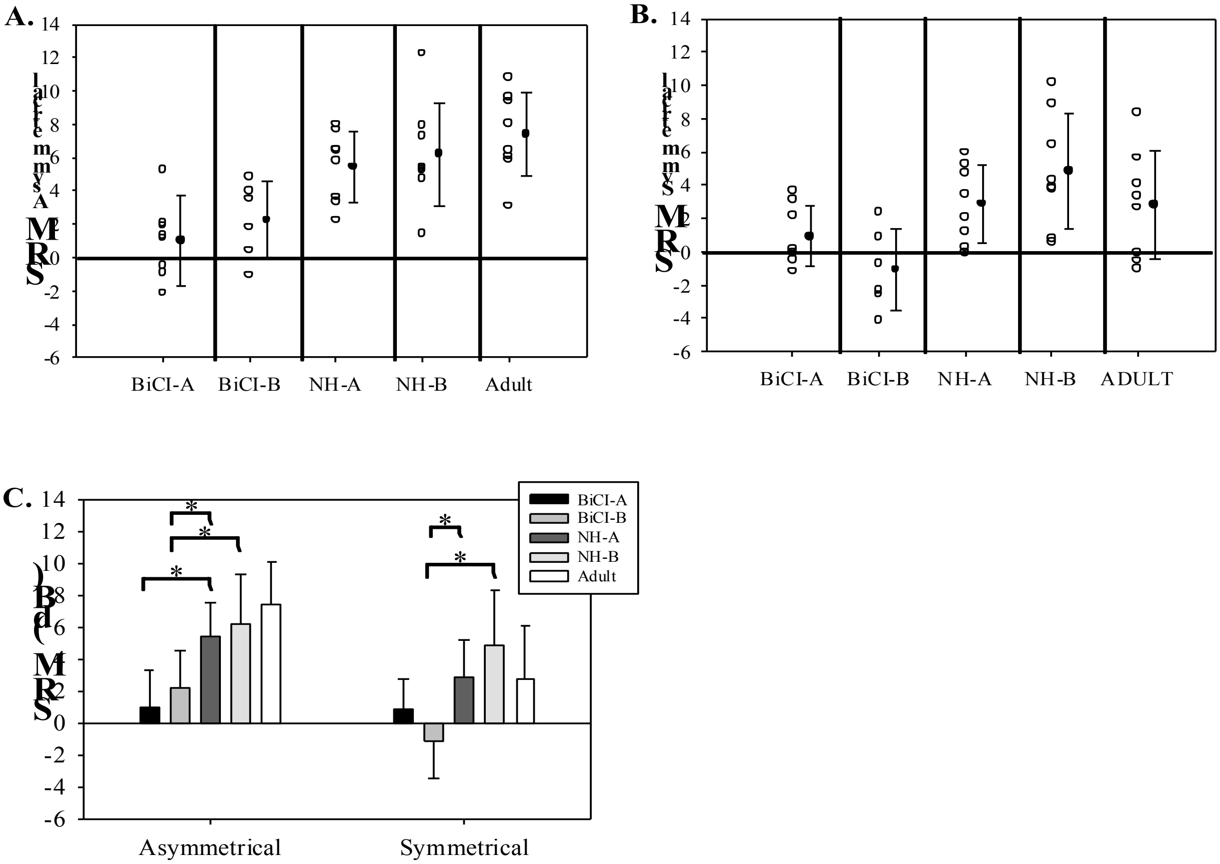

FIG. 4. SRM values are plotted for each group (BiCI-A, BiCI-B, NH-A, NH-B, Adult). Data points represent individual subjects. Group means (±SD) are shown to the right of all subjects in each group. (A) SRM Asymmetrical. (B) SRM Symmetrical. (C) Mean (±SD) SRM for both the Asymmetrical and Symmetrical conditions are summarized for each group in each condition. Significant differences ($p<0.05$) are bracketed and indicated with an asterisk (*).

Figure 5

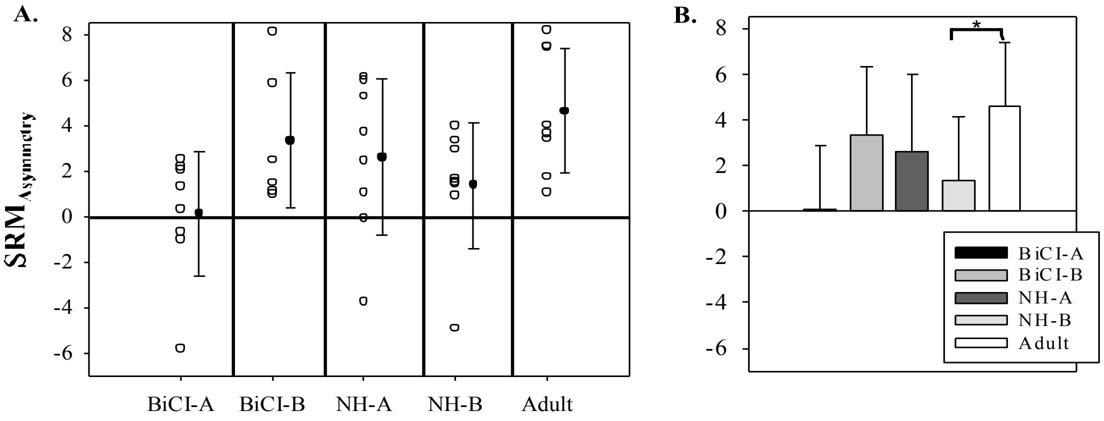

FIG. 5. SRM asymmetry is shown for each group. (A) Data points represent individual subjects. Group means (±SD) are shown to the right of all subjects in each group. (B) Mean (±SD) data are summarized. Significant differences ($p<0.05$) are bracketed and indicated with an asterisk (*).

VI. REFERENCES

1. van Hoesel, R.J. and R.S. Tyler, *Speech perception, localization, and lateralization with bilateral cochlear implants.* J Acoust Soc Am, 2003. **113**(3): p. 1617-30.

2. Nopp, P., P. Schleich, and P. D'Haese, *Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants.* Ear Hear, 2004. **25**(3): p. 205-14.

3. Litovsky, R.Y., A. Parkinson, and J. Arcaroli, *Spatial hearing and speech intelligibility in bilateral cochlear implant users.* Ear Hear, 2009. **30**(4): p. 419-31.

4. Schleich, P., P. Nopp, and P. D'Haese, *Head shadow, squelch, and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant.* Ear Hear, 2004. **25**(3): p. 197-204.

5. Loizou, P.C., et al., *Speech recognition by bilateral cochlear implant users in a cocktail-party setting.* J Acoust Soc Am, 2009. **125**(1): p. 372-83.

6. Van Deun, L., A. van Wieringen, and J. Wouters, *Spatial speech perception benefits in young children with normal hearing and cochlear implants.* Ear Hear, 2010. **31**(5): p. 702-13.

7. Grieco-Calub, T.M. and R.Y. Litovsky, *Sound localization skills in children who use bilateral cochlear implants and in children with normal acoustic hearing.* Ear Hear, 2010. **31**(5): p. 645-56.

8. Grieco-Calub, T.M. and R.Y. Litovsky, *Spatial acuity in 2-to-3-year-old children with normal acoustic hearing, unilateral cochlear implants, and bilateral cochlear implants.* Ear Hear, 2012. **33**(5): p. 561-72.

9. Grieco-Calub, T.M., R.Y. Litovsky, and L.A. Werner, *Using the observer-based psychophysical procedure to assess localization acuity in toddlers who use bilateral cochlear implants.* Otol Neurotol, 2008. **29**(2): p. 235-9.

10. Litovsky, R.Y., P.M. Johnstone, and S.P. Godar, *Benefits of bilateral cochlear implants and/or hearing aids in children.* Int J Audiol, 2006. **45 Suppl 1**: p. S78-91.

11. Godar, S.P. and R.Y. Litovsky, *Experience with bilateral cochlear implants improves sound localization acuity in children.* Otol Neurotol, 2010. **31**(8): p. 1287-92.

12. Culling, J.F., M.L. Hawley, and R.Y. Litovsky, *The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources.* J Acoust Soc Am, 2004. **116**(2): p. 1057-65.

13. Hawley, M.L., R.Y. Litovsky, and J.F. Culling, *The benefit of binaural hearing in a cocktail party: effect of location and type of interferer.* J Acoust Soc Am, 2004. **115**(2): p. 833-43.

14. Cullington, H.E. and F.G. Zeng, *Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects.* J Acoust Soc Am, 2008. **123**(1): p. 450-61.

15. Freyman, R.L., U. Balakrishnan, and K.S. Helfer, *Spatial release from informational masking in speech recognition.* J Acoust Soc Am, 2001. **109**(5 Pt 1): p. 2112-22.

16. Freyman, R.L., et al., *The role of perceived spatial separation in the unmasking of speech.* J Acoust Soc Am, 1999. **106**(6): p. 3578-88.

17. Culling, J.F., K.I. Hodder, and C.Y. Toh, *Effects of reverberation on perceptual segregation of competing voices.* J Acoust Soc Am, 2003. **114**(5): p. 2871-6.

18. Marrone, N., C.R. Mason, and G. Kidd, *Tuning in the spatial dimension: evidence from a masked speech identification task.* J Acoust Soc Am, 2008. **124**(2): p. 1146-58.

19. Zurek, P.M., *A note on onset effects in binaural hearing.* J Acoust Soc Am, 1993. **93**(2): p. 1200-1.

20. Jones, G.L. and R.Y. Litovsky, *A cocktail party model of spatial release from masking by both noise and speech interferers.* J Acoust Soc Am, 2011. **130**(3): p. 1463-74.

21. Garadat, S.N. and R.Y. Litovsky, *Speech intelligibility in free field: spatial unmasking in preschool children.* J Acoust Soc Am, 2007. **121**(2): p. 1047-55.

22. Johnstone, P.M. and R.Y. Litovsky, *Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults.* J Acoust Soc Am, 2006. **120**(4): p. 2177-89.

23. Litovsky, R.Y., *Speech intelligibility and spatial release from masking in young children.* J Acoust Soc Am, 2005. **117**(5): p. 3091-9.

24. Litovsky, R.Y., et al., *Bilateral cochlear implants in children: localization acuity measured with minimum audible angle.* Ear Hear, 2006. **27**(1): p. 43-59.

25. Cameron, S., H. Dillon, and P. Newall, *The listening in Spatialized Noise test: normative data for children.* Int J Audiol, 2006. **45**(2): p. 99-108.

26. Mok, M., et al., *Speech perception benefit for children with a cochlear implant and a hearing aid in opposite ears and children with bilateral cochlear implants.* Audiol Neurootol, 2010. **15**(1): p. 44-56.

27. Chadha, N.K., et al., *Speech detection in noise and spatial unmasking in children with simultaneous versus sequential bilateral cochlear implants.* Otol Neurotol, 2011. **32**(7): p. 1057-64.

28. Ashmead, D.H., et al., *Sound localization and sensitivity to interaural time differences in human infants.* Child Dev, 1991. **62**(6): p. 1211-26.

29. Van Deun, L., et al., *Bilateral cochlear implants in children: binaural unmasking.* Audiol Neurootol, 2009. **14**(4): p. 240-7.

30. Litovsky, R.Y., *Review of recent work on spatial hearing skills in children with bilateral cochlear implants.* Cochlear Implants Int, 2011. **12 Suppl 1**: p. S30-4.

31. Rothauser, E., et al., *IEEE recommended practice for speech quality measurements.* IEEE Trans. Audio Electroacoust, 1969. **17**(3): p. 225-246.

32. Lutfi, R.A., et al., *One factor underlies individual differences in auditory informational masking within and across age groups.* Percept Psychophys, 2003. **65**(3): p. 396-406.

33. Leibold, L.J. and A.Y. Bonino, *Release from informational masking in children: effect of multiple signal bursts.* J Acoust Soc Am, 2009. **125**(4): p. 2200-8.

34. Ching, T.Y., et al., *Spatial release from masking in normal-hearing children and children who use hearing aids.* J Acoust Soc Am, 2011. **129**(1): p. 368-75.

35. Litovsky, R.Y., A. Parkinson, and J. Arcaroli, *Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users.* Ear and Hearing, 2009. **30**(4): p. 419-431.

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

37. Ponton, C.W., M. Don, and J.J. Eggermont, *Place-specific derived cochlear microphonics from human ears.* Scand Audiol, 1992. **21**(3): p. 131-41.

38. Moore, J.K. and Y.L. Guan, *Cytoarchitectural and axonal maturation in human auditory cortex.* J Assoc Res Otolaryngol, 2001. **2**(4): p. 297-311.

39. van Hoesel, R.J., *Contrasting benefits from contralateral implants and hearing aids in cochlear implant users.* Hearing research, 2012. **288**(1): p. 100-113.

40. Durlach, N.I., et al., *Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity.* J Acoust Soc Am, 2003. **114**(1): p. 368-79.

41. Kidd, G., Jr., et al., *Release from masking due to spatial separation of sources in the identification of nonspeech auditory patterns.* J Acoust Soc Am, 1998. **104**(1): p. 422-31.

Chapter 3: Informational masking in children with bilateral cochlear implants and with normal hearing: Effect of target-interferer similarity

ABSTRACT

In complex auditory environments, it is often difficult to separate the target talker from interfering speech. For normal hearing (NH) adult listeners, this difficulty increases as the target and interfering speech become more similar. This study investigated performance of children with NH and with bilateral cochlear implants (BiCIs) when target and interferers were same-sex talkers (male target, male interferer) compared to when the target and interferers were different-sex talkers (male target, female interferer). Performance was worse for both groups with the same-sex vs. different-sex stimuli when sources were co-located in a Front condition. Additionally, effects of spatial cues were investigated by having the interferers positioned in azimuth either symmetrically (+/-90°) or asymmetrically (+90°). Improvement of speech intelligibility when target speech and interferers are spatially separated, as opposed to co-located, is an effect known as spatial release from masking (SRM). Results showed more SRM for NH than the BiCI listeners, suggesting that NH listeners are better able than listeners who use BiCIs to make use of spatial cues to improve speech understanding in noise. When comparing SRM within groups, there was a trend for more SRM with the same-sex vs. different-sex stimuli. Lack of significant differences is likely due to within-group variability.

I. INTRODUCTION

The ability to selectively attend to a target talker while ignoring interfering stimuli is a complex skill utilized in many daily listening environments. This skill is often contemplated within the context of Cherry's [1] early description of the "cocktail party problem," which highlights the difficulty that listeners experience when extracting information from a target speech source in a multi-talker environment. The cocktail party problem is defined in part by the consequences that arise as a result of the target speech competing with other auditory sources in the listening environment. The effect of this competition—described as "masking"—amounts to a reduced ability to hear and understand the target source. Over time, the concept of masking has been more specifically defined in the literature to help create common language around auditory phenomena related to the cocktail party problem. The field of auditory psychophysics differentiates between "energetic" and "informational" masking. The former occurs when target and interfering stimuli contain energy within the same critical bands at the same time [2, 3] and is thought to be accounted for by mechanisms in the auditory periphery. In contrast, it is theorized that informational masking is driven by auditory and attentional mechanisms mediated by neural levels beyond the periphery (i.e., the central auditory pathway), with confusability of the target and interferers under conditions in which both are audible. Thus, with informational masking, factors such as similarity or uncertainty of the sources result in the listener's increased difficulty in segregating them [4-6]. The inability to distinguish between target and interfering speech can cause confusion for listeners in many daily environments, including classrooms, social environments and work-related spaces. The questions that drive the present study are focused on conditions in which listeners can utilize spatial cues to segregate a target talker from other talkers.

In normal hearing (NH) adults, an increase in similarity between the target and interferer typically leads to an increase in masking by a larger amount than would be expected if only the overlap in energy of the target and interferers was being accounted for (i.e. energetic masking). Masking can be greater for speech targets when the interferers are also speech sounds, as opposed to noise, and in particular speech sounds that bear similarity such as same-sex vs. different-sex target and interferers, or same-talker vs. different-talker combinations [7-12]. Relevant to the current study is the finding that, in NH adults, when informational masking is robust, spatial cues are particularly helpful for target-interferer segregation. For example, when both the target and interferers are speech-based stimuli (i.e. speech and time-reversed speech) listeners show more improvement due to availability of spatial cues (spatial release from masking; SRM) than when the target is speech but the interferers are modulated noise [8-11, 13-17]. SRM is an effect whose size has been shown to depend on numerous factors, such as room acoustics [18], number of interferers [15], spatial configuration of sources [19] and perceived spatial separation between target and interferers [9, 16].

Recent studies on this topic have also focused on target-interferer segregation in children. The motivation behind that work is to understand developmental factors that might contribute to source segregation abilities in young children, with possible clinical applications emerging from this work [20-23]. For children who are deaf and use cochlear implants (CIs), the task of segregating speech from noise provides a compelling marker for the extent to which children can benefit from the use of bilateral cochlear implants (BiCIs) vs. unilateral stimulation. Unlike the plethora of data in NH adults, studies investigating the cocktail party problem in children remain small in number. Also notable is that, to date, these studies have used target-interferer stimuli that consist of different-sex talkers under the presumption that clarity of the task for the children (i.e., instructions to "listen to the female and ignore the male") would be maximized (see [21, 23-27]. The purpose of the present study was to identify the role of target-interferer similarity for SRM in NH children and in children who are deaf and fitted with BiCIs. In the previous study on these populations of children [23], it was likely that spectral properties of the different-sex stimuli (i.e. lower fundamental frequency of male versus female talkers) resulted in a listening task in which the target and interfering sources were easily distinguishable for listeners with NH. Therefore, in some cases, the introduction of spatial cues may have provided little additional benefit. In order to shift the focus of source segregation away from benefit possibly due to differences in spectral cues, the current study used same-sex target and interfering speech. This design was intended to improve our understanding of the effect of heightened similarity between the voices, and hence uncertainty about the content of the target, which in adults with NH can result in informational masking [5].

It was hypothesized that children with NH would demonstrate poorer speech reception thresholds (SRTs) with the task containing same-sex target and interferers. It has been shown that children who use BiCIs have difficulty in distinguishing target speech from interferers even with different-sex talkers [21, 23]. This may be due to the degraded clarity of spectral cues provided by the CI device, resulting in interferer and target voices that are perceptually similar in the same way that same-sex voices are. Other research has suggested that gender distinction is a challenge for CI users [28-30]. In order to test this idea we conducted further investigation with this population, with same-sex target and interferer, and predicted that if children with CIs indeed perceive same-sex and different-sex talkers as similar, then performance with either interferer in place would be similar. In other words, due to the limited number of spectral channels in a CI, children with BiCIs would likely experience more perceptual confusability with either set of stimuli and demonstrate more masking than NH children.

A second focus of this study was the age of participants: children with NH and with BiCIs were recruited so that they would be matched either according to hearing age (HA, the amount of time exposed to sound) or chronological age (CA). Relatively little is known about how children who use CIs and receive auditory input through electrical stimulation to the auditory nerve, without integration of inputs to the two ears, compare to children with NH in these multi-source environments. We hypothesized that even when matched for HA, the BiCI groups would perform worse on tasks of spatial hearing than the NH groups, due to the fact that the processors for each ear are not coordinated, resulting in minimal or absent binaural cues [31, 32]. We were also interested in examining age effects within populations of children with NH or BiCIs. Although NH listeners of all ages are able to use spatial cues for source segregation, developmental effects have been reported in free field [23] and for contralateral unmasking headphone stimulation [33, 34]. This difference could be due to the fact that information masking also depends on development of central auditory mechanisms and non-auditory factors such as attention [35-37].

A third question we asked in this study regarded the effect of interferer location on SRM. We were interested in the extent to which SRM occurs when interferers are directed toward one ear, where monaural cues in the opposite ear may facilitate source segregation, compared to

SRM in conditions in which interferers are placed symmetrically toward both ears, where listeners must rely more heavily on binaural cues. In our previous study [23] we reported that in NH children ages 5 and older, SRM occurs not only in the Asymmetrical condition, but also in the Symmetrical condition. In that population, we are now examining the extent to which same-sex target-interferers would result in enhanced SRM. Furthermore, in children with BiCIs, who demonstrate SRM in the Asymmetrical condition due to the availability of the monaural head shadow cue, we are interested in understanding whether they can also benefit from spatial cues in the Symmetrical conditions. It is reasonable to surmise that lack of coordination between the two CIs may render the conditions with monaural cues most relevant for SRM [21, 23, 38-40]. In fact, previous studies on the cocktail party problem in hearing impaired listeners and CI users have quantified benefits with reference to whether one or both ears are stimulated. Commonly, the “better ear effect” (or “monaural head shadow”) is observed and is thought to be due to a more favorable signal-to-noise ratio (SNR) in one ear [15, 41]. However, effects due to spatial separation are less commonly explored. The study was designed to test the hypothesis that less spatial unmasking would occur with interferers directed toward both ears (symmetrically) than directed only toward one ear (asymmetrically), and whether the increased informational masking with same-sex target and interferers would render spatial cues more effective.

In summary, the first issue of focus in the current project is the effect of using same- vs. different-sex talkers for spatial unmasking in children with NH. The second asks whether children who are fitted with BiCIs demonstrate similar effect sizes, and how that relates to HA and CA of the children with NH. The third question explores the effect of interferer location and whether SRM depends on the availability of the “better ear” or can also occur in children when interferers are distributed symmetrically around the head, and in particular when informational masking is induced with same-sex stimuli.

II. METHODS

This study was designed such that we could compare SRTs and SRM *within* (young (A) versus old (B)) and *between* (NH versus BiCI) groups. In addition, the design was guided by the goal of comparing the current data, using same-sex (male target, male interferers) stimuli, to those from Misurelli and Litovsky [23] where different-sex stimuli were used.

A. Listeners

The population of children tested consisted of 4 children with BiCIs, 11 children with NH, and 6 adults with NH who had participated in our recently published study with different-sex stimuli [23], and who were retested for this study with same-sex target and interferers. We recruited 9 additional children with BiCIs, 5 children with NH, and 2 adults with NH all of whom were tested with both sets of stimuli. The children whose data for the different-sex talkers that were published in the previous study are shown in Table I with an asterisk next to the subject code. Thirty-seven (37) native English-speaking listeners were recruited, and all received payment for their participation. Five groups of listeners were tested: two groups of children with NH, two groups of children with BiCIs, and one group of adults with NH. Children groups were recruited according to hearing age (HA), defined as the amount of time the listener had been exposed to sound (both acoustic and electric). For the groups with NH, HA was equivalent to chronological age (CA). For the children with BiCIs, HA was calculated from the time of activation of the first CI in addition to any acoustic experience prior to CI activation (i.e. cases of sudden or progressive hearing loss). There were 4 listeners who had acoustic experience prior to receiving their first CI: 3 in the younger BiCI group with varying histories (CICA-hearing loss identified at 2 years; CIET-enlarged vestibular aqueduct syndrome (EVAS)/sudden hearing loss; and CICF-progressive hearing loss due to meningitis) and 1 listener in the older BiCI group (CIEK who had EVAS/progressive hearing loss).

Children with BiCIs were recruited from CI centers throughout the United States and traveled to Madison, WI to participate. These children visited the lab for three consecutive days and participated in a number of experiments. During all testing sessions, children's CI processors were set to their clinically programmed every-day listening mode, as confirmed by parent and audiologist reports. Prior to the first testing session a subjective loudness balancing procedure was conducted, and volume and sensitivity controls were adjusted to best equalize the loudness between the two CI devices. These children were separated into two groups according to their HA (see Table 1 for demographic information). In the younger group (BiCI-A) there were 7 children, with a HA of 5;0 ± 0;10 (mean and standard deviation for years;months), and a CA of 6;4 ± 0;9 (mean and standard deviation for years;months). This group had bilateral experience of 3;2 ± 1;1 (mean and standard deviation for years;months). In the older group (BiCI-B) there were 6 children, with a HA of 7;2 ± 0;10 (mean and standard deviation for years;months), and a CA of 8;5 ± 0;6 (mean and standard deviation for years;months). This group had fairly similar amount of bilateral experience to that of the younger group (3;4 ± 0;9 mean and standard deviation for years;months). Group BiCI-A had three children who received their CIs simultaneously (CICA, CIEH, CIET) and four children who received their CIs sequentially. In the sequentially implanted children, the first CI was activated by 18 months of age for all but one child who had a late identification of hearing loss (CIDW). For group BiCI-B, all children received their CIs sequentially. In this group, all but two listeners (i.e. CIDJ-hereditary hearing loss and CIEK-EVAS/progressive hearing loss) had their first CI activated by 18 months of age. For both BiCI groups, all children had at least one year of experience listening with two cochlear implants (mean, min, max (years;months) = 3;3, 1;7, 4;5), and the main mode of communication noted by parents was oral. No children were known to have co-morbidity factors due to identified disabilities.

All children with NH were recruited locally from the Madison, WI area. Children with NH were selected to match BiCI listeners for age in such a way that we were able to compare performance based on HA or CA. NH children were recruited into two groups. In the younger group (NH-A) there were 8 children, with a CA of 5;2 ± 1;0 (mean and standard deviation for years;months). In the older group (NH-B) there were 8 children, with a CA 8;3 ± 0;11. The younger NH group (NH-A) matched the younger BiCI group (BiCI-A) for HA, and, on average, the BiCI-A was chronologically 1;2 older than NH-A. NH-B matched the older BiCI group (BiCI-B) for CA and, on average, the NH-B had 1;1 more hearing experience than BiCI-B. In addition, 8 NH adult listeners (ages 18-25 years) were recruited from the student population at the University of Wisconsin-Madison. All NH listeners were given a hearing screening consisting of pure-tone thresholds of 20 dB HL or less at octave frequencies between 250 and 8000 Hz. Children were also screened using tympanometry in order to rule out any middle ear anomalies or infection on the day of testing. Tympanometric results indicated normal peak-compensated acoustic static admittance at each testing session. No participants that were recruited were excluded from participation due to a failed hearing screening.

This research was approved by, and carried out in accordance with, the University of Wisconsin-Madison Human Subjects IRB regulations. Before testing commenced, all adult participants signed a consent form and, for children, caregivers signed a consent form. Children with a CA of 7 years or older signed an assent form.

B. Testing environment

Testing was conducted inside a standard Industrial Acoustics Company (IAC) sound booth (dimensions 2.8 x 3.25 m) with a reverberation time (RT_{60}) of 250 ms. During testing, listeners sat at a foam-covered desk in the center of a semicircular loudspeaker arc, facing a computer monitor located at 0° azimuth. Loudspeakers (Cambridge Soundworks, Center/Surround IV) were positioned at ear level approximately 1.5 m from the listener and were calibrated prior to each testing session. For NH listeners, testing was completed after two visits to the lab, with each visit lasting approximately one hour for adults and three hours for children. Children with BiCIs completed the task within their multi-day visits.

C. Stimuli

The target stimuli were a closed-set of 25 spondees (a two-syllable word with equal stress on both syllables) selected to be within the vocabulary of children ages 4 years and older [22], and were pre-recorded using a male talker (all root-mean-square levels equalized). Interfering sentences were taken from the Harvard IEEE corpus [42] and were pre-recorded separately using a male talker (different talker than the male target talker) and a female talker. Sentences were filtered to match the long-term average speech spectrum of the target spondees. Two-talker interferers were created by overlaying two recordings from the same talker, either male or female. Figure 1 shows spectral differences between the target stimuli and the male and female interferers.

D. Design and procedure

Speech reception thresholds (SRTs) were measured for target spondees in quiet and in the presence of interfering speech. For each listener, SRTs were measured in quiet with no interferers present and in 3 conditions with interferers fixed at 55 dB SPL. More specifically, SRTs measured in the 4 conditions are denoted as: $\mathrm{SRT}_{\mathrm{Quiet}}$ (target 0° front, no interferers), $\mathrm{SRT}_{\mathrm{Front}}$ (target and interferers both 0° front), $\mathrm{SRT}_{+90°/+90°}$ (target 0° front, interferers placed in an asymmetrical configuration at 90° to one side or another) and $\mathrm{SRT}_{+90°/-90°}$ (target 0° front, interferers placed in a symmetrical configuration with one at +90° and one at -90°). In the asymmetrical condition, interferers were placed 90° to the right for the NH listeners and 90° to the side of the first CI for the BiCI listeners.

SRTs for each of the conditions were obtained using both the female and male interferers in order to perform within group comparisons. To minimize any order effects, all conditions were randomized within interferer type. Due to time constraints, fewer than three SRTs per condition were measured on 9 NH listeners (NH-A: 3, NH-B: 6) and 8 BiCI listeners (BiCI-A: 4, BiCI-B: 4). Conditions in which this occurred varied between listeners. SRTs were averaged for data analysis, resulting in one SRT per interferer for each of the four conditions. As described in Misurelli and Litovsky [23], SRTs in children with NH and with BiCIs have been shown to be consistent between trials; therefore warranting averaging SRTs within each condition for each listener.

Each listener participated in a brief familiarization task before testing commenced in order to verify accurate identification of each target spondee with the corresponding icon. The experiment was designed with the intent of measuring SRTs in quiet and in noise, but not in order to test vocabulary. For that reason, any target spondee that an individual listener was not able to identify was not used in the experimental testing (similar approaches were used in our previous studies by Litovsky and colleagues). Out of the 37 listeners, only two used a target list less than 25 words, with no list less than 19 words (NH-A: 1 listener, BiCI-A: 1 listener). Practice in the Quiet and Front conditions was conducted for all listeners in order to allow each listener to become familiar with various aspects of the testing (i.e. listener position, computer controls, stimuli). Data collected during the practice session was not included in the analysis, and was used solely to assure each listener was comfortable with the experimental task. A trained tester accompanied all child listeners in the booth.

The experimental test consisted of a 4-alternative-forced-choice task [20-23, 43], in which the target spondee was identified from four possible icons. On trials with male interferers, listeners were given experience hearing the interferer sentences and the male target; subsequently they were instructed to ignore the “boy talkers” and to pay attention to the man’s voice (male target). On trials with female interferers, listeners were given experience hearing the interferer sentences and the male target; subsequently they were instructed to ignore the “lady talkers” and to pay attention to the man’s voice (male target). Trials began with the word “ready,” followed by one randomly chosen spondee from the list of 25. Similar to our previous studies, in conditions with interferers, the interferers were turned on first followed by the target. Interferers continued for approximately 1s after the target was turned off. After the target was presented four pictures were displayed on the computer monitor, 3 of the 4 pictures were randomly selected from the closed-set of 25 and 1 of the 4 pictures corresponded with the target spondee. The listener was then asked to identify the picture that corresponded to the spoken spondee. Feedback was given only after incorrect responses. The feedback was pre-recorded phrases such as, “let’s try another one” or “that must have been difficult”. Regardless of whether the child identified the correct picture or not, children were reinforced after each response by a computer display of one digitized puzzle piece. Children were also reinforced with stickers and small prizes, and “listening breaks” were given when necessary.

E. Speech reception threshold estimation

SRTs were measured using an adaptive tracking method. Target level was initially presented at 60 dB SPL, decreasing in intensity with correct responses and increasing in intensity with incorrect responses. Initially levels decreased by 8 dB following correct responses. Following the first incorrect response target levels were adjusted using a 3-down/1-up procedure, and step size was halved with each reversal of target level. Testing was terminated after four reversals. SRTs were calculated using Maximum Likelihood Estimation (MLE) methods [44, 45], which define threshold as the point on the psychometric function in which stimulus intensity level corresponded to a performance level of 79.4% correct. The methods used have been previously described [20, 22, 23, 43].

F. Spatial release from masking

SRM was calculated for each listener by subtracting the mean SRT in either the asymmetrical or symmetrical conditions from the mean SRT in the front condition. SRM was defined as the difference between SRT_{Front} and either the $SRT_{+90°/+90°}$ or $SRT_{+90°/-90°}$, such that:

1. $SRM_{Asymmetrical} = SRT_{Front} - SRT_{+90°/+90°}$
2. $SRM_{Symmetrical} = SRT_{Front} - SRT_{+90°/-90°}$

This approach is identical to that used by Misurelli and Litovsky [23] to measure SRTs with female interferers. Positive SRM values indicate an improvement in identification of the target when interferers are spatially separated versus co-located with the target. The larger the amount of SRM, the greater the perceived benefit with the sources spatially separated. Negative SRM indicates that a listener performed worse when the target and interferers were spatially separated.

III. RESULTS

A. Speech reception thresholds

Mean (±SD) SRTs are shown in Figure 2 comparing data collected with the male interferers (filled symbols) and data collected with the female interferers (open symbols). Visual inspection of the SRT data in Figure 2 suggests that the male interferer produced higher SRTs in general; statistical assessment of these effects was conducted within each listener group, in each condition, using planned paired t-test comparisons. Statistically significant differences are denoted (*) within each panel. Bonferroni corrections for 3 comparisons per planned comparison group were applied. Results of the planned t-test comparisons as shown in Table II identified a number of condition/group combinations in which the male (same-sex) interferers produced higher SRTs (worse performance) than the female (different-sex) interferer. This effect was observed: (a) for all subject groups in the spatial configuration with both target and masker in Front (co-located); (b) in the Symmetrical condition for the two groups of NH children; and (c) in the Asymmetrical condition for the older NH children and adults.

Figure 3 summarizes between-group SRT data. Results of planned t-test comparisons with Bonferroni corrections for 4 comparisons per group were applied and are summarized in Table III. Results showed slightly more significant differences between groups that have differences in hearing type (NH vs BiCI) than in groups that have differences within hearing type but differ in age (A vs B).

B. Spatial Release from Masking

Figure 4 displays individual data points for $\mathrm{SRM}_{\mathrm{Symmetrical}}$ vs. $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ for each listening group with each interferer. The diagonal line represents unity for $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ and $\mathrm{SRM}_{\mathrm{Symmetrical}}$. The data points below the diagonal line denote cases in which listeners demonstrated more SRM with interferers located in the asymmetrical condition; data points on the diagonal indicate unity, i.e., no effect of interferer symmetry vs. asymmetry; data points above the diagonal indicate that SRM was greater with the symmetrically distributed interferers. Moreover, data points in the right or top portions of each panel denote greater SRM. Planned paired t-tests were conducted within each listener group, to evaluate the effect of distributing the interferers symmetrically vs. asymmetrically relative to the listener's head. $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ was greater than $\mathrm{SRM}_{\mathrm{Symmetrical}}$ for all listener groups with the same-sex interferer (BiCI-A: *p=0.03, BiCI-B: *p=0.04, NH-A: *p=0.001, NH-B: *p=0.01, Adult: *p=0.04). However, with the different-sex interferers the same effect of interferer location was found only for the adult group (BiCI-A: p=0.34, BiCI-B: p=0.37, NH-A: p=0.14, NH-B: p=0.14, Adult: *p=0.004).

Figure 5 summarizes the mean (±SD) SRM data for each group in both the Asymmetrical (4A) and Symmetrical (4B) conditions. Planned paired t-test comparisons on differences in SRM within groups for the male vs. female interferers showed a statistically significant difference for only the NH-B group in the asymmetrical condition [t(7)=4.35, p=0.003], and the BiCI-A group in the symmetrical condition [t(6)=3.01, p=0.02]. Although all groups, on average, exhibited more SRM with the male interferers, large variability within groups was a likely determining factor in the lack of statistically significant differences in most cases.

The data in Figure 5 were re-plotted in Figure 6 in order to highlight the differences in SRM that were observed with the two different interferer configurations with each the male interferers (6A) and the female interferers (6B). Significant results from planned t-tests with a Bonferroni correction for 2 comparisons are denoted (*) and listed in Table IV. Results of the planned comparisons with the male interferers were as follows (Figure 6A): (1) No differences were found between the older and younger child groups for either the NH or BiCI groups. (2) Comparing children in the two older groups (matched for CA), NH-B had more SRM than BiCI-B in both conditions: $\text{SRM}_{\text{Asymmetrical}}$ [t(12)=4.11, p=0.001, two-tailed] and $\text{SRM}_{\text{Symmetrical}}$ [t(12)=4.28, p=0.001, two-tailed]. (3) Comparing the children in the two younger groups (matched for HA), NH-A had more SRM than BiCI-A in only the $\text{SRM}_{\text{Asymmetrical}}$ [t(13)=4.03, p=0.001, two-tailed]. (4) Comparing the younger NH group with the older BiCI group, NH-A had more SRM in both conditions: $\text{SRM}_{\text{Asymmetrical}}$ [t(12)=3.19, p=0.008, two-tailed] and

$SRM_{Symmetrical}$ [t(12)=2.58, p=0.024, two-tailed]. (5) There were no differences for either condition when comparing the younger and older NH children with the Adult group.

Results of the planned comparisons with the female interferers were as follows: (1) There were no differences between the older and younger groups in either the NH or BiCI children groups. (2) Comparing children in the two older groups (matched for CA), NH-B had more SRM than BiCI-B in both conditions: $SRM_{Asymmetrical}$ [t(12)=3.80, p=0.003, two-tailed] and $SRM_{Symmetrical}$ [t(12)=3.37, p=0.006, two-tailed]. (3) Comparing the children in the two younger groups (matched for HA), NH-A had more SRM than BiCI-A in both conditions: $SRM_{Asymmetrical}$ [t(13)=4.40, p=0.001, two-tailed] and $SRM_{Symmetrical}$ [t(13)=3.55, p=0.004, two-tailed]. (4) Comparing the younger NH group with the older BiCI group, NH-A had more SRM in both conditions: $SRM_{Asymmetrical}$ [t(12)=-3.74, p=0.003, two-tailed] and $SRM_{Symmetrical}$ [t(12)=-2.91, p=0.01, two-tailed]. (5) There was no difference for either condition when comparing the younger and older NH children with the Adult group.

IV. DISCUSSION

The ability to segregate sources and attend to a target talker is particularly important for children, for whom learning and development often occur in multi-source, generally noisy environments. Difficulties can arise when uncertainty exists regarding the particular sources to which they should attend, a challenge that can be exacerbated by similarity between sources (e.g., [5, 8, 20]). Here, we compared performance on speech intelligibility in the presence of speech interferers consisting of talkers that were either the same sex as the target talker, or a different sex than the target talker (the latter was also tested by Misurelli and Litovsky [23]).

Further, we investigated listener ability to take advantage of spatial cues for source segregation when spectral cues were limited (i.e., the target and interferer were both male talkers). For this study, we tested two groups of children with NH, two groups of children with BiCIs, and one group of adults with NH. The NH child groups were divided according to chronological age (CA) and the BiCI child groups varied in both CA and hearing age (HA).

A. Same-Sex vs. Different-Sex Stimuli

Talker discrimination is essential to successful verbal communication. Individuals must use a variety of cues in order to accurately identify and segregate target and interfering speech in noisy environments. One cue that NH individuals use to achieve this is spectral differences between sources (e.g., voice pitch). The NH literature has shown that overall, speech understanding in noise is worse when the target and interferers are more spectrally similar for both adults [5, 8] and children [20]. Furthermore, when NH individuals listen to processed CI speech through a 24-channel CI simulation, they demonstrate difficulty extracting the target speech from the noise [46], similar to that shown in BiCI users [47]. In CI users, individual talker discrimination has been shown to be poorer than in both NH children [30, 48] and NH adults [28, 29]. Specifically, this research has shown that CI users find it more difficult to distinguish between voices of the same sex [30, 48], but also have difficulty discriminating between voices of different sex talkers [28, 29]. This inability to clearly discriminate between talkers in quiet conditions is exacerbated in more realistic environments with numerous sources played simultaneously. It is not surprising that CI users have more difficulty discriminating between talkers and worse speech intelligibility in noise than NH listeners [23, 40, 49], given that the current clinically available CIs deliver signals to a reduced number of frequency channels. Our results show that for all groups in all conditions, on average, SRTs were higher with the male (same-sex) interferers than with the female (different-sex) interferers, and all listeners were able to identify target speech in Quiet at a much lower SRT than when interferers were present. That is, speech intelligibility was worse with interferers present and with more similar target-interferer stimuli. When we compared SRTs with same vs. different-sex stimuli, statistical significance between stimuli sets was reached for all groups in the Front condition, meaning that all groups needed a higher SNR in order to identify the target when the stimuli was spectrally more similar. It is likely that the lack of spectral difference between the target and the interferer with the same-sex stimuli introduced more informational masking, and therefore contributed to the poorer performance (i.e., higher SRTs) in these listening conditions [12, 13, 50, 51]. It is somewhat surprising that both BiCI groups showed significantly higher thresholds with the same-sex stimuli in the Front condition, given that spectral information is not strongly encoded in the current CI device [52]. It may be that that the spectral differences between the two male voices used in this study were different enough that some CI users were able to use weak envelope spectral cues to distinguish between the target and interferer.

When we spatially separated the target and interferers, significant differences in performance with same- vs. different-sex stimuli occurred between groups and conditions (Asymmetrical, Symmetrical). For groups with no difference in SRTs between the two sets of stimuli in the spatially separated conditions, no additional masking was demonstrated with the same-sex stimuli; that is, similar amounts of masking were demonstrated with either set [53]. It is interesting that when interferers were spatially separated, neither BiCI group showed a significant difference in SRTs with either same- or different-sex stimuli. It may be that, for CI users, spatially separating the stimuli poses a more difficult listening environment, independent of target-interferer similarity. NH listeners are able to use directional acoustic cues to differentiate minute changes in intensities of signals arriving from front and side locations. CI users lack this fine-tuned ability to consistently and accurately differentiate signals from various frontal hemifields because the directional microphones used in CI processors sit above the pinna and amplify a much more broad range of source locations.

All groups, on average, showed more SRM (asymmetrical and symmetrical) with the same sex stimuli vs. different sex stimuli, although, statistically significant differences were only demonstrated in the older NH ($\mathrm{SRM}_{\mathrm{Asymmetrical}}$) and the younger BiCI ($\mathrm{SRM}_{\mathrm{Symmetrical}}$) groups (see Fig. 5). This lack of significant differences in all groups is likely due to the within-group variability on the speech in noise task. It is not surprising to see large amounts of variability in the BiCI groups, given that variability on performance on auditory tasks is a consistently demonstrated hallmark of CI research literature. Much of this variability can be accounted for by demographic factors such as length of CI use, age at implantation, inter-CI gap, and bilateral experience. What is of particular interest here is the large variability that is also demonstrated within the NH child groups. It is known that auditory pathways continue to mature into adolescence [54, 55]. Specific to this study, informational masking has been shown to have a developmental component [33, 36]. Wightman *et al.* [37] tested NH populations ages 5 to 61 and showed that the most variability on a dichotic listening task involving informational masking was observed in NH children ages 6 to 12 years of age. This finding is congruent with our results, suggesting that the variability found in our NH children groups ages 5 to 8 years may be due to incomplete development of central auditory pathways.

B. Interferer Location

For this study we also sought to investigate spatial release from masking (SRM) with interferers directed either toward one ear (Asymmetrical), where monaural head shadow cues are available, or both ears (Symmetrical), where monaural cues are reduced and listeners must rely more on binaural cues.

Benefits of speech understanding in noise can be attributed to three factors: monaural head shadow, when interferers are directed to only one ear, and therefore the opposite ear has an advantageous SNR; binaural squelch, or benefit from adding the ear with the poorer SNR; and binaural summation, in which target and interferers occur from the same direction, and redundancy, rather than spatial separation, provides the listener with cues to improve speech understanding. Although NH listeners are able to use each of these cues to gain considerable benefit when listening to speech in noise, BiCI users do not show nearly as much benefit. This is likely because they have limited access to binaural cues, and therefore rely primarily on the monaural head shadow cue. Adults who use BiCIs generally show better speech understanding in noise than people with only one CI [39, 40], but due to the lack of coordination between the two CI devices, binaural-driving benefits are still reduced compared to NH listener. Speech understanding in noise for BiCI children [56] and adults [39, 40, 49] can be largely accounted for by monaural head shadow.

Previous work has shown that NH adults demonstrate more $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ than $\mathrm{SRM}_{\mathrm{Symmetrical}}$ with same-sex [23, 57] and different-sex [23] stimuli. Although SRM is largest in asymmetrical conditions where both monaural and binaural cues are available, we have also recently shown that NH children as young as 4 years of age demonstrate SRM when sources are symmetrically distributed [23]. This result provides evidence that by age 4, NH children are able to take advantage of binaural cues to benefit from spatial separation when interferers are directed towards both the right and left ears [23]. The results of the current study show that significantly less $SRM_{Symmetrical}$ vs. $SRM_{Asymmetrical}$ is demonstrated for all groups when same-sex stimuli were used; this same effect was only demonstrated in the adult group when using different-sex stimuli. The larger difference in $SRM_{Asymmetrical}$ and $SRM_{Symmetrical}$ with the same-sex stimuli suggests that, when spectral cues are reduced (and therefore masking is increased), spatial cues are more effective. It is likely that informational masking contributed to the increased difference between the Symmetrical and Asymmetrical SRM, such that in the Symmetrical condition not only were monaural cues reduced, but the stimuli also lacked spectral differences, creating an extremely difficult listening environment.

In summary, our results show that BiCI groups were better able to use spatial cues with asymmetrical vs. symmetrical interferers, and NH groups were better able than BiCI groups to use spatial cues to segregate the target from the interferers in both the Asymmetrical and Symmetrical conditions.

C. Age Comparisons

In conditions where spectral cues are easily distinguishable (e.g. different-sex stimuli) to a NH listener, BiCI users demonstrate worse performance when matched to NH listeners for either HA or CA [23]. The current results are consistent with previous findings that both BiCI groups showed worse performance (i.e. higher SRTs) than NH groups when identifying a target with either interferer. Other work with children CI users shows similar results, in that the target speech must be played at a louder level compared to the interferers in order for correct identification of target stimuli [21, 56]. This may be attributed to poor temporal and spectral resolution encoding, as well as the limited number of electrodes in the CI device [28, 58, 59]. When matched for HA, the younger BiCI and NH groups only differed in the Quiet and Symmetrical conditions. This suggests that children with BiCIs are able to use monaural head shadow cues to achieve similar performance to their NH peers, when matched for HA, in only the Asymmetrical condition. This is consistent with research showing that CI users use primarily monaural head shadow cues to function in noisy environments [39, 40, 49, 56]. When we matched NH and BiCI groups for CA (with the BiCI group having approximately one year less HA), the BiCI users had significantly worse performance (i.e. higher SRTs) in all of the conditions except the Front condition. The lack of significant difference in the Front condition may be caused by the difficulty in segregating target and interfering speech that listeners with NH also show when target and interferers are both co-located and spectrally similar.

When matched for either HA (NH-A and BiCI-A) or CA (NH-B and BiCI-B), the NH groups demonstrated more SRM with either set of stimuli. Even when matching the younger NH group with the older BiCI group, in which the BiCI group was chronologically approximately 3 years older and had approximately 2 years more auditory experience, the NH group demonstrated significantly more SRM. This significant difference in SRM between all groups of NH and BiCI listeners suggests that cues necessary for spatial hearing are not delivered by current CI processor technology [14]. Additionally, the independent functionality of the current processors likely cause binaural cues to be either absent or inconsistent in BiCI users [60].

We also compared SRTs and SRM with the same-sex stimuli *within* each hearing type (young vs. old) to investigate whether speech understanding is improved as children mature. No significant differences in SRTs were found between younger and older NH children in the Asymmetrical condition, suggesting that when both monaural and binaural cues are available, even with limited spectral cues, NH children as young as 4 years of age are able to significantly benefit from spatial separation of the target and interferers. However, consistent with our previous findings [23], SRTs were higher for the older NH children in only the Symmetrical condition, and all conditions with interferers were better for adults than for the older NH children. This indicates that when monaural head shadow is reduced and listeners must rely more heavily on binaural cues, older NH children are able to identify the target talker at a lower level than younger NH children. Unlike in the NH children groups, no differences in SRTs in any conditions were shown between the younger and older BiCI groups. A lack of differences in SRTs in BiCI groups suggests that as children gain more experience with their CIs, the ability to segregate the target source at a lower level relative to the interferer does not significantly improve. Therefore, it is most likely the limitations of the CI device, rather than experience, which limit access of BiCI users to cues that aid in identification of the target source in noisy environments [60, 61].

Although age effects were shown for SRTs in the NH groups, no significant effects of age were shown for SRM (asymmetrical or symmetrical) in either the NH or BiCI groups. There was an overall trend in the NH listeners for older children with NH to demonstrate more SRM than the younger NH group, but significance was not reached. Presumably, the effects of SRT were not strong enough to affect SRM because SRM is a derived quantity from SRTs in the

Front and spatially separated conditions. This is consistent with Misurelli and Litovsky [23], where we used different-sex stimuli, in that no significant within hearing type age effects of SRM were found.

V. CONCLUSIONS

(1) For all groups, speech reception thresholds (SRTs) were better when listening in Quiet than with interferers.

(2) On average, SRTs for all groups in all conditions were poorer when both the target and the interferers were same-sex talkers, compared to when the target was a male talker and the interferers were female talkers. This suggests that using same-sex target and interfering stimuli introduces more masking.

(3) When $\mathrm{SRM}_{\mathrm{Asymmetrical}}$ and $\mathrm{SRM}_{\mathrm{Symmetrical}}$ with the same-sex stimuli are compared within each listening group, all groups showed significantly more asymmetrical SRM. These results indicate that with same-sex stimuli, the Symmetrical condition—where binaural cues are necessary to segregate sources—was significantly more difficult than the Asymmetrical condition, in which both monaural cues and binaural cues are available.

(4) Within-group comparisons showed a trend for more SRM with the same-sex than with the different-sex stimuli, suggesting that spatial cues were more beneficial in an environment intended to create more informational masking. Although there is a trend for more SRM with same-sex stimuli, statistically significant results between the two interferers were often not reached. This may be due to the large amount of within-group variability for both the NH and BiCI listeners.

(5) When matched for CA (older groups), HA (younger groups), and when comparing children with BiCIs who are chronologically older but have less hearing experience than the younger NH children, in all instances more SRM was shown in the NH groups. Because HA does not appear to explain the difference between children with NH and children with BiCIs, it may be that spectral and temporal fine structure binaural cues accessible only to NH listeners may be the key factor in the ability to hear target speech in the presence of interferers.

TABLES

Table 1: Subject Demographics

	Subject Code	Chronological Age (yr;mo)	Hearing Age (yr;mo)	Bilateral experience (yr;mo)	Etiology	Age first CI activation (yr;mo)	Age second CI activation (yr;mo)	Time between first and second CI (yr;mo)	First CI (device, ear)	Second CI (device, ear)
BiCI-A (younger BiCI users)	*CIBW	7;0	6;0	3;3	Connexin-26	1;0	3;9	2;9	N24C, R	Freedom Contour, L
	CICA	6;6	4;1	4;1	Unknown, late ID	2;5	2;5	simultaneous	Med-El Pulsar, R	Med-El Pulsar, L
	CIEH	5;0	3;11	3;11	Hereditary	1;1	1;1	simultaneous	Nucleus Freedom, R	Nucleus Freedom, L
	CIET	6;4	5;6	1;7	EVAS/sudden HL	4;9	4;9	simultaneous	Med-EL Sonata, R	Med-EL Sonata, L
	CICF	6;7	6;2	4;3	Meningitis	1;6	2;4	0;10	Freedom Contour, R	Freedom Contour, L
	CICN	5;11	4;7	3;0	Connexin-26	1;4	2;11	1;6	Nucleus Freedom, R	Nucleus Freedom, L
	CIDW	7;0	5;0	2;0	Unknown, late ID	2;3	5;0	2;9	Advanced Bionics HiRes 90K, L	Advanced Bionics HiRes 90K, R
		mean= 6;4	**mean= 5;0**							
BiCI-B (older BiCI users)	CIDJ	9;0	7;4	3;11	Hereditary	1;8	5;1	3;5	N24C, R	Nucleus Freedom, L
	*CIBU	8;3	7;2	3;3	Connexin-26	1;1	5;0	3;11	Med-El Tempo, L	Med-EL Sonata, L
	*CIDQ	8;8	7;11	4;5	Unknown	0;10	4;3	3;6	N24C, R	Nucleus Freedom, L
	*CICY	7;11	6;11	3;1	Unknown	1;0	4;8 (reimplanted)	3;8	Advanced Bionics HiRes 90K, R	Advanced Bionics HiRes 90K, L
	CIEF	8;0	6;9	3;2	Unknown	1;4	4;10	3;6	Nucleus Freedom, R	Nucleus Freedom, L
	CIEK	7;5	6;8	2;2	EVAS/Progressive HL	4;9	5;2	0;5	Advanced Bionics HiRes 90K, L	Advanced Bionics HiRes 90K, R
		mean= 8;5	**mean= 7;2**							
NH-A (younger NH)	*CKX	6;5								
	*CNJ	6;5								
	COU	4;7								
	CMG	4;11								
	CPE	4;7								
	CKS	6;2								
	CPG	4;3								
	*CNS	4;2								
		mean= 5;2								
NH-B (older NH)	*CKB	8;7								
	*CKG	8;8								
	*CNW	8;3								
	*CNU	7;11								
	*CNB	6;7								
	*CNG	9;10								
	*CNH	8;8								
	*CNV	7;6								
		mean= 8;3								

* children who also particpated in Misurelli & Litovsky, 2012

Table 2

Table II. Results of planned comparisons between male and female SRTs (*significance)

Group	Front	Asymmetrical	Symmetrical
BiCI-A	*p=0.003	p=0.47	p=0.17
BiCI-B	*p=0.01	p=0.13	p=0.02
NH-A	*p<0.001	p=0.16	*p=0.013
NH-B	*p=0.005	*p=0.01	*p=0.008
Adult	*p=0.002	*p=0.008	p=0.08

Table 3

Table III. Results of planned comparisons for between group SRTs (*significance)

Groups	Quiet	Front	Asymmetrical	Symmetrical
BiCI-A vs. BiCI-B	p=0.384	p=0.385	p=0.193	p=0.71
NH-A vs. NH-B	p=0.09	p=0.14	p=0.041	*p=0.007
NH-B vs. Adult	p=0.02	*p<0.001	*p<0.001	*p<0.001
BiCI-A vs. NH-A (HA equivalent and CA greater in BiCI than NH by 1;2)	*p=0.002	p=0.24	*p=0.009	p=0.09
BiCI-B vs. NH-B (CA equivalent and HA greater in NH than BiCI by 1;1)	*p<0.001	p=.09	*p<0.001	*p<0.001
BiCI-B vs. NH-A (HA greater in BiCI than in NH by ~2yrs)	p=0.02	p=0.68	p=0.013	*p=0.012

Table 4

Table IV. Results of planned comparisons for between group SRM (*significance)

Groups	Asymmetrical (male)	Symmetrical (male)	Asymmetrical (female)	Symmetrical (female)
BiCI-A vs. BiCI-B	p=0.76	p=0.18	p=0.96	p=0.88
NH-A vs. NH-B	p=0.42	p=0.07	p=0.96	p=0.36
NH-B vs. Adult	p=0.07	p=0.24	p=0.54	p=0.20
BiCI-A vs. NH-A (HA equivalent and CA greater in BiCI than NH by 1;2)	*p=0.001	p=0.20	*p=0.001	*p=0.004
BiCI-B vs. NH-B (CA equivalent and HA greater in NH than BiCI by 1;1)	*p=0.001	*p=0.001	*p=0.003	*p=0.006
BiCI-B vs. NH-A (HA greater in BiCI than in NH by ~2yrs)	*p=0.008	*p=0.02	*p=0.003	*p=0.01

FIGURES

Figure 1

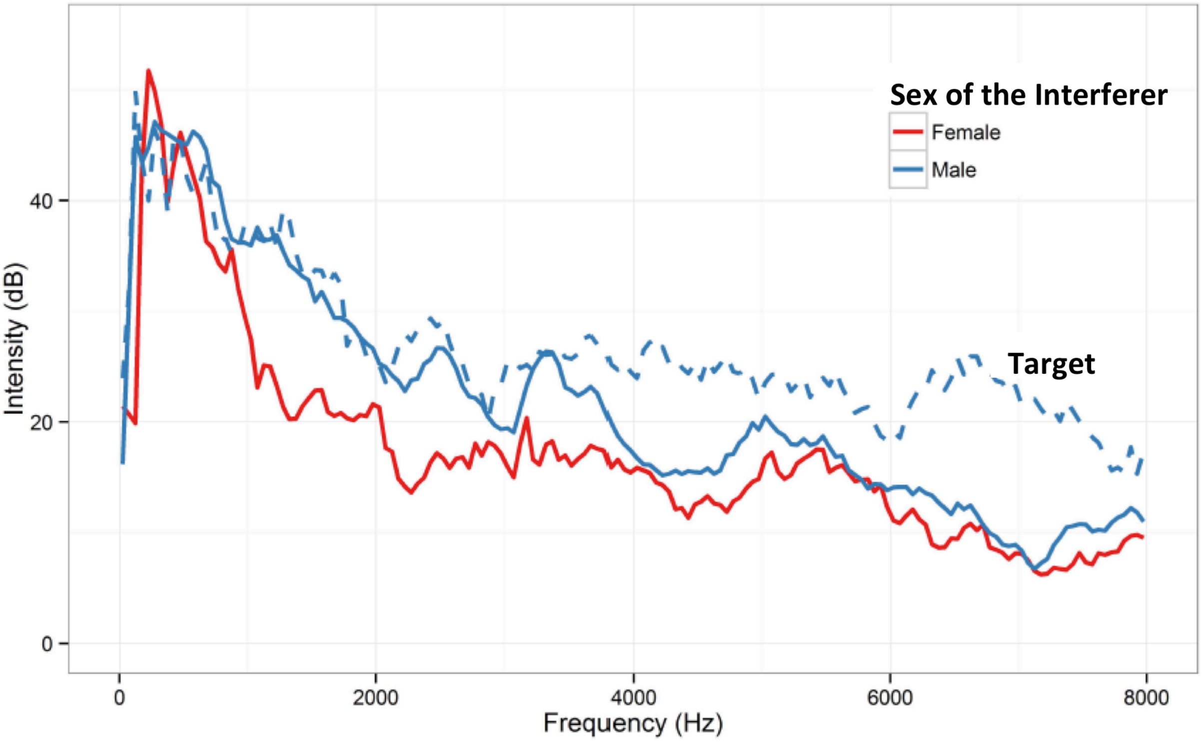

FIG. 1. The overall average spectrum are plotted for the target and interfering stimuli. The dashed blue line represents the target stimuli. The solid lines represent the interfering stimuli (red line=female interferers, blue line=male interferers).

Figure 2

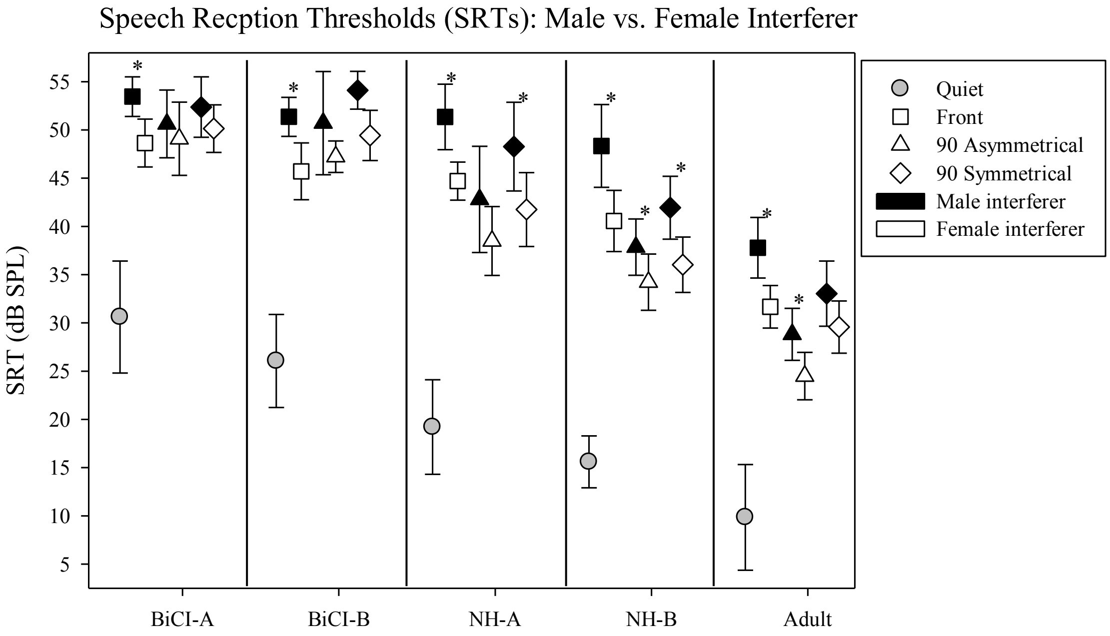

FIG. 2. Mean (±SD) SRTs are plotted for all groups in each condition with both the male and female interferers. The gray circles represent SRTs in the Quiet condition. Filled symbols represent SRTs with the male interferers. Open symbols represent SRTs with the female interferers. Within each group, each condition is represented by a different symbol (square: Front, triangle: 90A, diamond: 90S).

Figure 3

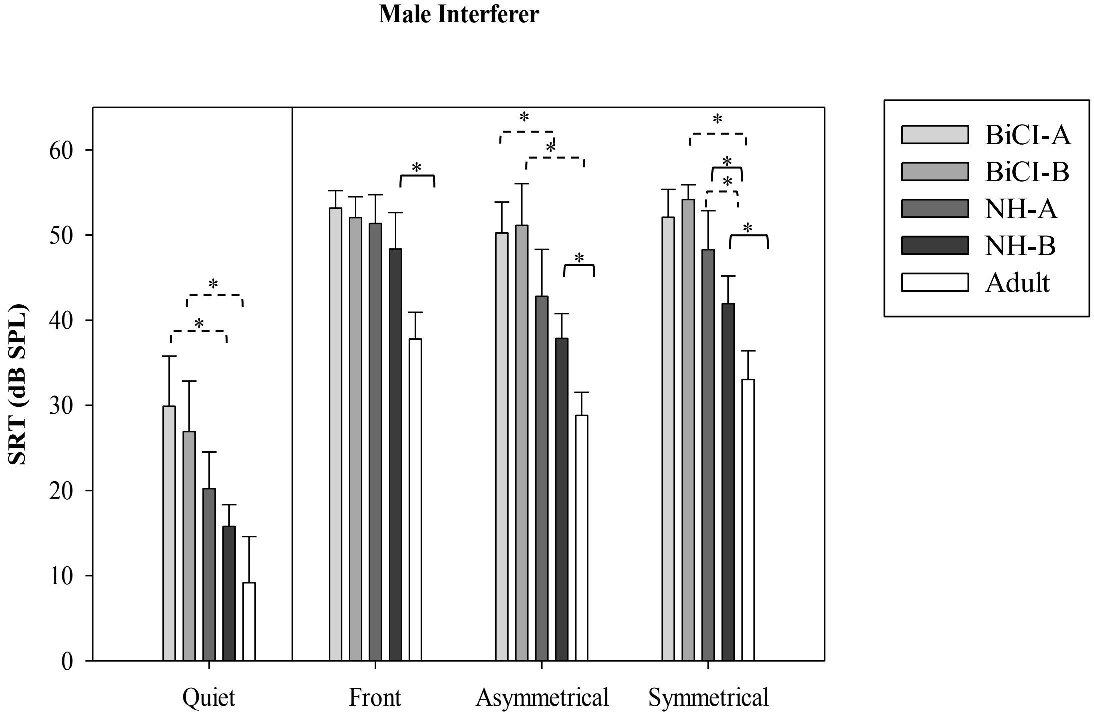

FIG. 3. Mean (±SD) SRTs are shown for each group in each condition (i.e., Quiet, Front, Asymmetrical, Symmetrical). Significant differences (*p* < 0.0125) are bracketed and indicated with an asterisk (*). Solid brackets indicate significant differences within hearing type (i.e., A vs B). Dashed brackets indicate differences between hearing type (i.e., NH vs BiCI).

Figure 4

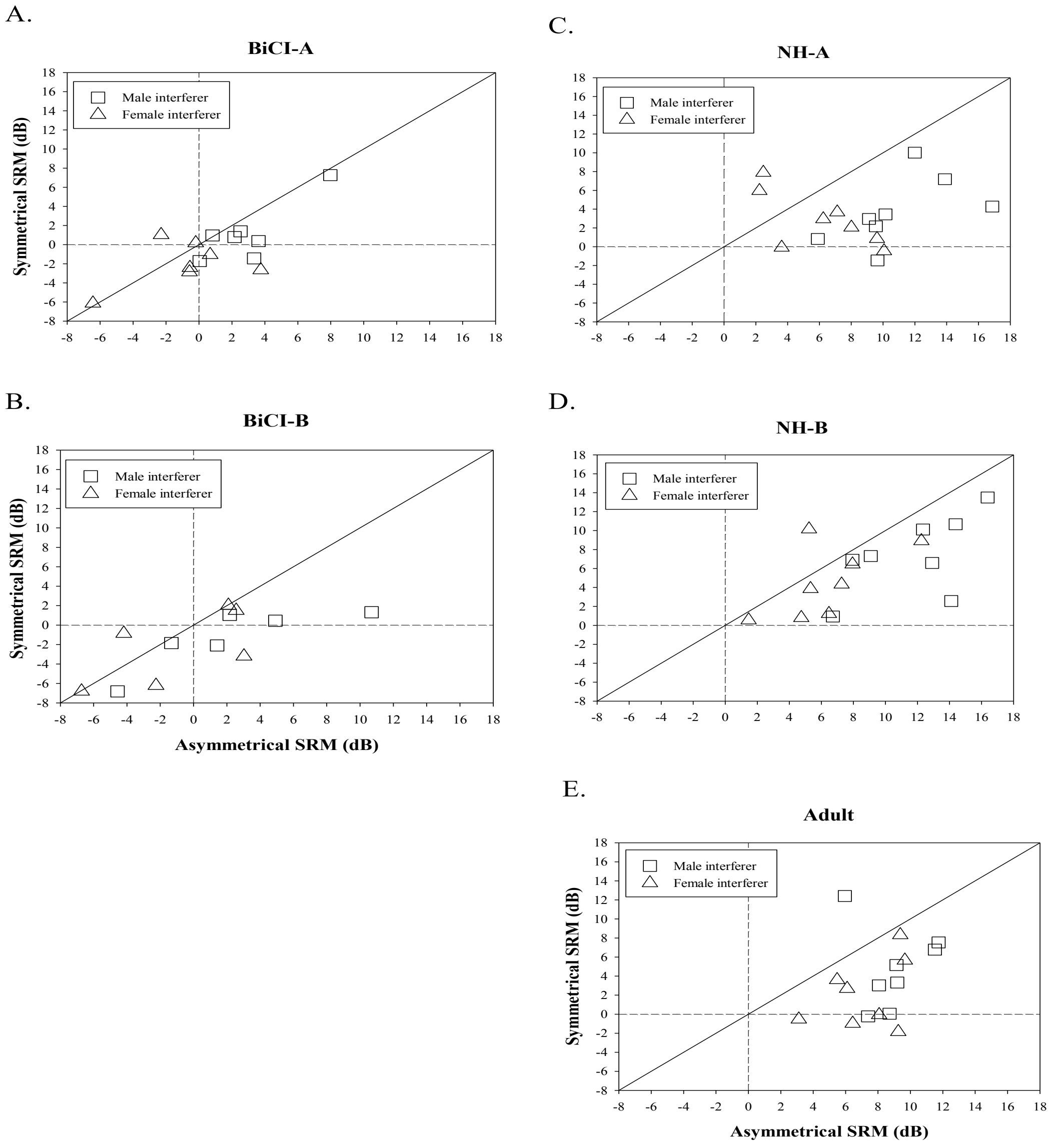

FIG. 4. SRM values for Asymmetrical and Symmetrical conditions are compared for each individual listener group (A: BiCI-A, B: BiCI-B, C: NH-A, D: NH-B, E: Adult). Each data point represents SRM for an individual listener. The squares represent SRM with the male interferers, and the triangles represent SRM with the female interferers. The diagonal line corresponds to equivalent SRM in the Asymmetrical and Symmetrical conditions.

Figure 5

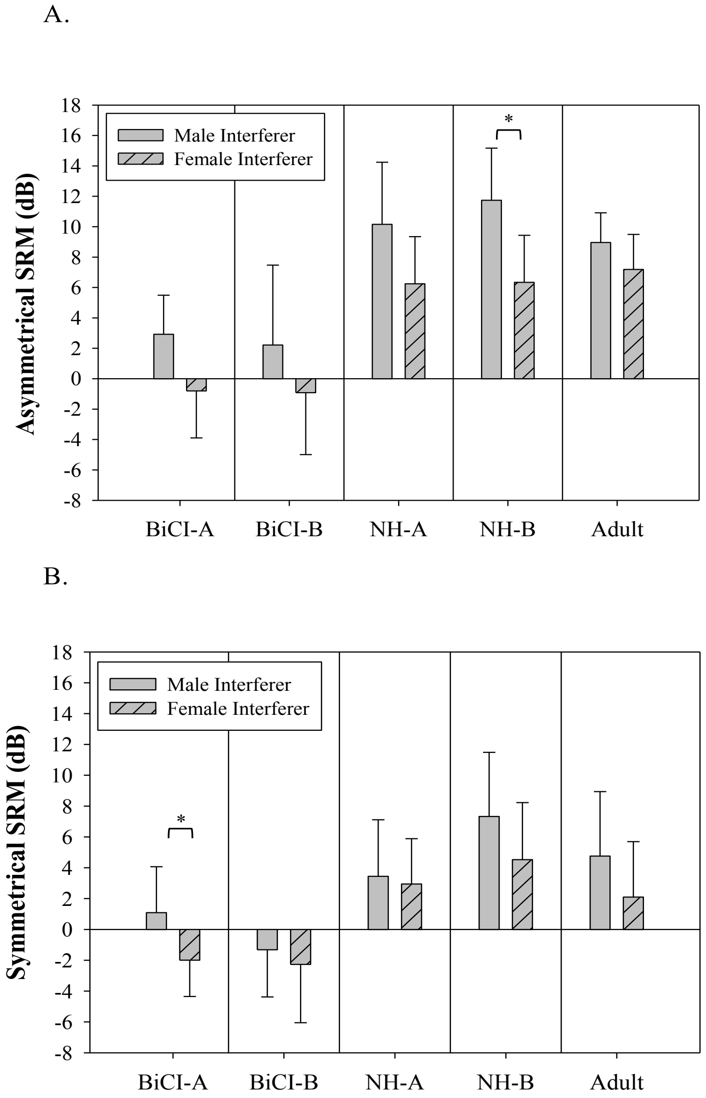

FIG. 5. Mean (±SD) SRM values for both the Asymmetrical (4A) and Symmetrical (4B) conditions are plotted for each group (i.e., BiCI-A, BiCI-B, NH-A, NH-B, Adult). Solid bars represent SRM with male interferers. Dashed bars represent SRM with female interferers. Significant differences ($p < 0.05$) are highlighted (*).

Figure 6

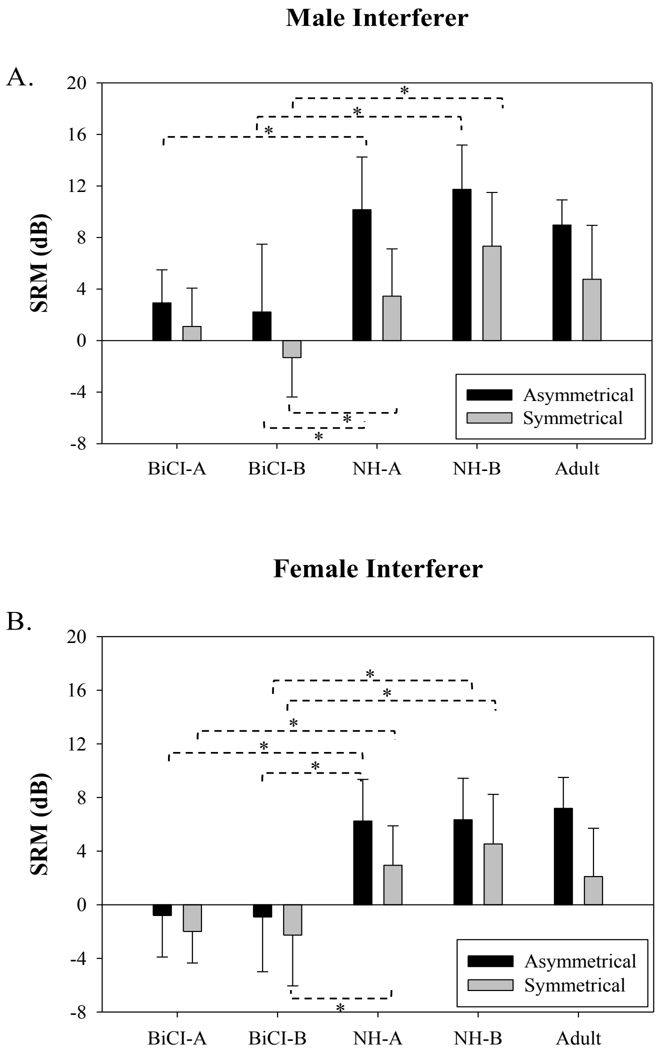

FIG. 6. Mean (±SD) SRM values are plotted for each group (i.e., BiCI-A, BiCI-B, NH-A, NH-B, Adult) for both the male (6A) and female (6B) interferers. Black bars represent Asymmetrical SRM. Gray bars represent Symmetrical SRM. Significant differences ($p < 0.025$) are bracketed and indicated with an asterisks (*). Dashed brackets indicate difference between hearing type (i.e., NH vs BiCI).

VI. REFERENCES

1. E.C., C., *Some experiments on the recognition of speech, with one and with two ears.* J. Acoust. Soc. Am., 1953. **25**: p. 975-979.

2. Watson, C.S., W.J. Kelly, and H.W. Wroton, *Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty.* J Acoust Soc Am, 1976. **60**(5): p. 1176-86.

3. Neff, D.L. and T.M. Dethlefs, *Individual differences in simultaneous masking with random-frequency, multicomponent maskers.* J Acoust Soc Am, 1995. **98**(1): p. 125-34.

4. Lutfi, R.A., *A model of auditory pattern analysis based on component-relative-entropy.* J Acoust Soc Am, 1993. **94**(2 Pt 1): p. 748-58.

5. Durlach, N.I., et al., *Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity.* J Acoust Soc Am, 2003. **114**(1): p. 368-79.

6. Watson, C.S., *Some comments on informational masking.* Acta Acustica United with Acustica, 2005. **91**(3): p. 502-512.

7. Hawley, M.L., R.Y. Litovsky, and H.S. Colburn, *Speech intelligibility and localization in a multi-source environment.* J Acoust Soc Am, 1999. **105**(6): p. 3436-48.

8. Brungart, D.S., *Informational and energetic masking effects in the perception of two simultaneous talkers.* J Acoust Soc Am, 2001. **109**(3): p. 1101-9.

9. Freyman, R.L., et al., *The role of perceived spatial separation in the unmasking of speech.* J Acoust Soc Am, 1999. **106**(6): p. 3578-88.

10. Cullington, H.E. and F.G. Zeng, *Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects.* J Acoust Soc Am, 2008. **123**(1): p. 450-61.

11. Jones, J.A. and R.L. Freyman, *Effect of priming on energetic and informational masking in a same-different task.* Ear Hear, 2012. **33**(1): p. 124-33.

12. Brungart, D.S., et al., *Informational and energetic masking effects in the perception of multiple simultaneous talkers.* J Acoust Soc Am, 2001. **110**(5 Pt 1): p. 2527-38.

13. Arbogast, T.L., C.R. Mason, and G. Kidd, Jr., *The effect of spatial separation on informational and energetic masking of speech.* J Acoust Soc Am, 2002. **112**(5 Pt 1): p. 2086-98.

14. Culling, J.F., M.L. Hawley, and R.Y. Litovsky, *The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources.* J Acoust Soc Am, 2004. **116**(2): p. 1057-65.

15. Hawley, M.L., R.Y. Litovsky, and J.F. Culling, *The benefit of binaural hearing in a cocktail party: effect of location and type of interferer.* J Acoust Soc Am, 2004. **115**(2): p. 833-43.

16. Freyman, R.L., U. Balakrishnan, and K.S. Helfer, *Spatial release from informational masking in speech recognition.* J Acoust Soc Am, 2001. **109**(5 Pt 1): p. 2112-22.

17. Arbogast, T.L., C.R. Mason, and G. Kidd, Jr., *The effect of spatial separation on informational masking of speech in normal-hearing and hearing-impaired listeners.* J Acoust Soc Am, 2005. **117**(4 Pt 1): p. 2169-80.

18. Culling, J.F., K.I. Hodder, and C.Y. Toh, *Effects of reverberation on perceptual segregation of competing voices.* J Acoust Soc Am, 2003. **114**(5): p. 2871-6.

19. Peissig, J. and B. Kollmeier, *Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners.* J Acoust Soc Am, 1997. **101**(3): p. 1660-70.

20. Johnstone, P.M. and R.Y. Litovsky, *Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults.* J Acoust Soc Am, 2006. **120**(4): p. 2177-89.

21. Litovsky, R.Y., P.M. Johnstone, and S.P. Godar, *Benefits of bilateral cochlear implants and/or hearing aids in children.* Int J Audiol, 2006. **45 Suppl 1**: p. S78-91.

22. Litovsky, R.Y., *Speech intelligibility and spatial release from masking in young children.* J Acoust Soc Am, 2005. **117**(5): p. 3091-9.

23. Misurelli, S.M. and R.Y. Litovsky, *Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry.* J Acoust Soc Am, 2012. **132**(1): p. 380-91.

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

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

26. Garadat, S.N. and R.Y. Litovsky, *Speech intelligibility in free field: Spatial unmasking in preschool children.* Journal of the Acoustical Society of America, 2007. **121**(2): p. 1047-1055.

27. Runge, C.L., et al., *Aiding and occluding the contralateral ear in implanted children with auditory neuropathy spectrum disorder.* J Am Acad Audiol, 2011. **22**(9): p. 567-77.

28. Fu, Q.J., S. Chinchilla, and J.J. Galvin, *The role of spectral and temporal cues in voice gender discrimination by normal-hearing listeners and cochlear implant users.* J Assoc Res Otolaryngol, 2004. **5**(3): p. 253-60.

29. Fu, Q.J., et al., *Voice gender identification by cochlear implant users: the role of spectral and temporal resolution.* J Acoust Soc Am, 2005. **118**(3 Pt 1): p. 1711-8.

30. Cleary, M. and D.B. Pisoni, *Talker discrimination by prelingually deaf children with cochlear implants: preliminary results.* Ann Otol Rhinol Laryngol Suppl, 2002. **189**: p. 113-8.

31. van Hoesel, R.J., *Exploring the benefits of bilateral cochlear implants.* Audiol Neurootol, 2004. **9**(4): p. 234-46.

32. Litovsky, R.Y., et al., *Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory.* J Am Acad Audiol, 2012. **23**(6): p. 476-94.

33. Wightman, F.L. and D.J. Kistler, *Informational masking of speech in children: effects of ipsilateral and contralateral distracters.* J Acoust Soc Am, 2005. **118**(5): p. 3164-76.

34. Wightman, F., D. Kistler, and D. Brungart, *Informational masking of speech in children: auditory-visual integration.* J Acoust Soc Am, 2006. **119**(6): p. 3940-9.

35. Lutfi, R.A., et al., *One factor underlies individual differences in auditory informational masking within and across age groups.* Percept Psychophys, 2003. **65**(3): p. 396-406.

36. Leibold, L.J. and A.Y. Bonino, *Release from informational masking in children: effect of multiple signal bursts.* J Acoust Soc Am, 2009. **125**(4): p. 2200-8.

37. Wightman, D.J. Kistler, and A. O'Bryan, *Individual differences and age effects in a dichotic informational masking paradigm.* J Acoust Soc Am, 2010. **128**(1): p. 270-9.

38. Nopp, P., P. Schleich, and P. D'Haese, *Sound localization in bilateral users of MED-EL COMBI 40/40+ cochlear implants.* Ear Hear, 2004. **25**(3): p. 205-14.

39. van Hoesel, R.J. and R.S. Tyler, *Speech perception, localization, and lateralization with bilateral cochlear implants.* J Acoust Soc Am, 2003. **113**(3): p. 1617-30.

40. Litovsky, R.Y., A. Parkinson, and J. Arcaroli, *Spatial hearing and speech intelligibility in bilateral cochlear implant users.* Ear Hear, 2009. **30**(4): p. 419-31.

41. Zurek, P.M., *A note on onset effects in binaural hearing.* J Acoust Soc Am, 1993. **93**(2): p. 1200-1.

42. Rothauser, E., et al., *IEEE recommended practice for speech quality measurements.* IEEE Trans. Audio Electroacoust, 1969. **17**(3): p. 225-246.

43. Garadat, S.N. and R.Y. Litovsky, *Speech intelligibility in free field: spatial unmasking in preschool children.* J Acoust Soc Am, 2007. **121**(2): p. 1047-55.

44. Wichmann, F.A. and N.J. Hill, *The psychometric function: I. Fitting, sampling, and goodness of fit.* Percept Psychophys, 2001. **63**(8): p. 1293-313.

45. Wichmann, F.A. and N.J. Hill, *The psychometric function: II. Bootstrap-based confidence intervals and sampling.* Percept Psychophys, 2001. **63**(8): p. 1314-29.

46. Qin, M.K. and A.J. Oxenham, *Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers.* J Acoust Soc Am, 2003. **114**(1): p. 446-54.

47. Cullington, H.E. and F.G. Zeng, *Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification.* Ear Hear, 2011. **32**(1): p. 16-30.

48. Cleary, M., D.B. Pisoni, and K.I. Kirk, *Influence of voice similarity on talker discrimination in children with normal hearing and children with cochlear implants.* J Speech Lang Hear Res, 2005. **48**(1): p. 204-23.

49. Loizou, P.C., et al., *Speech recognition by bilateral cochlear implant users in a cocktail-party setting.* J Acoust Soc Am, 2009. **125**(1): p. 372-83.

50. Oh, E.L. and R.A. Lutfi, *Effect of masker harmonicity on informational masking.* J Acoust Soc Am, 2000. **108**(2): p. 706-9.

51. Assmann, P.F. and Q. Summerfield, *Modeling the perception of concurrent vowels: vowels with different fundamental frequencies.* J Acoust Soc Am, 1990. **88**(2): p. 680-97.

52. Zeng, F.G., *Temporal pitch in electric hearing.* Hear Res, 2002. **174**(1-2): p. 101-6.

53. Stickney, G.S., et al., *Cochlear implant speech recognition with speech maskers.* J Acoust Soc Am, 2004. **116**(2): p. 1081-91.

54. Moore, J.K. and F.H. Linthicum, Jr., *The human auditory system: a timeline of development.* Int J Audiol, 2007. **46**(9): p. 460-78.

55. Boothroyd, A., *Auditory development of the hearing child.* Scand Audiol Suppl, 1997. **46**: p. 9-16.

56. Van Deun, L., A. van Wieringen, and J. Wouters, *Spatial speech perception benefits in young children with normal hearing and cochlear implants.* Ear Hear, 2010. **31**(5): p. 702-13.

57. Jones and R.Y. Litovsky, *A cocktail party model of spatial release from masking by both noise and speech interferers.* J Acoust Soc Am, 2011. **130**(3): p. 1463-74.

58. Friesen, L.M., et al., *Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants.* J Acoust Soc Am, 2001. **110**(2): p. 1150-63.

59. Shannon, R.V., J.J. Galvin, 3rd, and D. Baskent, *Holes in hearing.* J Assoc Res Otolaryngol, 2002. **3**(2): p. 185-99.

60. van Hoesel, R.J. and R.Y. Litovsky, *Statistical bias in the assessment of binaural benefit relative to the better ear.* J Acoust Soc Am, 2011. **130**(6): p. 4082-8.

61. Litovsky, R.Y., *Review of recent work on spatial hearing skills in children with bilateral cochlear implants.* Cochlear Implants Int, 2011. **12 Suppl 1**: p. S30-4.

Chapter 4: Auditory source segregation in normal hearing children and adults and the effects of executive function

ABSTRACT

In noisy environments, it is difficult to attend to a talker while simultaneously ignoring background speech and noise. Compared to adults, children are less able to extract target speech from interfering noise and demonstrate greater variability in performance on source segregation tasks. Little is known about factors of executive function that could help account for this variability. Four groups of normal hearing (NH) listeners were tested, ages 7-23 years, on a complex speech-in-noise task intended to elicit mechanisms of executive function. Target sentences were either coherent or anomalous; participants listened to target speech in the presence of speech or noise interferers either co-located or spatially separated with the target speech at each of four signal-to-noise ratios (SNRs). Improvement in performance due to spatial separation of sources is known as spatial release from masking (SRM). Results showed that SRM was largest for all groups at the negative SNRs, and that performance on the task improved with age. Coherent semantic content was advantageous at the least favorable SNRs, especially when spatial cues were not available. Results also showed that measures of executive function were all predictors of SRM at the least favorable SNR when no other demographic predictors were considered.

I. INTRODUCTION

In complex auditory environments it is difficult to attend to one speaker while simultaneously disregarding irrelevant background speech and noise. This ability to segregate sources and selectively attend to a target speaker enables normal hearing (NH) adults to function relatively well in complex auditory environments compared to children with NH. Much of the research to date has focused on peripheral sensory processes affecting the ability to identify closed-set sentences or one-word target speech amongst interfering speech and environmental noise. However, with the emergence of what has come to be called “cognitive hearing science,” more recent work has begun to examine how “bottom-up” peripheral processes interact with “top-down” central processing in these complex listening situations [1]. It has been shown that in these situations, both auditory and cognitive factors play a role in the ability to segregate and focus on the signal of interest [2]. The current study investigated the ability to hear complex speech in noise, either spatially co-located or separated, for a large age range of NH listeners (7-23 years of age), and examined the relationship of executive function in these environments.

A. Masking

The deleterious effects of auditory masking have the potential to hinder successful communication in noisy situations. When listening to speech in noise, two types of auditory masking can occur: energetic or informational (either individually or in combination). Energetic masking occurs at the level of the peripheral auditory system when the target and interfering stimuli contain energy that overlaps in the frequency and time domain [3, 4]. This type of masking has been described by auditory filter models. These models are based on well-studied demonstrations that the human peripheral auditory system is essentially comprised of a series of overlapping band-pass filters [5, 6]. When an incoming signal reaches our ears, we identify and attend to the filter which has the highest likelihood of containing the signal [7].

Informational masking is thought to be mediated by auditory mechanisms beyond the periphery; stimuli in such masking experiments are characterized by similarity of the target and interferers, or uncertainty as to the particular stimuli to which listeners should attend [3, 8, 9]. In contrast to energetic masking, informational masking cannot be accounted for by peripheral auditory filter models [10]. Additionally, children are more vulnerable to informational masking than adults. That is, relative to adults, children show greater amounts of masking under these conditions compared with the amount of masking seen with energetic masking [11, 12]. This is problematic because children must frequently function in complex auditory environments (e.g., classrooms, cafeterias, playgrounds) where competing sound sources have the characteristics of informational maskers.

B. Release from Masking

The present study focused on measuring release from masking in NH children and adults. Listeners can use a multitude of cues, including visual cues, individual talker characteristics, previous knowledge of the topic, and location cues to help alleviate some difficulty when listening to speech in noise [13]. Location cues will be the focus of the work described here. A listener with NH is able to rely on binaural cues, whereby integration of cues received at the right and left ear occurs. Binaural hearing is especially advantageous when listening in noisy environments in order to make use of differences in timing and level of stimuli between the two ears. There are three acoustic cues that help with source segregation in noise: (1) “monaural head shadow" refers to an effect in which the target and interferer are spatially separated and the signal-to-noise ratio (SNR; level of target relative to the level of the interfering speech) is more favorable in the ear farthest from interferers, due to the fact that the head acts as an acoustic shadow [14]; (2) "binaural squelch" refers to a comparison between monaural and binaural conditions, where target and interferer are spatially separated, and improvement in performance is demonstrated when the ear with the worse SNR is added [15]; and (3) "binaural summation" occurs when the target signal is presented from directly in front of the listener, and the target is easier to hear due to signal reaching both ears at the same time and intensity (and therefore is "summed").

Binaural hearing allows access to cues that are useful when listening to speech in noise. The term spatial release from masking (SRM) is used to refer to the improvement in speech understanding when the target and interfering stimuli are spatially separated as opposed to when they are co-located. SRM in NH adult listeners can be as large as 12-15 dB [16, 17]. SRM is also demonstrated in NH school-aged children, although it is more variable across listeners compared with adults [18-21]. More recently, SRM has been demonstrated in children as young as 2 years of age [22]. This work suggests that although variable, even very young NH children are able to take advantage of spatial cues to aid in listening in noisy environments. Previous work also suggests that both adults and children demonstrate more SRM with informational, rather than energetic, maskers [20, 23]. In the current study, half of the participants listened to target speech amongst either modulated speech-shaped-noise (noise) or amongst a speech interferer, the latter of which was intended to create an environment with more informational masking. It is expected that more masking will be shown in the younger groups, and that older groups will be better, than younger groups, at using spatial cues to help segregate auditory sources. Effects of interferer type on SRM are expected to be most pronounced at the least favorable SNRs, where source segregation is the most difficult.

Another aspect of release from masking that was investigated was the influence of semantic content (coherent vs. anomalous) when listening to speech in noise. Previous studies have used either single-word or closed-set stimuli to investigate the ability of children to hear target speech in the presence of background noise. The current study uses an open-set corpus of sentences that vary in semantic content. Individuals were either presented with sentences that were semantically coherent or semantically anomalous. It has been shown that adults are more accurate at identifying words in contextually meaningful sentences than words in isolation [24]. It was hypothesized that the semantically coherent sentences would be easier to identify, and that listeners will show more SRM with coherent sentences, specifically in conditions with less favorable SNRs. The above was hypothesized because coherent semantic content provides a cue, in addition to spatial separation, which could aid in release from masking.

A. Auditory Development

It is known that children perform worse than adults on source segregation tasks; that is, children need much more favorable SNRs in order to successfully segregate and comprehend target speech [12, 21, 25-27]. For individuals with NH, the peripheral auditory system is structurally intact and functional at birth [28]. Although the basic structures of the peripheral auditory system are intact at birth, both peripheral and central components continue to mature after. For example, neural connections continue to develop such that transmission of auditory signals becomes more efficient, and auditory pathways reorganize [29]. Cognitive processing mechanisms also continue to develop and become more refined [28, 30, 31]. Together, immature peripheral and central processing continues to develop and become more refined throughout adolescence; it is likely that both contribute to the variability on auditory tasks. Thus far, no studies have examined a continuum of NH listeners from childhood to adults using a complex speech-in-noise task. The current study used stimuli that requires processing of complex sounds, which was intended to enable examination of higher-order aspects of auditory processing, and also examine variability across a large age range (i.e. 7-23 years). Measures of executive function, such as memory and attention, which may serve as predictors of performance in complex auditory environments and aid in explaining some sources of variability, were investigated.

D. Executive Function

Executive function is a term that embodies a multitude of cognitive abilities which organize information required for goal-directed, purposeful, and non-automatic behavior. These abilities can include things such as memory, attention, planning, and inhibition [32]. For this study the relationship of executive function and performance on a speech-in-noise task was examined. Working memory involves the conscious storage, manipulation, and integration of information [33-36]. In contrast, short-term memory is an active space for merely storing incoming information. Working memory is often referred to as a mental “work bench” because it holds information that receives immediate attention and processing. It is thought that information held in working memory controls “top-down” cognitive processes and knowledge that aids in modifying “bottom-up” sensory and peripheral signals [37]. When listening to speech-in-noise, the representation of the target signal is degraded due to the interfering speech, and more resources are needed to first recognize and segregate the signal and then encode, store, and process the information [38]. Previous research suggests that a limited working memory span negatively affects ability to segregate sources in noisy environments [39]. A set of experiments conducted by Rabbitt [40] demonstrated an increase in processing demands when stimuli was heard in noise. In his study, Rabitt instructed listeners to repeat sentences in which the first half of the sentence was heard in quiet and the second half of the sentence was either heard in quiet or in noise. Results showed that when the second half of the sentences were presented in noise listeners had a more difficult time correctly repeating the sentence, suggesting that an increase in processing demands (i.e. with a sentence heard in noise) limited the resources available to retain the entire sentence.

It has also been suggested that executive function skills develop over time and influence other skills [41]. Specifically, improvement and development of working memory with age and training has been shown to be associated with improvement in other areas, such as in development of language skills [42-44]. Recent work has also shown that, for adults, working memory is a better predictor of performance on speech-in-noise tasks than measures of general cognitive ability such as IQ [45]. Working memory capacity is hypothesized to influence performance on our speech-in-noise task, where participants must process and verbally repeat complex auditory stimuli in the presence of either interfering speech or noise, such that those with better working memory, or larger working memory capacity, will also perform better on the speech-in-noise task.

Attention is also hypothesized to influence ability to hear speech in noise, such that individuals who are better able to inhibit interfering stimuli and shift attention should benefit more from spatial separation of sources (i.e. SRM). Before information can be stored and manipulated in working memory, auditory information must first be selected and attended to [46]. Attending to an auditory target occurs simultaneously with ignoring the interfering stimuli.

For the purposes of this study, two aspects of attention will be investigated: attention shifting and attention inhibition. When presented with multiple stimuli, listeners must be able to quickly shift their attention to the target of interest while concurrently inhibiting the distracting stimuli. In general, it is hypothesized that the older groups will have higher scores on the tasks of working memory and attention than the younger groups because it is known that executive function continues to develop throughout adolescence.

In summary, there were two overarching goals of the present study. The first goal was to examine source segregation abilities (i.e. SRM) in children with NH. NH listeners were recruited into one of four groups to examine SRM over a close continuum of ages: 7-10, 11-14, 15-17, and 18-23 years of age. It was also of interest to investigate how interferer type and semantic content influenced performance when listening to speech in noise. The second main goal of this study was to identify measures of executive function that may serve as predictors of performance on tasks of source segregation. For children with NH, there is a wide range in ability to selectively attend to a target source in noisy environments, and there is evidence to suggest that variable performance may derive from central processes.

II. METHODS

A. Participants

A total of 92 typically developing native English speakers were recruited to participate in this study. Participants were divided into three groups of NH children between the ages of 7-17 years, and one group of NH adults ages 18-23 years. Groups were divided according to age and school placement: elementary school (7-10 years, n=24); middle-school (11-14 years, n=24); high-school (15-17 years, n=20); and adult (18-23 years, n=24). Prior to testing each subject underwent a hearing screening to screen for hearing sensitivity within normal limits. All participants passed the hearing screening as indicated by a repeatable response at pure-tone air conduction intensities at 20 dB HL for octave frequencies between 250-4000 Hz. Participants completed all testing in 2-3 hours within 1-2 visits. All participants were recruited from the Madison, WI community and University of Wisconsin-Madison undergraduate population.

B. Environment

All auditory testing was conducted in a carpeted standard IAC sound booth (2.8 x 3.25m) with reverberation time (RT_{60}) of 250ms. Participants were seated at a foam-covered desk facing a computer monitor at 0^{o} azimuth in the center of a speaker array, approximately 1.5m from surrounding loudspeakers. All executive function measures were administered in a quiet room that contained a dual-monitor computer, which is required for some of the executive functioning measures (NIH toolbox version 1.0, www.nihtoolbox.org).

C. Auditory Stimuli

Target stimuli were spoken by a male talker and included a total of 200 sentences between 6 and 13 words in length [47]. Sentences were originally recorded using a male speaker of British English, and were re-recorded for this study by a male speaker and a female speaker of Standard American English. Sentences were designed to contain either semantically coherent content (e.g., “the whole sky was full of birds”) (n=100) or semantically anomalous content (e.g., “the success moved to hope the milk) (n=100). Coherent and anomalous sentences were matched for phonological, lexical, and syntactic properties, but the latter lack coherent meaning. The sentences were recorded to be between 1-3 seconds in duration. All sentences were clearly spoken and digitized at a sampling rate of 44.1KHz (all root-mean-square levels equalized).

Interferers were either amplitude modulated speech-shaped noise or speech. The speech-shaped noise was created by taking the long-term spectrum and imposing an amplitude modulation of the coherent and anomalous sentences. The noise was first scaled to the same root-mean-square value as the speech sentences and then the noise was segmented to match the length of the speech competitor. The envelope of the speech interferer was then extracted and used to modulate the noise. Two-talker interferer speech stimuli were created by overlaying two recordings spoken by the same female talker (interfering speech was composed of the sentences as described above).

D. Design and Procedure

I. Auditory Task

Before testing began, each participant was randomly assigned to either the noise or speech interferer. For non-quiet conditions, the interferers were fixed at 55 dB SPL and the level of the target was adjusted accordingly for the desired SNR for each trial. Results were calculated based on percent correct identification of words in the target sentence. For each listener, performance was measured in Quiet (target 0° front, no interferers). In addition, testing was conducted with interferers in two conditions: Front (target and interferer both at 0° front) and a 90 Asymmetrical (target 0° front, interferers at 90° to the right). In each of the conditions with interferers, testing was conducted at four signal-to-noise ratios (SNRs; -16, -8, 0 and +8 dB). In conditions with interferers, the target speech began approximately one second after the onset of the interferer. The interferer continued approximately one second after the offset of the target sentence. Participants were instructed to listen to the male target speaker and ignore the background noise or the female speech, depending on the interferer to which they had been assigned. All participants verbally repeated the target sentences at the offset of the auditory stimuli, and an experimenter sat behind them and transcribed the verbal responses. Audio recordings of verbal responses were also made and analyzed off-line to ensure that all transcriptions were correct. No feedback was provided during testing. Breaks were given intermittently throughout testing in order to keep the participants motivated and alert; in addition, children were reinforced throughout testing with small prizes.

II. Measures of Executive Function

Tests of working memory and attention were administered in order to investigate the relationship between executive function and performance when listening to speech in noise. A non-verbal IQ and expressive vocabulary test were also administered to rule out any extraneous relationships due to these factors.

Two computer-based tests from the National Institutes of Health (NIH) toolbox measures (NIH toolbox version 1.0, www.nihtoolbox.org) were used to assess attention. Participants sat in a quiet room in front of a computer with a dual screen monitor for the following measures: Dimensional Change Card Sort Test (DCCS) and the Flanker Inhibitory Control and Attention Test. The DCCS is a standardized measure of cognitive flexibility. Two target pictures are presented that vary along two dimensions (e.g., shape and color). Participants were instructed to match a series of bivalent test pictures (e.g., yellow balls and blue trucks) to the target pictures, first according to one dimension (e.g., color) and then, after a number of trials, according to the other dimension (e.g., shape). Scoring is done online and based on a combination of accuracy and reaction time. The test takes approximately 4 minutes to administer. The Flanker test is a standardized task that measures both attention and inhibitory control by requiring the participant to focus on a given stimulus while inhibiting attention to flanking stimuli. The participant is presented with a series of arrows and is instructed to indicate, on a keyboard using only their index finger, which way the middle arrow is pointing. The middle stimulus either points in the same direction as the “flankers” (congruent) or in the opposite direction (incongruent). Scoring is done online and based on a combination of accuracy and reaction time. The test takes approximately 3 minutes to administer. For analyses including the Flanker and DCCS the unadjusted scaled scores were used. This score is calculated and given via the output from the toolbox, and compares the score of the participant to those in the entire NIH toolbox’s nationally represented normative sample excluding age, gender, or any other demographic variables. This score was used in order to be able to compare across various ages.

Additionally, three non-computer based tests were administered. Subtests from the *Weschler Intelligence Scale* (Weschler, 1991) were used to assess working and short-term memory. The forward and backward digit span sub-tests were administered to all participants. These tests were presented using live-voice in auditory mode, with an experimenter speaking at a rate of approximately one digit per second [48]. In the forward digit span task, the participant is instructed to repeat a list of digits in the order that they were presented (maximum amount of digits possible=8). In the backward digit span task, the participant is instructed to repeat the list of digits in the reverse order (maximum amount of digits possible=7). Lists begin with two items and increase in length with correct repetition of all the digits in the presented list. Testing is terminated when two lists, at a given length, are repeated incorrectly. Participants receive one point for each correct trial when all digits in a given list are repeated in the correct temporal order; no partial credit is given. Scores for the digit span forward and backward were defined, separately, as the longest length sequence of digits that the participant could accurately repeat. The Kaufman Brief Intelligence test, Second Edition (KBIT-2) was also administered. The KBIT-2 is a standardized test that measures both verbal and nonverbal intelligence. The nonverbal intelligence quotient (IQ) was used for this study. The nonverbal score comprises one test, Matrices, which assesses ability to perceive relationships and complete visual analogies. Lastly, the Expressive Vocabulary Test, Second Edition (EVT-2) was administered. The EVT is a standardized test that measures expressive vocabulary knowledge using two types of items, Labeling and Synonyms. The participant is shown a picture, and must respond with one word that is an acceptable label for the picture, one word that answers a specific question about the picture, or a word that is a synonym for the pictured context. This test requires no reading or writing. For analyses including the KBIT-2 and EVT the raw scores were used. This score was used in order to be able to compare across various age groups.

III. RESULTS and DISCUSSIONS

A. Auditory Speech-in-Noise Task

Within age groups the following were compared: (1) difference in percent correct when the target and interferers were co-located (i.e. Front) verses spatially separated (i.e. 90 Asymmetrical); (2) effect of SNR on SRM; (3) difference in percent correct when the target sentence was either coherent or anomalous; (4) effect of semantic content on SRM; and (5) effect of interferer type on SRM. *Between* age groups, SRM with an increase in chronological age was examined.

Percent correct for each group at each SNR in both the Front and 90 Asymmetrical conditions is shown in Figure 1 (A: noise interferer; B: speech interferer). A paired samples t-test was conducted at each SNR for each group, with each interferer, to compare the percent correct in the Front and 90 Asymmetrical conditions. Bonferroni corrections for 4 comparisons per planned comparison group were applied. Results were the same for both interferers. All groups showed a significant improvement in percent correct in the 90 Asymmetrical condition at the -16 and -8 dB SNRs (results listed in Table I). Significant differences varied between groups for the 0 dB SNR, and no significant differences were found for any of the groups at the +8 dB SNR.

SRM for each SNR was determined by subtracting the percent correct in the Front condition from percent correct in the 90 Asymmetrical condition. Percent correct in the 90 Asymmetrical condition should be higher than percent correct in the Front condition, so a positive SRM would be expected.

To explore within group effects of SNR on SRM, a within subjects ANOVA was conducted for each age group with each interferer to determine if SRM was larger at lower SNRs. There was a significant main effect of SNR for all groups for the noise (7-10 yrs: $F(3,9)=78.8$, $p<0.001$; 11-14 yrs: $F(3,9)=105.0$, $p<0.001$; 15-17 yrs: $F(3, 6)=51.6$, $p<0.001$; 18-23 yrs: $F(3,9)=86.5$, $p<0.001$) and the speech interferer (7-10 yrs: $F(3,9)=43.0$, $p<0.001$; 11-14 yrs: $F(3,9)=95.7$, $p<0.001$; 15-17 yrs: $F(3, 4)=155.9$, $p<0.001$; 18-23 yrs: $F(3,9)=122.1$, $p<0.001$). Pairwise comparisons using a Bonferroni correction for 6 comparisons per group were conducted in order to investigate within group differences of SNR (results are shown in Table II; noise interferer= Table IIa, speech interferer= Table IIb). The results show that significantly more SRM was demonstrated for the -16 vs. 0 dB, -16 vs. +8 dB, -8 vs. 0 dB, and -8 vs. +8 dB SNRs within all groups, for both interferers, suggesting that for all groups spatial separation of sources was most useful when the target speech was played at a lower level than the interferers. Difference in SRM between the -16 vs. -8 dB SNR and the 0 vs. +8 dB SNR was not significant for all groups, for either interferer. This suggests that the difference in benefit of spatially separating the sources between the two least favorable and two most favorable SNRs may not have a large perceptual effect for all groups.

Percent correct comparisons between the coherent and anomalous sentences are shown in Figure 2 (A: noise interferer, Front, B: noise interferer, 90 Asymmetrical, C: speech interferer, Front, D: speech interferer, 90 Asymmetrical). A paired samples t-test was conducted, at every SNR for each group with each interferer in both the Front and 90 Asymmetrical conditions, to compare the percent correct with coherent sentences versus percent correct with anomalous sentences. Bonferroni corrections for 5 comparisons per planned comparison group were applied. (results displayed in Table IIIa=Front, Table IIIb=90 Asymmetrical). All groups demonstrated a trend for higher percent correct for coherent versus anomalous sentences in both conditions at all SNRs, significant differences varied between groups and conditions.

Figure 3 shows SRM for the semantically coherent and semantically anomalous sentences separately for each group at each SNR. Panel A shows SRM for the noise interferer and Panel B shows SRM for the speech interferer. Paired t-tests were conducted to compare SRM with coherent vs. anomalous sentences (results shown in Table IV). Bonferroni corrections for 4 comparisons per planned comparison group were applied. With the -16 dB SNR, all groups, except the 7-10 year olds with the speech interferer, showed significantly more SRM with the coherent sentences. This suggests that when the target is much lower than the interferers spatial separation of sources and semantic content both have an effect on performance, regardless of interferer. Lack of significance in the youngest group is most likely due to the within group variability. For the -8, 0, and +8 dB SNRs, more significant differences in SRM were demonstrated for the speech interferer rather than the noise interferer.

Figure 4 shows SRM within each interferer separately for each group (A: 7-10 years, B: 11-14 years, C: 15-17 years, D: 18-23 years). Independent t-tests were conducted to compare SRM for the noise and speech interferers. Bonferroni corrections for 4 comparisons per planned comparison group were applied. Within the two youngest groups (7-10 yrs and 11-14 yrs) there were no significant differences in SRM demonstrated between interferer types, suggesting equal benefit with both. There was a trend for more SRM with the speech interferer at the most difficult SNR, -16 dB, although only the oldest group showed significantly more ($p<0.001$). Results for SRM at the -16 dB SNR suggest that, although variability exists, spatial separation of sources in the least favorable listening conditions is most advantageous for speech interferers. The 11-14, 15-17, and 18-23 year olds showed significantly more SRM for the noise interferer at the -8 dB SNR ($p=0.006$, $p=0.003$, $p<0.001$ respectively).

Between groups, we were interested in examining SRM with an increase in chronological age. Figure 5 (5A: noise interferers, 5B: speech interferers) compares SRM between groups in each condition. To explore *between group* effects, a one-way analysis of variance (ANOVA) was conducted to compare effect of group on SRM for each SNR separately, totaling four ANOVAs per interferer type. There was a significant main effect of group, for both interferer types, for only the -16 dB (noise: $F(3,41)=6.3$, $p=0.001$, speech: $F(3,43)=7.5$, $p<0.001$) and the -8 dB (noise: $F(3,41)=4.0$, $p=0.014$, speech: $F(3,43)=3.7$, $p=0.019$) SNRs. The pairwise comparisons, with a Bonferroni correction for 4 comparisons, for the -16 dB SNR showed the following: with the noise interferer, 7-10 year olds demonstrated significantly lower SRM ($^*p=0.001$) than the 15-17 year olds; with the speech interferer, 7-10 year olds demonstrated significantly lower SRM than both the 15-17 year olds ($^*p=0.006$) and the 18-23 year olds ($^*p<0.001$). The pairwise comparisons for -8 dB SNR showed no significant differences between any of the groups after a Bonferroni correction for 4 comparisons was applied. It is most likely that the within group variability limited more significant differences. No between group main effects were shown at the 0 or +8 dB SNR.

B. Discussion: Auditory Speech-in-noise Task

This study was the first of its kind to examine SRM using a complex speech-in-noise task, for a large age range of NH children and adults. Performance when listening to speech in noise was significantly better when the target and interferer were spatially separated (i.e. 90 Asymmetrical) relative to when they were co-located (i.e. Front) in all groups, at SNRs where the target was played at a lower level than the interferers (i.e. -16 and -8 dB). This is consistent with previous literature showing that performance is better in spatially separated vs. co-located conditions for both NH adults [17, 23, 49] and NH children [19-22]. Significant differences between Front and 90 Asymmetrical percent correct varied between groups at the 0 dB SNR, and no significant differences were found at the +8 dB SNR. This is most likely due to the fact that when the target is played at the same level as the interferer or 8 dB louder than the interferer, the target can be clearly heard even in the Front condition, and therefore spatial cues are not as useful. At positive SNRs, performance in the Front and 90 Asymmetrical conditions was nearly identical and near-perfect. Previous work has also shown that at SNR values above +9 dB, performance is near-perfect for NH listeners [50]. The difference in performance in the Front versus 90 Asymmetrical conditions (i.e. higher percent correct in the 90 Asymmetrical condition) demonstrates that NH children and adults were able to take advantage of auditory cues that aid in source segregation. Performance on the speech-in-noise task improved with age, such that percent correct increased for each age group from 7-10 to 18-23 years, especially in the 90 Asymmetrical condition. This is consistent with the literature, suggesting that there is a development component in ability to segregate sources [7, 51].

When comparing within group SRM at different SNRs, SRM was largest for the least favorable SNRs. This is consistent with the NH adult literature showing that spatial cues are most important in the most challenging listening conditions [49, 52].

SRM between interferer types—noise vs. speech—was also compared. When comparing within group SRM between interferer types, significant differences were only found in the 11-14, 15-17, and 18-23 year olds. This result of no effects of interferer type on SRM for the youngest group is consistent with previous work [21]. Although there was a trend for more SRM with the speech interferer (which was intended to introduce more informational masking) at the least favorable SNR (-16 dB), the only significant difference was shown in the adult group. Previous literature suggests that NH adults show more SRM with maskers that have more informational masking [23]. At the -8 dB SNR, all but the 7-10 year olds showed more SRM with the noise interferer. Spatial separation of sources may have been more beneficial for the noise interferer than the speech interferer at this SNR because the spectrum of the noise does not vary over the course of the presentation. The lack of frequency fluctuations in the noise interferer may have rendered spatial cues more advantages relative to the speech interferer, in which, both amplitude and frequency fluctuations occur. With the speech interferer, it is likely that participants were able to use the brief gaps in the temporally fluctuating interferer to “glimpse” portions of the target sentence without the interferer, which helped in segregating the target sentence [53]. It is also possible that at the -8 dB SNR the male target and female interferers were perceptually easier to separate due to the fact that the target and interferers were different-sex talkers. The difference in fundamental frequency between the voices of the two talkers may have created a task that reduced the utility of spatial cues. Because each participant was tested with either the noise or the speech interferer, within group variability likely exists. It would be beneficial to test each participant on both interferer types to be able to better compare between the two interferers.

Another major finding was the trend showing that semantically coherent sentences were easier to identify than semantically anomalous sentences for all groups. It has been shown that listeners demonstrate better performance for sentences with coherent meaning [47, 54, 55]. In the Front condition at -16 dB SNR and in Quiet, there was no difference in performance, with either interferer, for any groups when presented with either the coherent or anomalous sentences. This suggests either that when spatial cues are not available and the SNR is poor enough, or that in an easy listening condition with no interferers present, semantic cues are negated. In the 90 Asymmetrical condition at the least favorable SNR (-16 dB), all groups except the 7-10 year olds (with the speech interferer) showed significantly better performance with the semantically coherent sentences. All groups showed significantly better performance with the semantically coherent sentences at the -8 dB SNR. This demonstrates that when the target and interferers are spatially separated semantic cues can help even at extremely poor SNRs. Together, these results suggest that the availability of spatial cues is more important than semantic cues when the SNR is poor (i.e., no differences in performance with coherent vs. anomalous were shown in the Front condition at the -16 dB SNR). When spatial cues are available at poor SNRs (i.e., 90 Asymemtrical) semantic cues can also help in release from masking. When target and interferers were played at the same level (i.e. 0 dB SNR), performance improved with coherent sentences relative to anomalous sentences for all groups in the Front, but not the 90 Asymmetrical conditions. This suggests that at more favorable SNRs semantic content helps with release from masking for all groups only in more challenging conditions, such as when spatial cues are unavailable (i.e., Front). Additional benefits of semantic content were not shown for all groups at the +8 dB SNR in either condition. This is most likely because the level of the target was favorable relative to the level of the interferer, and for some groups no additional cues were necessary to successfully segregate sources.

SRM for coherent vs. anomalous sentences was also compared, with different patterns of performance shown between each interferer. At the least favorable SNR (i.e. -16 dB), SRM with coherent sentences was significantly higher with both interferers for all but the 7-10 year olds (with the speech interferer). The lack of significant difference between SRM with the coherent and anomalous sentences in the youngest group is likely due to the within group variability. For the -8 dB SNR, significantly more SRM with the coherent sentences was shown only with the speech interferer. This suggests that semantic content has more of an effect for the speech vs. the noise interferer, implying that semantic content can be useful in perceptual segregation of sources particularly for interferers that were intended to create an environment with more informational masking. “Glimpsing” allowed for brief moments when the listener was able to isolate a few words in the target speech and then, with the coherent semantic cues, attempt to fill in the gaps using context through top-down knowledge [56]. Because semantic content is thought to elicit cognitive processes, top-down processes impacted performance on the speech-in-noise task such that coherent sentences were easier to identify (i.e. higher percent correct and larger SRM with coherent sentences) at the least favorable SNRs. At the 0 and +8 dB SNR there was a trend for more SRM with the anomalous sentences, suggesting that when the SNR is favorable spatial separation of sources is more beneficial when other cues, such as semantic content, are not available.

Lastly, one of the main interests was to investigate SRM across a continuum of age groups. NH adults demonstrate as much as 12-15 dB of SRM [17, 23]. This result has not been repeated in studies that have included both children and adults with NH [19-21], most likely because the task that is used in the latter studies (4-alternative forced choice word identification task) is fairly easy for the adult participants even with the target and interferers co-located. The current study aimed to incorporate a more difficult listening task, using open-set sentences that could be effectively used with a continuum of ages from school-aged children to adults. When SRM was compared across age groups, different results were found for the noise and speech interferer. For the noise interferer, there was a trend for SRM to increase from 7-17 years and then stabilize. Differences between groups were only significant for the 7-10 and 15-17 year olds at the -16 dB SNR. A lack of significant differences between all groups in the negative SNRs is most likely due to the large within-group variability, suggesting that some individuals are better able to take advantage of spatially separated sources than others even when matched for age. It is interesting that more SRM with the noise interferer was observed at the -8 dB SNR than the -16 dB SNR. This is most likely due to the fact that at the most difficult SNR (i.e. -16 dB) even when sources were spatially separated (i.e. 90 Asymmetrical) listening was extremely difficult. For the speech interferer at the -16 dB SNR, a systematic increase in SRM is demonstrated from the youngest to the oldest groups. A lack of significant differences between all groups in the -16 dB SNR is likely due to the within group variability. Unlike with the noise interferer, where SRM was greater for all groups at the -8 dB SNR than the -16 dB SNR, within group patterns of performance with speech interferers show that individuals demonstrate a consistent increase in SRM (more SRM) as SNR is worsened. The 7-10 year olds were the only group that demonstrated no additional benefit of spatially separating sources after -8 dB SNR, whereas all other groups showed less SRM for the -8 dB SNR than the -16 dB, suggesting that for the older groups the benefit of SRM may persist beyond even a -16 dB SNR.

The trend for a smaller amount of SRM demonstrated in the younger, compared to the older groups, may be due to lack of development in regions of the brain that may help with top down processing [37, 57]. From previous work, it is known that young children are more susceptible to informational masking [7, 12, 25, 26, 58]. This may be due to the lack of development of auditory attention that is necessary to segregate sources [31]. Older children and adults have more optimal listening strategies where they are better able to use spatial cues, and in addition have a more developed ability to selectively attend to the target [7, 11, 12]. Physiological development may also affect differences in performance between children and adults on speech-in-noise tasks, such that adults benefit more from spatial cues than children. Myelination connecting the right and left hemispheres of the corpus callosum, which are important for perception of spoken language, continues to develop into adolescence [28]. There is also evidence suggesting that neural connections become more efficient after birth [29]. All of the above may contribute to the increased susceptibility of children to auditory masking. No between-group effects are shown with either interferer at the more favorable SNRs, suggesting that there are no effects of age on ability to make use of spatial cues when the target is played at a positive SNR.

C. Relationship of Executive Function Skills and Performance on Speech Measures

Measures of attention and working memory were included to investigate the relationship between executive function and ability to take advantage of spatial separation of sources (i.e., SRM).

Table V displays the mean (±SD) and range for the measures of executive function for all 4 groups. The DCCS and Flanker tests show the unadjusted score (score not taking into account any demographic variables). In addition, the Digit Span scores shown reflect the longest span of digits accurately repeated. Table VI displays the mean (±SD) and range of the raw scores for the Kbit and EVT. Measures of IQ, expressive vocabulary, and executive function were moderately correlated when all groups were combined (N=92) (values shown in Table VII).

Regression analyses were used to test if executive function, expressive vocabulary, or IQ significantly predicted SRM. For the analyses described, SRM was collapsed across interferer type (noise and speech) and semantic content (coherent and anomalous). The first analysis for each SNR was a simple linear regression and included only the measure of executive function: DCCS, Flanker, Longest Digit Span forward, or Longest Digit Span Backward. The second analysis for each SNR was a three stage hierarchical multiple regression model that included group (age), expressive vocabulary (EVT) and IQ (Kbit), and the measures of executive function (longest digit span forward, longest digit span backward, Flanker, and DCCS) as predictors. For the second analysis, group was entered as stage one of the regression to control for age, expressive vocabulary and IQ were entered at stage two, and all measures of executive function were entered at stage three. The third analysis for each SNR was also a three stage hierarchical multiple regression model that included group (age), expressive vocabulary (EVT) and IQ (Kbit), and the measures of executive function (longest digit span forward, longest digit span backward, Flanker, and DCCS) as predictors. For the third analysis, group was entered in stage one of the regression to control for age, and in contrast to the second analysis, all measures of executive function were entered in stage two, and expressive vocabulary and IQ were entered in stage three. The predictor variables for the second and third analysis were entered in reverse order for stage two and three of the regression models to investigate if there was an effect of executive function over and above language or vice versa.

The first simple linear regression analysis (with only the measure of executive function as a predictor), followed by the hierarchical multiple regression analyses (with added predictors, in various orders) will be described below.

Simple linear regression models, with each measures of executive function as the only predictor of SRM, are described for each SNR. Figures 6 and 7 show simple linear regressions for each measure of executive function (Fig.6: DCCS and Flanker, Fig. 7: Longest Digit Span Forward and Longest Digit Span Backward), with all groups combined, plotted against SRM for the -16 dB SNR (Panels A) and -8 dB SNR (Panels B), where the most SRM was observed. The simple linear regression equation with DCCS (r^2=0.11, *F (1,89)*=10.92*,* **p*=0.001) significantly predicted SRM for the -16 dB SNR (Fig. 6A). A non-significant correlation between DCCS and SRM was observed for the -8 dB SNR (r^2=0.00, *F (1,89)*=0.023*, p*=0.88) (Fig. 6B), 0 dB SNR (r^2=0.03, *F (1,89)*=2.5*, p*=0.12), and +8 dB SNR (r^2=0.04, *F (1,89)*=4.0*, p*=0.051). These results show that DCCS was only a significant predictor of SRM at the least favorable SNR when no other variables were included in the model. Similar results were found for the Flanker. The regression equation with Flanker (r^2=0.12, *F (1,89)*=12.45*,* **p*=0.001) significantly predicted SRM for the -16 dB SNR (Fig. 6A). A non-significant correlation between Flanker and SRM was observed for the -8 dB SNR (r^2=0.005, *F (1,89)*=0.41*, p*=0.52) (Fig. 6B), 0 dB SNR (r^2=0.004, *F (1,89)*=0.32*, p*=0.057), and +8 dB SNR (r^2=0.014, *F (1,89)*=1.2*, p*=0.27). These results show that Flanker was only a significant predictor of SRM at the least favorable SNR when no other variables were included in the model. The simple linear regression equation with Digit Span Forward (r^2=0.076, *F (1,89)*=7.2*,* **p*=0.009) significantly predicted SRM for the -16 dB SNR (Fig. 7A). A non-significant correlation between Digit Span Forward and SRM was observed for the -8 dB SNR ($r^2=0.00$, *F (1,89)*=0.04*, p*=0.84) (Fig. 7B), 0 dB SNR ($r^2=0.004$, *F (1,89)*=0.38*, p*=0.54), and +8 dB SNR ($r^2=0.006$, *F (1,89)*=0.49*, p*=0.48). These results show that Digit Span Forward was only a significant predictor of SRM at the least favorable SNR when no other variables were included in the model. The analyses with the Digit Span Backward show similar results to the Digit Span Forward. The regression equation with Digit Span Backward ($r^2=0.09$, *F (1,89)*=8.7*, *p*=0.004) significantly predicted SRM for the -16 dB SNR. A non-significant correlation between Digit Span Backward and SRM was observed for the -8 dB SNR ($r^2=0.002$, *F (1,89)*=0.19*, p*=0.66), 0 SNR ($r^2=0.001$, *F (1,89)*=0.05*, p*=0.82), and +8 dB SNR ($r^2=0.001$, *F (1,89)*=0.09*, p*=0.76). These results show that Digit Span Backward was only a significant predictor of SRM at the least favorable SNR when no other variables were included in the model.

A three stage hierarchical multiple regression was conducted for the second and third analyses, at each SNR, with SRM as the dependent variable. The independent variables that were included were group (age), expressive vocabulary and IQ, and measures of executive function. For both the second and third analyses group was entered in stage one to control for age. For the second analysis, expressive vocabulary and IQ were entered in stage two, and measures of executive function were entered in stage three. Stage two and three were reversed in the third analyses (stage two: measures of executive function, stage three: expressive vocabulary and IQ). The second analysis as described in the following: the hierarchical multiple regression for SRM at -16 dB SNR revealed that at stage one, age significantly contributed to the regression model, *F* (1,87)=21.59, **p*<0.001 and accounted for 19.9% of the variation in SRM. Adding expressive vocabulary and IQ to the regression model explained an additional 7.3% of the variation in SRM at -16 dB SNR and this change in R^2 was significant, F (1,87)=10.56, *p<0.05. Finally, the addition of the measures of executive function in stage three of the regression model did not significantly explain any additional variation in SRM, F (1,87)=4.5, p=0.91.

For the third analysis, when measures of executive function were entered in stage two and expressive vocabulary and IQ were entered in stage 3, expressive vocabulary and IQ had an effect on SRM above the measures of executive function. The third analysis revealed that in stage one, group (age), significantly contributed to the regression model, as shown above using the same predictors as the second analysis (age accounted for 19.9% of variation in SRM). In stage two, when adding measures of executive function, the regression model did not explain any additional variance, F (1,87)=4.6, p=0.75. Finally, in stage three, the addition of expressive vocabulary and IQ significantly contributed to the regression model, F (1,87)=4.51, *p<0.05 and accounted for an additional 6.3% of the variation in SRM. Together, all variables (age, executive function, expressive vocabulary, and IQ) accounted for 28.1% of the variance in SRM at -16 dB SNR.

None of the hierarchical regression models were significant at any of the stages for the -8, 0, or +8 dB SNRs. Results for the second analysis (stage 2: expressive vocabulary and IQ, stage 3: measures of executive function) are as follows: -8 dB SNR: stage one= r^2=0.002, F (1,87)=0.15, p=0.70, stage two= r^2=0.004, F (1,87)=0.12, p=0.95, all variables= r^2=0.016, F (1,87)=0.19, p=0.99; 0 dB SNR: stage one= r^2=0.008, F (1,87)=0.68, p=0.41, stage two= r^2=0.03, F (1,87)=0.89, p=0.45, all variables= r^2=0.07, F (1,87)=0.87, p=0.53; +8 dB SNR, stage one= r^2=0.00, F (1,87)=0.001, p=0.97, stage two= r^2=0.02, F (1,87)=0.57, p=0.64, all variables= r^2=0.107, F (1,87)=1.38, p=0.23. Results for the third analysis (stage 2: measures of executive function, stage 3: expressive vocabulary and IQ) are as follows: -8 dB SNR: stage one= $r^2=.002$, $F(1,87)=0.15$, $p=0.70$, stage two= $r^2=0.012$, $F(1,87)=0.19$, $p=0.96$, all variables= $r^2=0.016$, $F(1,87)=0.19$, $p=0.99$; 0 dB SNR: stage one= $r^2=0.008$, $F(1,87)=0.68$, $p=0.41$, stage two= $r^2=0.053$, $F(1,87)=0.93$, $p=0.47$, all variables= $r^2=0.07$, $F(1,87)=0.87$, $p=0.53$; +8 dB SNR: stage one= $r^2=0.000$, $F(1,87)=0.001$, $p=0.97$, stage two= $r^2=0.094$, $F(1,87)=1.7$, $p=0.14$, all variables= $r^2=0.11$, $F(1,87)=1.38$, $p=0.23$.

Lastly, it was of interest to examine measures of executive function in relation to SRM for the coherent sentences and the anomalous sentences separately. SRM for only the -16 dB SNR was analyzed, as this was the only SNR that proved significant in the analyses described above. Figure 8 displays SRM (coherent and anomalous separated) as a function of performance on each measure of executive function (A: DCCS, B: Flanker, C: Longest Digit Span Forward, D: Longest Digit Span Backward). The circles (dashed line) represent the simple linear regression for SRM with coherent sentences, and the triangles (solid line) represent simple linear regression for SRM with anomalous sentences. Although the simple linear regression equation comparing the relationship between each measure of executive function with SRM for the coherent sentences showed a trend for a stronger relationship with the measures of executive function than the SRM for the anomalous sentences, statistically significant differences between the coherent and anomalous correlation values was not reached in any of the cases. Results are as follows: DCCS coherent SRM ($r^2=0.34$, $F(1,88)=6.4$, $^*p=0.001$) and anomalous SRM ($r^2=0.24$, $F(1,88)=5.2$, $^*p=0.024$), coherent vs. anomalous $p=0.47$; Flanker coherent SRM ($r^2=0.39$, $F(1,88)=15.5$, $^*p<0.001$) and anomalous SRM ($r^2=0.22$, $F(1,88)=4.6$, $^*p=0.035$), coherent vs. anomalous $p=0.21$; Longest Digit Span Forward coherent SRM ($r^2=0.28$, $F(1,88)=7.6$, $^*p=0.007$) and anomalous SRM ($r^2=0.22$, $F(1,88)=3.5$, $p=0.06$), coherent vs. anomalous $p=0.67$; Longest

Digit Span Backward coherent SRM (r^2=0.27, *F* (1,88)=7.1, **p*=0.009) and anomalous SRM (r^2=0.26, *F* (1,88)=6.4, **p*=0.013), coherent vs. anomalous p=0.94.

D. Discussion: Relationship of Executive Function Skills and Performance on Speech Measures

In general, performance on all measures of executive function showed a trend of improvement with age. This aligns with previous research suggesting that the areas of the brain that are thought to control executive function continue to develop throughout adolescence [37, 57].

This study investigated why some listeners function exceptionally well in realistic adverse environments and why others do not [2]. The current study aimed to examine specific components of executive function that were predicted to impact SRM. This was of interest because our previous work has shown variability in benefit of spatial cues within groups of NH listeners (as described in Chapters 2 and 3 of this dissertation). Open-set sentences, varying in semantic content, were used to create a listening task in which top-down processes interact with peripheral “bottom-up” sensory processing [56]. Results showed that specific measures of executive function (i.e. DCCS, Flanker, Digit Span Forward, and Digit Span backward) were only predictors of SRM when no other predictors were accounted for, and at the least favorable SNR. After adding predictor variables that account for age, expressive vocabulary, and IQ, the significant effects of executive function disappeared.

Working memory is a component of executive function that has been found to be relevant to the identification of target speech in noisy environments. When processing complex speech in adverse listening environments, both short-term and working memory are necessary [59]. A listener must actively store incoming information (short-term memory), while at the same time integrate and manipulate information that is continually being received (working memory). The quantity of information that can be held in working memory is limited, and is affected by clarity of the incoming signal [40, 60]. As a result, working memory can be negatively impacted when listening in adverse environments relative to when listening in more favorable environments (e.g. Quiet or favorable SNRs).

It has been shown that better working memory is associated with better performance on word recognition tasks for children who are hearing impaired and use electronic hearing devices such as cochlear implants [61, 62]. Similarly, growth and training of working memory has been shown to have a relationship with improvement in language skills [42-44]. This work suggests that those who have a larger capacity for information in working memory perform better on tasks that involve working memory. Therefore, it was hypothesized that the listeners who had larger spans of both short-term and working memory would be better at segregating sources at poor SNRs because the complex auditory task used in this study was designed to incorporate activation of these processes. With all age groups collapsed together and measures of memory accounted for individually, both short-term and working memory were significant predictors of SRM at only the -16 dB SNR and when no other predictors were included. The measures of memory were significant predictors only at the least favorable SNRs tested in this experiment, which suggests that short-term and working memory capacity is most sensitive in extremely difficult listening situations. Our results coincide with previous findings for older adult listeners in which memory is most affected in the most complex listening situations [63]. The significant relationships between SRM and measures of memory disappeared after age, expressive vocabulary, and IQ were added as predictors. In the latter model, age, expressive vocabulary, and

IQ predicted significant variability in SRM above and beyond executive function. This finding is not surprising given that other studies have shown that tests of working and short-term memory are highly correlated with language outcomes [43, 64].

Both attention shifting and inhibition are imperative when listening to speech in noise. The listener must be able to shift and direct attention to the target while simultaneously inhibiting other interferers. It is thought that attentional processes are key to successful management and organization of working memory [65]. Therefore, it is suggested that attention and working memory are interconnected. For this study, analyses of these factors were conducted separately. It was hypothesized that listeners who were better able to shift attention and inhibit irrelevant stimuli would be better at segregating sources at the least favorable SNRs because they are better able to focus on the target while inhibiting the interferers. With all age groups collapsed together, and measures of attention accounted for individually, both shifting and inhibition were significant predictors of SRM at -16 dB SNR only when no other predictors were included in the model. The measures of attention were significant predictors only at the least favorable SNRs tested in this experiment, which, as discussed above in the context of memory, may suggest that attentional skills are most engaged in the least favorable listening environments [63]. Although, results from this study suggest that age, IQ and expressive vocabulary impact performance on spatial hearing tasks more so than attention.

The relationship of measures of attention and memory did not significantly differ with SRM when participants listened to coherent verses anomalous sentences, although there was a trend for SRM with the coherent sentences to have a stronger relationship with the measures of attention and memory. When bottom-up sensory processing is degraded (target in the presence of interferers), top-down processing may help to alleviate some difficulty in extracting the target speech [66]. It was of interest to investigate if supportive semantic content cues would enhance performance when the sensory signal was degraded. The amount of top-down and bottom-up processing is likely different for the coherent and anomalous sentences. The coherent sentences give the listener a chance to predict words within the sentence using cues of semantic content and knowledge. This places more emphasis on top-down vs. bottom-up processing. Unlike with coherent sentences, top-down processing is not as useful for the anomalous sentences, where semantic content cues are unavailable. Other studies have demonstrated that NH adult listeners improve on auditory tasks that incorporate additional cues in adverse conditions because they are able to take advantage of top-down processing [67]. In the current study, the additional cue of semantic content (coherent sentences) is helpful when examining release from masking. Results of the current study show that ability to take advantage of spatial separation, regardless of semantic cues, was better for the participants who performed higher on the measures of executive function. Because the relationship between SRM and executive functioning skills is stronger with the coherent sentences, results suggest that listeners with better executive functioning skills are better able to employ top-down processing and make use of the semantic content cues when available to benefit more from SRM.

VI. FUTURE DIRECTIONS

These findings have implications for both normal developing and clinical populations. In the classroom, these results may help in the organization and design of educational settings. Favorable SNRs are needed so that children are able to effectively listen and learn in noisy classroom environments. There is evidence to suggest that younger children require a more favorable SNR than older children and adults.

These results may also have clinical implications. Predictors of performance, such as measures of executive function, on tasks of auditory source segregation in the future may assist in clinical decisions as to what amplification devices to fit. They may also help to predict success and intervention needs for children with hearing impairments. In addition, these measures of executive function could be used in a training paradigm to improve ability to listen to speech in noise for NH and clinical populations.

The proposed measures of executive function were not sensitive enough to capture the variability in performance on the auditory task in more favorable SNRs, above -16 dB. In future studies, it may also be beneficial to include a processing speed measure, dual-task methods (e.g. that demand greater cognitive load than an auditory task alone), and measure exact auditory thresholds, all of which have been shown to partially predict performance on auditory measures for older adults. It may also be beneficial to measure listening effort, with pupilometry or eye-tracking movement measures, in tasks of auditory source segregation in order to isolate conditions in which listeners exert the most effortful listening. It may be these conditions that reveal the most significant relationships with measures of executive function. Likewise, it would also be beneficial to compare our results to imaging studies to better elucidate which specific mechanisms are involved when segregating sources in a complex auditory environment in order to develop an improved measure that could be utilized in the clinical setting.

TABLES

Table 1

Table I. Results of planned comparisons: Front vs.
90 Asymmetrical % correct (*significance)

Group (years)	Interferer	-16 SNR	-8 SNR	0 SNR	+8 SNR
7 - 10	noise	*p<0.001	*p<0.001	*p=0.002	p=0.12
11 - 14	noise	*p<0.001	*p<0.001	p=0.02	p=0.12
15 - 17	noise	*p<0.001	*p=0.007	p=0.02	p=0.11
18 - 23	noise	*p<0.001	*p<0.001	*p=0.003	p=0.06
7 - 10	speech	*p<0.001	*p<0.001	p=0.03	p=0.20
11 - 14	speech	*p<0.001	*p<0.001	*p<0.001	p=0.08
15 - 17	speech	*p<0.001	*p<0.001	*p=0.003	p=0.09
18 - 23	speech	*p<0.001	*p<0.001	p=0.43	p=0.014

Table 2

Table IIa. Results of within group effects of SNR on SRM_noise interferer(*significance)

Group (years)	SNR comparisons					
	-16 vs. -8	-16 vs. 0	-16 vs. +8	-8 vs. 0	-8 vs. +8	0 vs. +8
7 - 10	p=0.03	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=0.19
11 - 14	p=0.46	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=0.16
15 - 17	p=1.0	*p=0.001	*p=0.001	*p=0.001	*p=0.001	p=1.0
18 - 23	p=0.03	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=0.19

Table IIb. Results of within group effects of SNR on SRM_speech interferer(*significance)

Group (years)	SNR comparisons					
	-16 vs. -8	-16 vs. 0	-16 vs. +8	-8 vs. 0	-8 vs. +8	0 vs. +8
7 - 10	p=0.58	*p=0.003	*p<0.001	*p<0.001	*p<0.001	p=0.17
11 - 14	p=0.10	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=0.03
15 - 17	*p=0.006	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=0.45
18 - 23	*p<0.001	*p<0.001	*p<0.001	*p<0.001	*p<0.001	p=1.0

Table 3

Table IIIa. Results of planned comparisons: Coherent vs. Anomalous % correct_Front (*significance)

Group (years)	Interferer	-16 SNR	-8 SNR	0 SNR	+8 SNR	Quiet
7 - 10	noise	t(11)=2.2, p=0.344	t(11)=7.2, *p<0.001	t(11)=8.1, *p=0.001	t(11)=3.9, p=0.03	t(11)=3.1, p=0.08
11 - 14	noise	t(11)=1.7, p=0.15	t(11)=6.9, *p<0.001	t(11)=5.9, *p<0.001	t(11)=8.1, *p<0.001	t(11)=2.5, p=0.02
15 - 17	noise	t(8)=2.02, p=0.40	t(8)=4.8, p=0.02	t(8)=4.5, *p<0.001	t(8)=0.7, *p<0.001	t(8)=0.7, p=0.83
18 - 23	noise	t(11)=0.7, p=0.38	t(11)=4.4, *p<0.001	t(11)=3.1, *p=0.003	t(11)=1.9, p=0.02	t(11)=1.5, p=0.04
7 - 10	speech	t(11)=2.4, p=0.03	t(11)=4.8, p=0.013	t(11)=5.9, *p<0.001	t(11)=5.5, p=0.03	t(11)=2.5, p=0.03
11 - 14	speech	t(11)=3.1, p=0.12	t(11)=9.6, *p<0.001	t(11)=8.1, *p<0.001	t(11)=3.7, *p=0.002	t(11)=3.0, p=0.02
15 - 17	speech	t(10)=4.1, p=0.03	t(10)=10.03, *p<0.001	t(10)=4.9, *p<0.001	t(10)=2.5, *p<0.001	t(10)=4.2, p=0.04
18 - 23	speech	t(11)=4.4, p=0.013	t(11)=8.1, *p<0.001	t(11)=7.1, *p<0.001	t(11)=7.4, *p<0.001	t(11)=3.1, p=0.08

Table IIIb. Results of planned comparisons: Coherent vs. Anomalous % correct_90A (*significance)

Group (years)	Interferer	-16 SNR	-8 SNR	0 SNR	+8 SNR
7 - 10	noise	t(11)=5.3, *p=0.003	t(11)=5.9, *p<0.001	t(11)=4.3, p=0.03	t(11)=3.4, p=0.03
11 - 14	noise	t(11)=8.8, *p<0.001	t(11)=8.8, *p<0.001	t(11)=3.7, p=0.02	t(11)=4.3, *p=0.008
15 - 17	noise	t(8)=12.6, *p<0.001	t(8)=4.2, *p=0.001	t(8)=6.9, *p=0.001	t(8)=1.9, p=0.20
18 - 23	noise	t(11)=8.4, *p<0.001	t(11)=8.6, *p<0.001	t(11)=4.8, *p=0.005	t(11)=4.6, *p=0.004
7 - 10	speech	t(11)=4.1, p=0.014	t(11)=7.8, *p<0.001	t(11)=11.2, *p=0.004	t(11)=3.2, p=0.011
11 - 14	speech	t(11)=8.0, *p<0.001	t(11)=6.5, *p<0.001	t(11)=2.9, *p=0.009	t(11)=4.1, *p<0.001
15 - 17	speech	t(10)=10.9, *p<0.001	t(10)=4.8, *p<0.001	t(10)=2.4, p=0.02	t(10)=4.2, *p<0.001
18 - 23	speech	t(11)=12.2, *p<0.001	t(11)=9.2, *p<0.001	t(11)=5.7, *p<0.001	t(11)=4.8, *p=0.005

Table 4

Table IV. Results of planned comparisons:coherent vs. anomalous SRM (*significance)

Group (years)	Interferer	-16 SNR	-8 SNR	0 SNR	+8 SNR
7 - 10	noise	t(11)=4.6, *p=0.002	t(11)=0.9, p=0.30	t(11)=-3.4, p=0.01	t(11)=-0.7, p=0.45
11 - 14	noise	t(11)=6.5, *p<0.001	t(11)=0.2, p=0.87	t(11)=-2.4, p=0.03	t(11)=-2.6, *p=0.007
15 - 17	noise	t(6)=11.1 ,*p<0.001	t(6)=-0.2, p=0.82	t(6)=-2.0, p=0.043	t(6)=-2.7, p=0.02
18 - 23	noise	t(11)=8.3, *p<0.001	t(11)=1.0, p=0.39	t(11)=-1.4, p=0.18	t(11)=-1.2, p=0.23
7 - 10	speech	t(11)=2.8, p=0.6	t(11)=7.8, *p<0.001	t(11)=-3.04, p=0.03	t(11)=0.9, p=0.35
11 - 14	speech	t(11)=4.1, *p=0.007	t(11)=6.5, *p=0.002	t(11)=-6.6, *p<0.001	t(11)=-0.5, p=0.61
15 - 17	speech	t(10)=4.4, *p=0.003	t(10)=-0.5, p=0.71	t(10)=-0.8, p=0.31	t(10)=-0.9, p=0.34
18 - 23	speech	t(11)=6.2, *p<0.001	t(11)=4.0, p=0.02	t(11)=-0.7, p=0.52	t(11)=-4.1, *p<0.001

Table 5

Table V. Mean scores (±SD) and range for measures of executive function

Group (years)	DCCS (unadjusted score)	Flanker (unadjusted score)	Digit Span Forward (longest)	Digit Span Backward (longest)
7 - 10	100.11 (8.14) 85.01 - 117.13	95.15 (6.81) 79.47 - 107.34	5.79 (1.06) 4 - 8	4.08 (1.06) 3 - 6
11-14	108.60 (7.06) 94.54 - 124.39	106.01 (9.46) 91.59 - 142.11	6.00 (1.24) 4 - 8	4.91 (1.31) 3 - 7
15-17	118.64 (10.99) 106.05 - 143.94	122.26 (10.77) 103.40 - 142.11	7.05 (0.10) 5 - 8	4.9 (1.25) 3 - 7
18-23	124.42 (11.82) 109.24 - 143.94	124.98 (11.42) 101.21 - 142.11	7.08 (0.89) 5 - 8	5.46 (1.06) 3-7

Table 6

Table VI. Mean scores (±SD) and range for IQ and expressive vocabulary

Group (years)	Kbit (raw scores)	EVT (raw scores)
7 - 10	33.33 (5.84) 16 - 41	131.29 (21.94) 75 - 183
11-14	36.70 (4.45) 28 - 43	144.70 (19.27) 92 - 171
15-17	37.70 (4.81) 27 - 45	156.50 (17.27) 131 - 196
18-23	39.75 (4.53) 28 - 45	168.33 (9.24) 148 - 187

Table 7

Table VII. Correlations for IQ, expressive vocabulary, and all measures of executive function

	Kbit (raw scores)	EVT (raw scores)	DCCS (unadjusted score)	Flanker (unadjusted score)	Digit Span Forward (longest)	Digit Span Backward (longest)
Kbit (raw scores)		r (90)=0.54, *p< 0.01	r (90)=0.30, *p< 0.01	r (90)=0.54, *p< 0.01	r (90)=0.33, *p< 0.01	r (90)=0.30, *p< 0.01
EVT (raw scores)	r (90)=0.54, *p< 0.01		r (90)=0.49, *p< 0.01	r (90)=0.57, *p< 0.01	r (90)=0.49, *p< 0.01	r (90)=0.47, *p< 0.01
DCCS (unadjusted score)	r (90)=0.42, *p< 0.01	r (90)=0.49, *p< 0.01		r (90)=0.80, *p< 0.01	r (90)=0.50, *p< 0.01	r (90)=0.42, *p< 0.01
Flanker (unadjusted score)	r (90)=0.54, *p< 0.01	r (90)=0.57, *p< 0.01	r (90)=0.80, *p< 0.01		r (90)=0.41, *p< 0.01	r (90)=0.34, *p< 0.01
Digit Span Forward (longest)	r (90)=0.33, *p< 0.01	r (90)=0.49, *p< 0.01	r (90)=0.50, *p< 0.01	r (90)=0.41, *p< 0.01		r (90)=0.55, *p< 0.01
Digit Span Backward (longest)	r (90)=0.30, *p< 0.01	r (90)=0.47, *p< 0.01	r (90)=0.42, *p< 0.01	r (90)=0.34, *p< 0.01	r (90)=0.55, *p< 0.01	

FIGURES

Figure 1

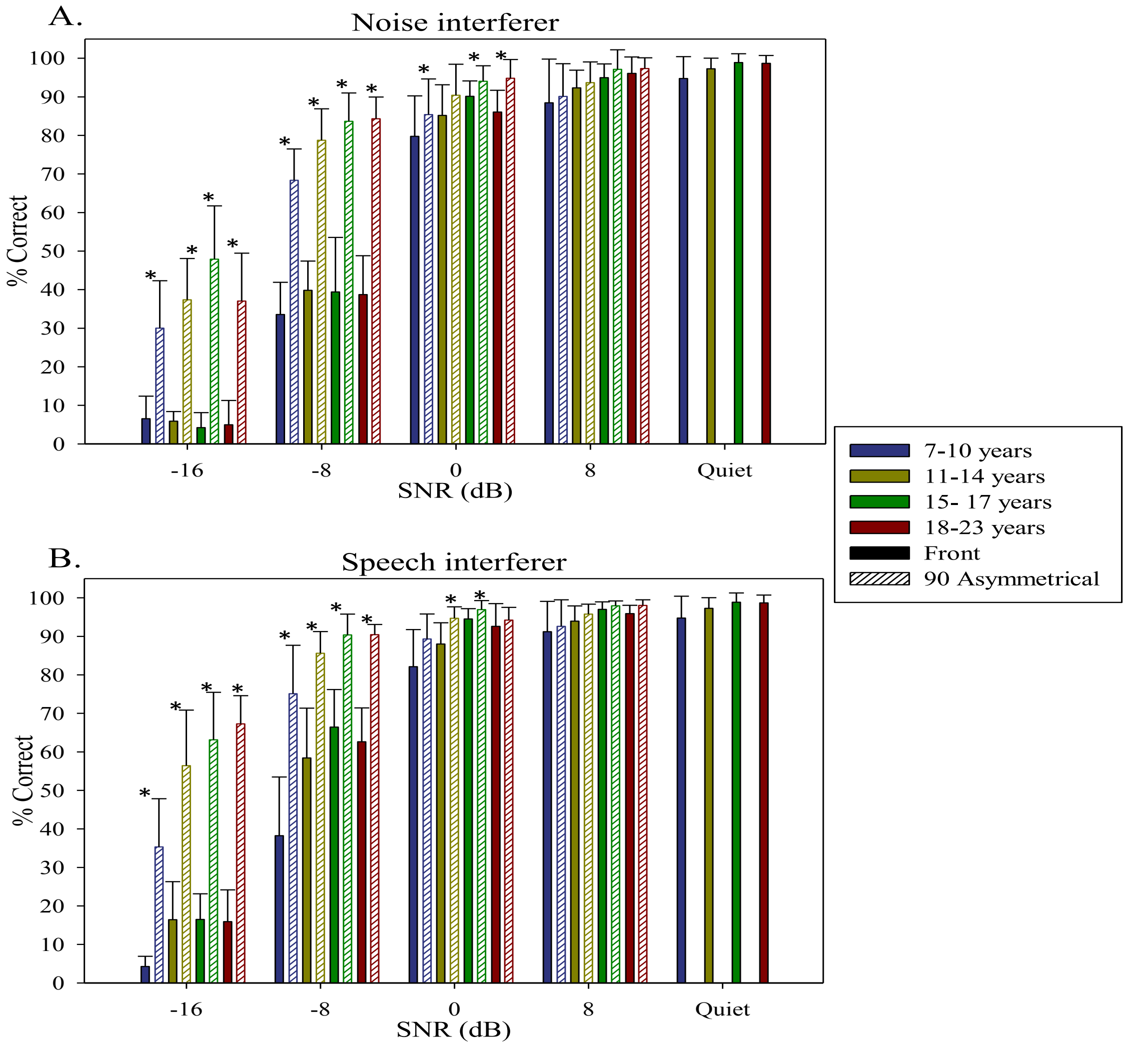

FIG. 1.

Mean (±SD) percent correct for both the noise interferer (1A) and the speech interferer (1B) are plotted, at each SNR, for each group (i.e., 7-10, 11-14, 15-17, 18-23). Solid bars represent percent correct in the Front condition. Dashed bars represent percent correct in the 90 Asymmetrical condition. Significant differences within each group between the Front vs. 90 Asymmetrical conditions (*p* < 0.0125) are highlighted (*).

Figure 2

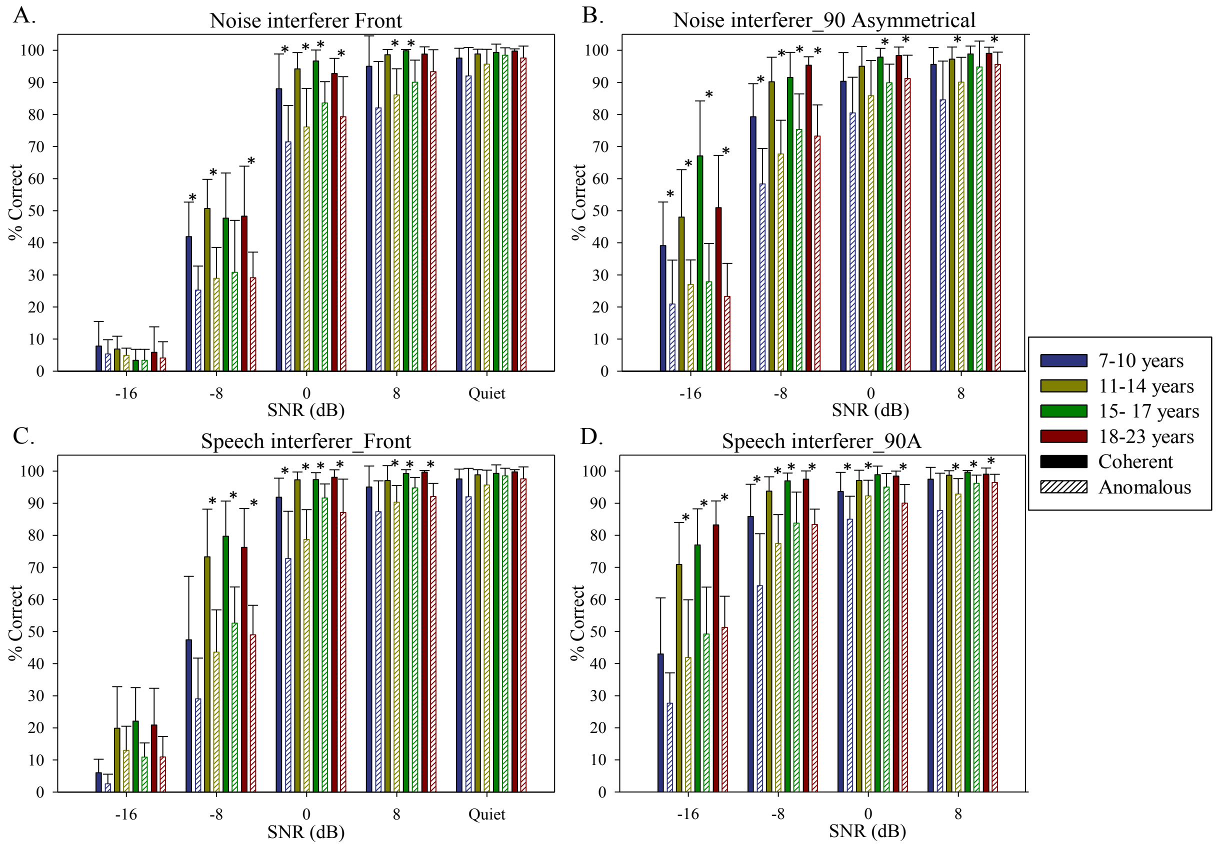

FIG. 2. Mean (±SD) percent correct for both the noise interferer (2A and 2B) and the speech interferer (2A and 2B) are plotted, at each SNR, for each group (i.e., 7-10, 11-14, 15-17, 18-23). Solid bars represent percent correct with the coherent sentences. Dashed bars represent percent correct with the anomalous sentences. Significant differences within each group between the coherent vs. anomalous sentences (*p* < 0.01) are highlighted (*).

Figure 3

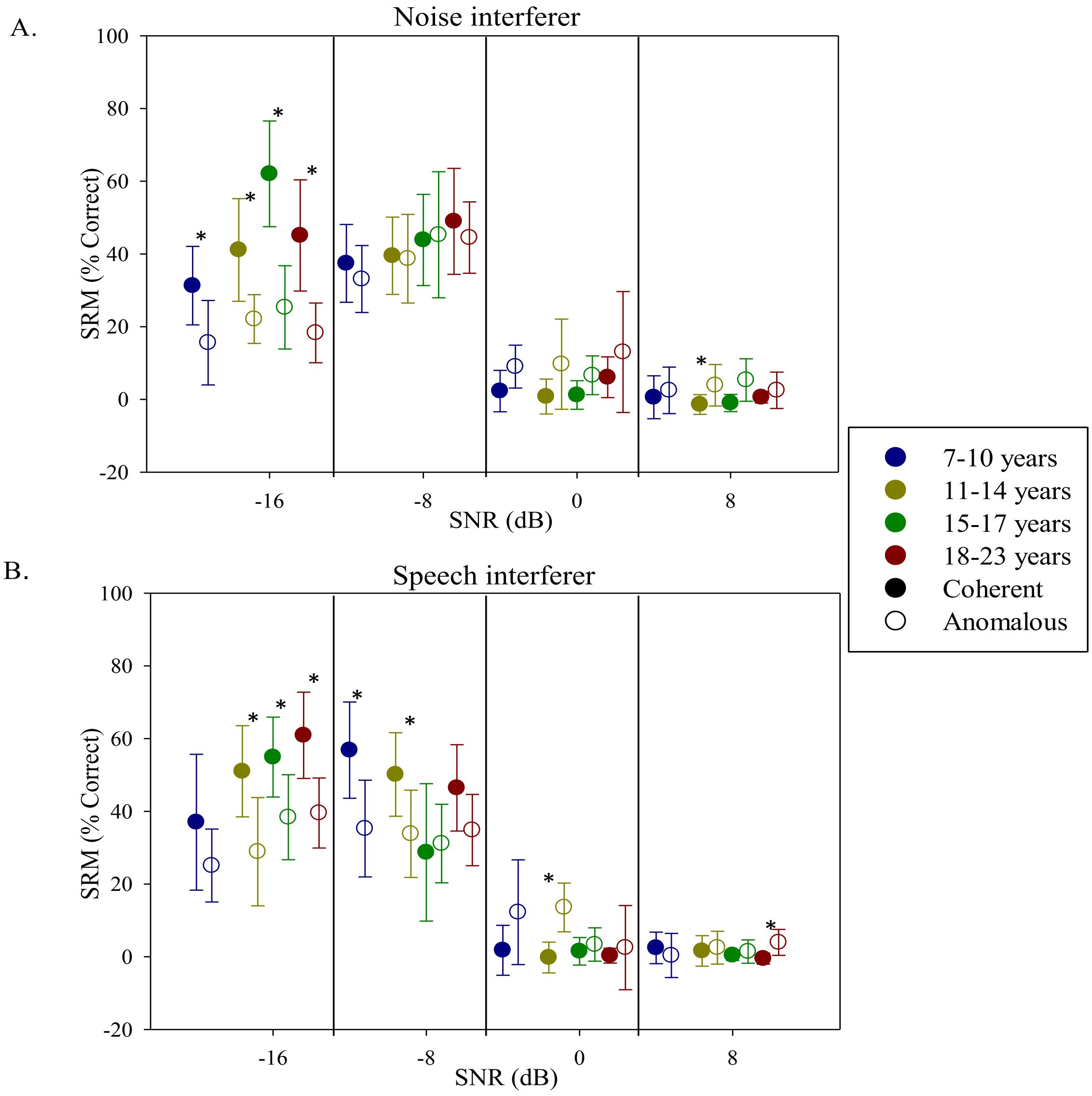

FIG. 3. Mean (±SD) SRM (change in percent correct between 90 Asymmetrical and Front conditions) are plotted, at each SNR, for all groups both the coherent and anomalous sentences. The filled circles represent SRM with the coherent sentences. The open circles represent SRM with the anomalous sentences. Significant differences within each group between the coherent SRM vs. anomalous SRM ($p < 0.0125$) are highlighted (*).

Figure 4

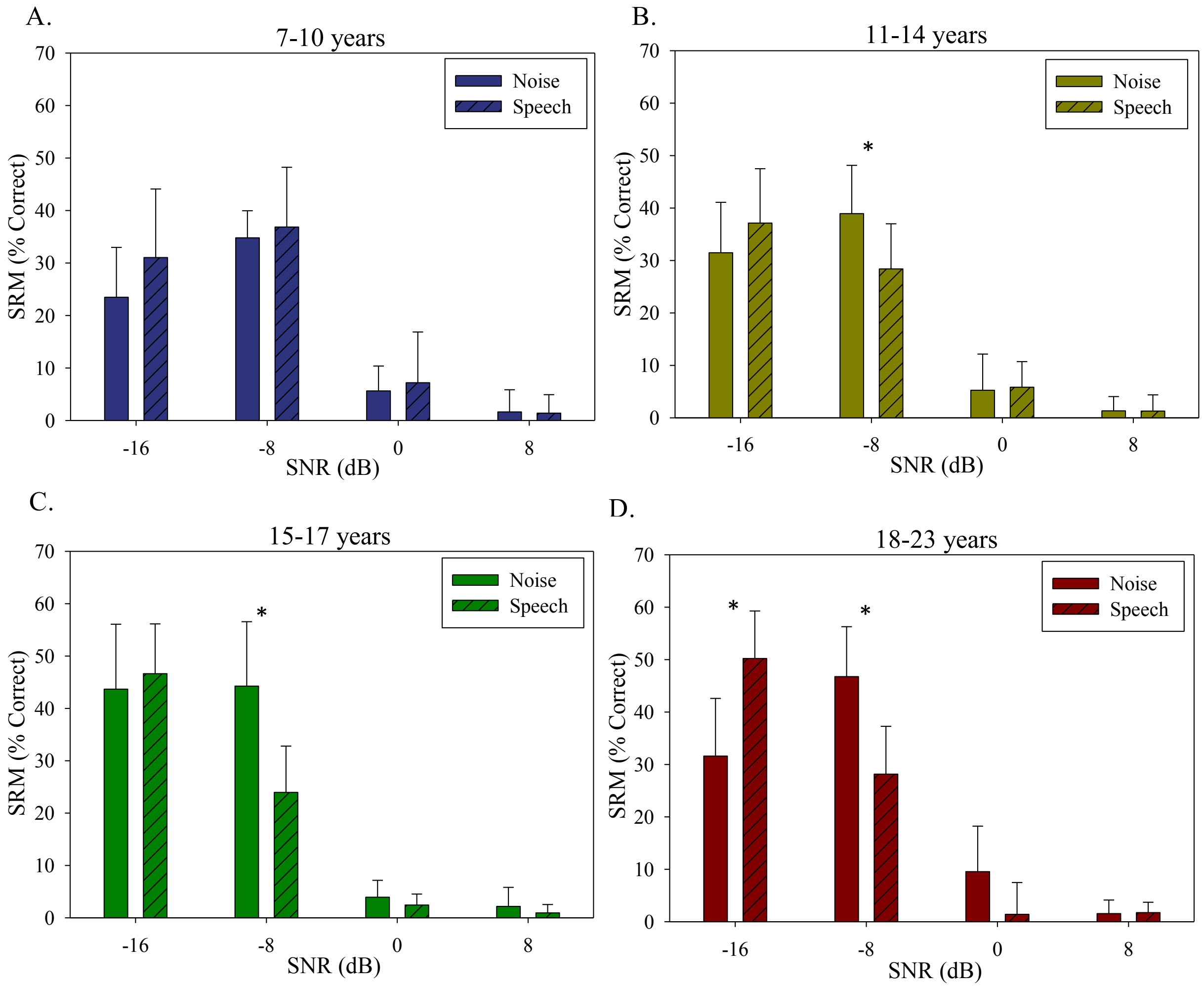

FIG. 4. Mean (±SD) SRM (change in percent correct between 90 Asymmetrical and Front conditions) are plotted, at each SNR, for each group (A: 7-10, B: 11-14, C: 15-17, D: 18-23). Solid bars represent SRM with noise interferers. Dashed bars represent SRM with speech interferers. Significant differences ($p < 0.0125$) are highlighted (*).

Figure 5

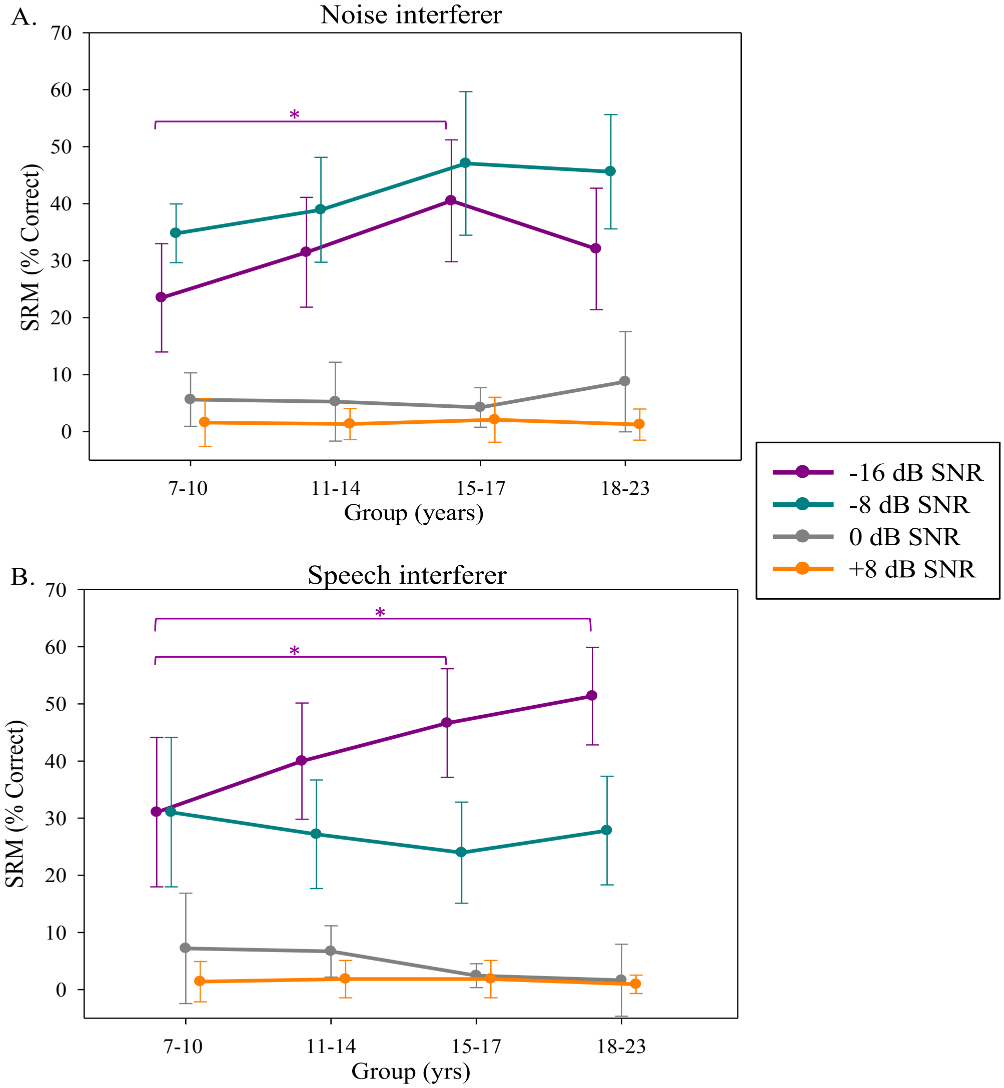

FIG. 5. Mean (±SD) SRM (change in percent correct between 90 Asymmetrical and Front conditions) are plotted, for each group, with both the noise (5A) and speech (5B) interferers. Each line represents a different SNR (dB). Significant differences between group for each SNR ($p < 0.0125$) are highlighted (*).

Figure 6

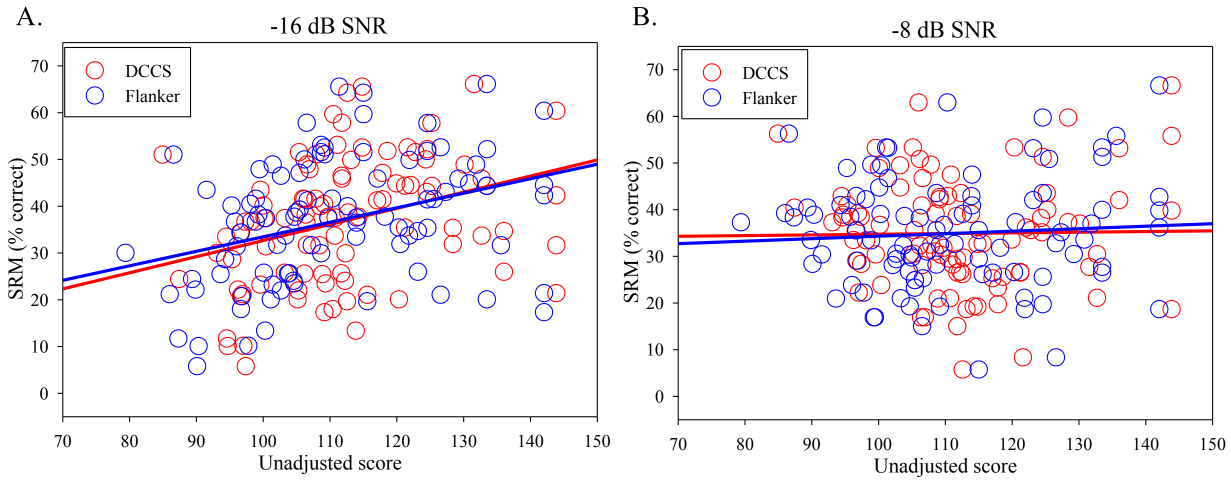

FIG. 6. Simple linear regression showing the relationship between SRM (change in percent correct between 90A and front conditions) and scores on DCCS and Flanker for all groups combined (n=92 participants). For panel A: *DCCS:* $r^2=0.27, p<0.001$; *Flanker:* $r^2=0.12, p=0.001$.

Figure 7

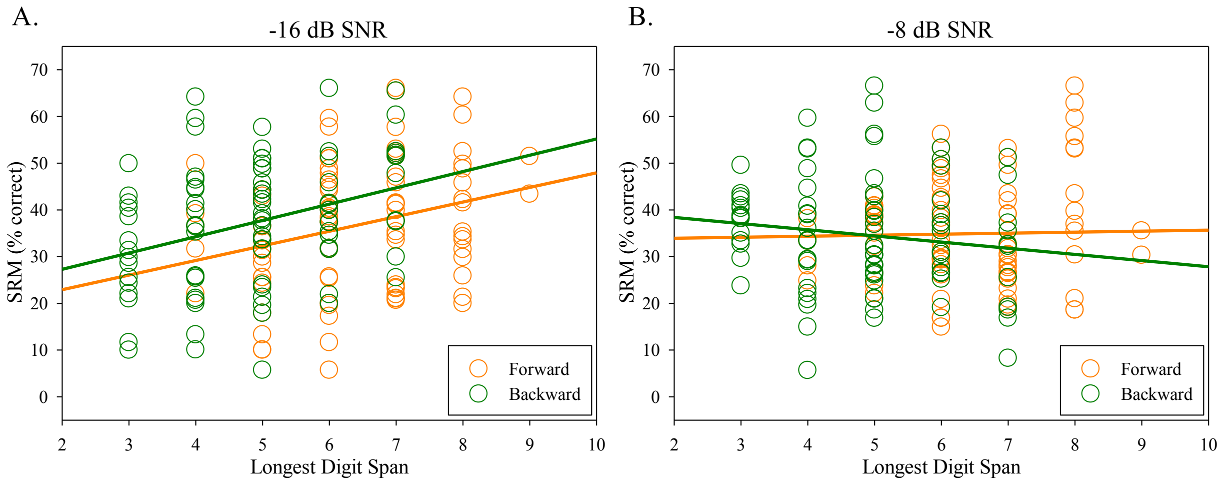

FIG. 7. Simple linear regression showing the relationship between SRM (change in percent correct between 90A and front conditions) and scores on the Digit Span Forward and Backward for all groups combined (n=92 participants). For panel A: *Digit Span Forward:* $r^2=0.076$, *p=0.009; Digit Span Backward:* $r^2=0.09$, *p=0.004*.

Figure 8

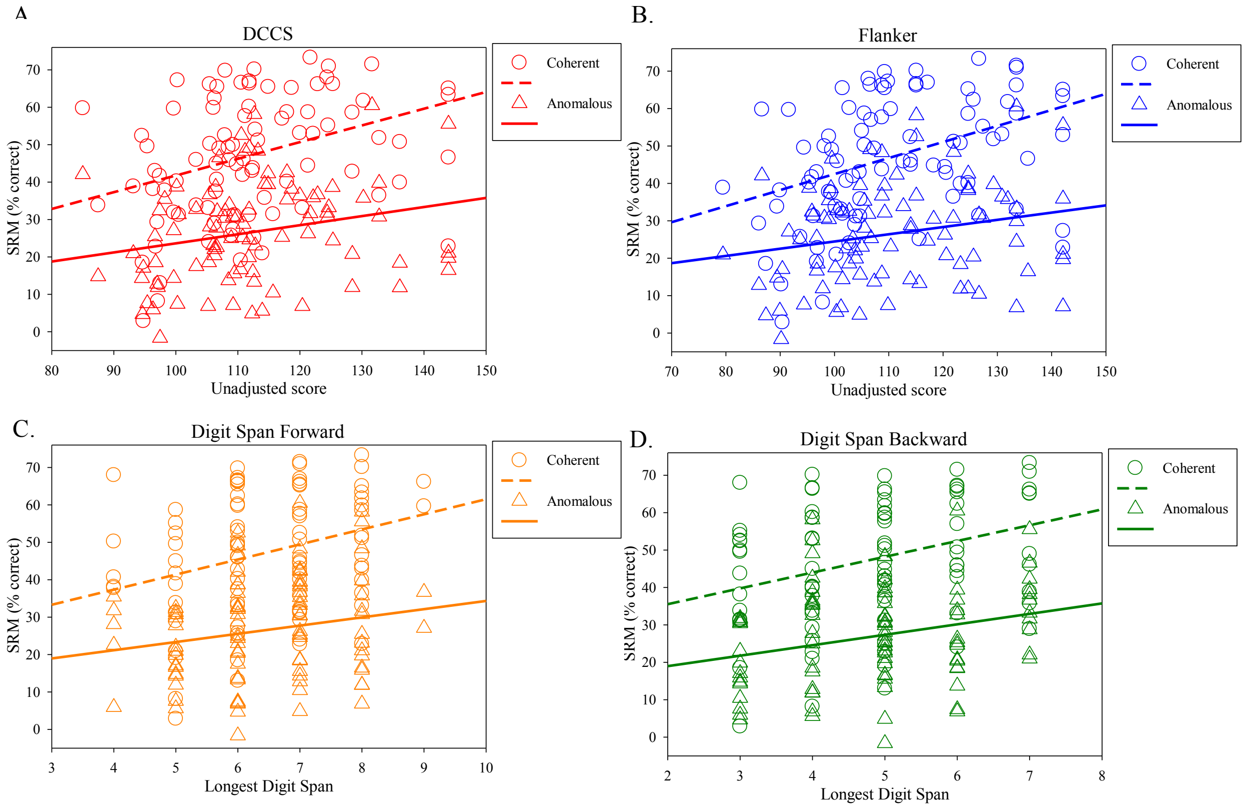

FIG. 8. Simple linear regression showing the relationship between SRM at the -16 dB SNR, for the coherent and anomalous sentences, and scores on each measure of executive function (A: DCCS, B: Flanker, C: Digit Span Forward, D: Digit Span Backward) for all groups combined (n=92 participants). SRM with the coherent sentences, in each panel, is represented by the circles and the dashed line. SRM with the anomalous sentences, in each panel, is represented by the triangles and the solid line.

VII. REFERENCES

1. Arlinger, S., et al., *The emergence of cognitive hearing science.* Scand J Psychol, 2009. **50**(5): p. 371-84.

2. Bregman, A.S., *Auditory scene analysis : the perceptual organization of sound*. 1990, Cambridge, Mass.: MIT Press. xiii, 773 p.

3. Watson, C.S., W.J. Kelly, and H.W. Wroton, *Factors in the discrimination of tonal patterns. II. Selective attention and learning under various levels of stimulus uncertainty.* J Acoust Soc Am, 1976. **60**(5): p. 1176-86.

4. Neff, D.L. and T.M. Dethlefs, *Individual differences in simultaneous masking with random-frequency, multicomponent maskers.* J Acoust Soc Am, 1995. **98**(1): p. 125-34.

5. Lutfi, R.A., *A model of auditory pattern analysis based on component-relative-entropy.* J Acoust Soc Am, 1993. **94**(2 Pt 1): p. 748-58.

6. Patterson, R.D., et al., *The deterioration of hearing with age: frequency selectivity, the critical ratio, the audiogram, and speech threshold.* J Acoust Soc Am, 1982. **72**(6): p. 1788-803.

7. Lutfi, R.A., et al., *One factor underlies individual differences in auditory informational masking within and across age groups.* Percept Psychophys, 2003. **65**(3): p. 396-406.

8. Arbogast, T.L., C.R. Mason, and G. Kidd, Jr., *The effect of spatial separation on informational and energetic masking of speech.* J Acoust Soc Am, 2002. **112**(5 Pt 1): p. 2086-98.

9. Durlach, N.I., et al., *Informational masking: counteracting the effects of stimulus uncertainty by decreasing target-masker similarity.* J Acoust Soc Am, 2003. **114**(1): p. 368-79.

10. Schneider, B.A., L. Li, and M. Daneman, *How competing speech interferes with speech comprehension in everyday listening situations.* J Am Acad Audiol, 2007. **18**(7): p. 559-72.

11. Oh, E.L., F. Wightman, and R.A. Lutfi, *Children's detection of pure-tone signals with random multitone maskers.* J Acoust Soc Am, 2001. **109**(6): p. 2888-95.

12. Wightman, F.L., et al., *Children's detection of pure-tone signals: informational masking with contralateral maskers.* J Acoust Soc Am, 2003. **113**(6): p. 3297-305.

13. E.C., C., *Some experiments on the recognition of speech, with one and with two ears.* J. Acoust. Soc. Am., 1953. **25**: p. 975-979.

14. Zurek, P.M., *A note on onset effects in binaural hearing.* J Acoust Soc Am, 1993. **93**(2): p. 1200-1.

15. Drullman, R. and A.W. Bronkhorst, *Multichannel speech intelligibility and talker recognition using monaural, binaural, and three-dimensional auditory presentation.* J Acoust Soc Am, 2000. **107**(4): p. 2224-35.

16. Jones, G.L. and R.Y. Litovsky, *A cocktail party model of spatial release from masking by both noise and speech interferers.* J Acoust Soc Am, 2011. **130**(3): p. 1463-74.

17. Jones, G.L. and R.Y. Litovsky, *Role of masker predictability in the cocktail party problem.* J Acoust Soc Am, 2008. **124**(6): p. 3818-30.

18. Garadat, S.N. and R.Y. Litovsky, *Speech intelligibility in free field: spatial unmasking in preschool children.* J Acoust Soc Am, 2007. **121**(2): p. 1047-55.

19. Misurelli, S.M. and R.Y. Litovsky, *Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry.* J Acoust Soc Am, 2012. **132**(1): p. 380-91.

20. Johnstone, P.M. and R.Y. Litovsky, *Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults.* J Acoust Soc Am, 2006. **120**(4): p. 2177-89.

21. Litovsky, R.Y., *Speech intelligibility and spatial release from masking in young children.* J Acoust Soc Am, 2005. **117**(5): p. 3091-9.

22. Hess, C.L., *Speech discrimination and spatial hearing in toddlers with bilateral cochlar implants*. iv, 110 leaves.

23. Hawley, M.L., R.Y. Litovsky, and J.F. Culling, *The benefit of binaural hearing in a cocktail party: effect of location and type of interferer.* J Acoust Soc Am, 2004. **115**(2): p. 833-43.

24. Stewart, R. and A. Wingfield, *Hearing loss and cognitive effort in older adults' report accuracy for verbal materials.* J Am Acad Audiol, 2009. **20**(2): p. 147-54.

25. Wightman, F., D. Kistler, and D. Brungart, *Informational masking of speech in children: auditory-visual integration.* J Acoust Soc Am, 2006. **119**(6): p. 3940-9.

26. Wightman, F.L. and D.J. Kistler, *Informational masking of speech in children: effects of ipsilateral and contralateral distracters.* J Acoust Soc Am, 2005. **118**(5): p. 3164-76.

27. Hall, J.W., 3rd, et al., *Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children.* Ear Hear, 2002. **23**(2): p. 159-65.

28. Boothroyd, A., *Auditory development of the hearing child.* Scand Audiol Suppl, 1997. **46**: p. 9-16.
29. Salamy, A., Eggermont, J., & Eldredge, L., *Neurodevelopment and auditory function in preterm infants*. Principles & Applications In Auditory Evoked Potentials. 1994: Allyn and Bacon. 287-312.

30. Werner, L., R.R. Fay, and A.N. Popper, *Human Auditory Development*, in *Springer Handbook of Auditory Research, 42*. 2012, Springer,: Dordrecht. p. 1 online resource (293 p.).

31. Gomes, H., et al., *The development of auditory attention in children.* Front Biosci, 2000. **5**: p. D108-20.

32. Lyon, G.R. and N.A. Krasnegor, *Attention, memory, and executive function*. 1996, Baltimore: P.H. Brookes Pub. Co. xvii, 424 p.

33. Baddeley, A., *Modularity, mass-action and memory.* Q J Exp Psychol A, 1986. **38**(4): p. 527-33.

34. Baddeley, A.D. and J.D. Larsen, *The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative.* Q J Exp Psychol (Hove), 2007. **60**(4): p. 497-504.

35. Baddeley, A., *Working memory: looking back and looking forward.* Nat Rev Neurosci, 2003. **4**(10): p. 829-39.

36. Daneman, M. and P.M. Merikle, *Working memory and language comprehension: A meta-analysis.* Psychon Bull Rev, 1996. **3**(4): p. 422-33.

37. Miller, E.K. and J.D. Cohen, *An integrative theory of prefrontal cortex function.* Annu Rev Neurosci, 2001. **24**: p. 167-202.

38. McCoy, S.L., et al., *Hearing loss and perceptual effort: downstream effects on older adults' memory for speech.* Q J Exp Psychol A, 2005. **58**(1): p. 22-33.

39. Parbery-Clark, A., et al., *Musician enhancement for speech-in-noise.* Ear Hear, 2009. **30**(6): p. 653-61.

40. Rabbitt, P.M., *Channel-capacity, intelligibility and immediate memory.* Q J Exp Psychol, 1968. **20**(3): p. 241-8.

41. Davidson, M.C., et al., *Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching.* Neuropsychologia, 2006. **44**(11): p. 2037-78.

42. Kronenberger, W.G., et al., *Working memory training for children with cochlear implants: a pilot study.* J Speech Lang Hear Res, 2011. **54**(4): p. 1182-96.

43. Pisoni, D.B. and M. Cleary, *Measures of working memory span and verbal rehearsal speed in deaf children after cochlear implantation.* Ear Hear, 2003. **24**(1 Suppl): p. 106S-20S.

44. Pickering, S.J., et al., *Development of memory for pattern and path: further evidence for the fractionation of visuo-spatial memory.* Q J Exp Psychol A, 2001. **54**(2): p. 397-420.

45. Akeroyd, M.A., *Are individual differences in speech reception related to individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults.* Int J Audiol, 2008. **47 Suppl 2**: p. S53-71.

46. Knudsen, E.I., *Fundamental components of attention.* Annu Rev Neurosci, 2007. **30**: p. 57-78.

47. Davis, M.H., et al., *Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI.* J Cogn Neurosci, 2011. **23**(12): p. 3914-32.

48. Cleary, M. and D.B. Pisoni, *Talker discrimination by prelingually deaf children with cochlear implants: preliminary results.* Ann Otol Rhinol Laryngol Suppl, 2002. **189**: p. 113-8.

49. Bronkhorst, A.W., *The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions.* Acustica, 2000. **86**(1): p. 117-128.

50. Stubbs, R.J. and Q. Summerfield, *Algorithms for separating the speech of interfering talkers: evaluations with voiced sentences, and normal-hearing and hearing-impaired listeners.* J Acoust Soc Am, 1990. **87**(1): p. 359-72.

51. Leibold, L.J. and A.Y. Bonino, *Release from informational masking in children: effect of multiple signal bursts.* J Acoust Soc Am, 2009. **125**(4): p. 2200-8.

52. Peissig, J. and B. Kollmeier, *Directivity of binaural noise reduction in spatial multiple noise-source arrangements for normal and impaired listeners.* J Acoust Soc Am, 1997. **101**(3): p. 1660-70.

53. Stickney, G.S., et al., *Cochlear implant speech recognition with speech maskers.* J Acoust Soc Am, 2004. **116**(2): p. 1081-91.

54. Boothroyd, A. and S. Nittrouer, *Mathematical treatment of context effects in phoneme and word recognition.* J Acoust Soc Am, 1988. **84**(1): p. 101-14.

55. Kalikow, D.N., K.N. Stevens, and L.L. Elliott, *Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability.* J Acoust Soc Am, 1977. **61**(5): p. 1337-51.

56. McClelland, J.L. and J.L. Elman, *The TRACE model of speech perception.* Cogn Psychol, 1986. **18**(1): p. 1-86.

57. Casey, B.J., et al., *Imaging the developing brain: what have we learned about cognitive development?* Trends Cogn Sci, 2005. **9**(3): p. 104-10.

58. Wightman, D.J. Kistler, and A. O'Bryan, *Individual differences and age effects in a dichotic informational masking paradigm.* J Acoust Soc Am, 2010. **128**(1): p. 270-9.

59. Pichora-Fuller, M.K., *Cognitive aging and auditory information processing.* Int J Audiol, 2003. **42 Suppl 2**: p. 2S26-32.

60. Pichora-Fuller, M.K., et al., *Temporal jitter disrupts speech intelligibility: a simulation of auditory aging.* Hear Res, 2007. **223**(1-2): p. 114-21.

61. Geers, A., C. Brenner, and L. Davidson, *Factors associated with development of speech perception skills in children implanted by age five.* Ear Hear, 2003. **24**(1 Suppl): p. 24S-35S.

62. Pisoni, D.B. and A.E. Geers, *Working memory in deaf children with cochlear implants: correlations between digit span and measures of spoken language processing.* Ann Otol Rhinol Laryngol Suppl, 2000. **185**: p. 92-3.

63. Tun, P.A. and A. Wingfield, *One voice too many: adult age differences in language processing with different types of distracting sounds.* J Gerontol B Psychol Sci Soc Sci, 1999. **54**(5): p. P317-27.

64. Pisoni, D.B., et al., *Measures of digit span and verbal rehearsal speed in deaf children after more than 10 years of cochlear implantation.* Ear Hear, 2011. **32**(1 Suppl): p. 60S-74S.

65. Hasher, L., R.T. Zacks, and C.P. May, *Inhibitory control, circadian arousal, and age.* Attention and Performance Xvii, 1999. **17**: p. 653-675.

66. Craik, F.I., *The role of cognition in age-related hearing loss.* J Am Acad Audiol, 2007. **18**(7): p. 539-47.

67. Kathleen Pichora-Fuller, M., *Use of supportive context by younger and older adult listeners: Balancing bottom-up and top-down information processing.* International Journal of Audiology, 2008. **47**(S2): p. S72-S82.

CHAPTER 5: Conclusions

This dissertation focuses on the ability of children with NH and with BiCIs to function in noisy environments. As described in Chapter 1, the goals of this work were to: (1) Investigate source segregation abilities in children with NH and with BiCIs for interferers spatially distributed either asymmetrically or symmetrically; (2) Examine the effect of target interferer similarity for these groups; (3) Explore measures of executive function and the relationship with performance on speech-in-noise tasks for NH listeners; (4) Investigate source segregation abilities in NH listeners, from school-age children to adults.

Chapter 2 compared the ability of children with NH and children with BiCIs to use spatial cues in order to segregate target speech from interfering speech. Groups of children with BiCIs or NH were matched according to either chronological age or their amount of auditory experience ("hearing age"). Using spondaic words as target speech, speech reception thresholds (SRTs) were measured in a "Front" condition (both target and interfering speech in front at 0 degrees), in an "Asymmetrical" condition (target in front and interferers asymmetrically distributed toward the right ear at +90 degrees), and in a "Symmetrical" condition (target in front and interferers symmetrically distributed toward both the right and left ears at +/- 90 degrees). In the Asymmetrical condition, both monaural "better ear" and binaural cues are available. In the Symmetrical condition, monaural cues are reduced and listeners must rely heavily on binaural cues to segregate target speech. SRM was computed as the difference between SRTs in the Front condition and either the Asymmetrical or Symmetrical conditions. Results showed that the asymmetrical SRM was smaller in BiCI users than in NH children. Additionally, NH children showed symmetrical SRM, suggesting they are able to use binaural cues for source segregation, whereas children with BiCIs demonstrated minimal or absent Symmetrical SRM. Limitations in the CI device likely also play a role in the reduced SRM.

Chapter 3 explored performance of children with NH and with BiCIs when both the target and interferers were spoken by a talker of the same sex (both male). This contrasts with the previously described study [1], in which the target and interferer were different-sex talkers (target = male talker; interferers = female talker). This comparison allowed investigation of speech intelligibility within an auditory environment that contained different amounts of informational masking (i.e., same-sex target and interferer results in more informational masking). SRM has been shown to be greater for NH adults in tasks that contain speech interferers than in tasks that contain non-linguistic interferers [2]. This result suggests that when informational masking is present, spatial cues become more important.

It has been shown that NH adult listeners demonstrate a reduction in speech intelligibility (increase in SRTs) on tasks with same-sex target and interferer stimuli [3], and that SRM tends to be larger in those conditions [4, 5]. We investigated this in children with NH and with BiCIs. Children with NH demonstrated an increase in SRTs with the male interferer, particularly in the most difficult listening condition, with the target and interferers co-located. Although children with NH showed a trend of more SRM with the same-sex interferers vs. different-sex interferers, only the older group showed a significant difference. Lack of significant differences in SRM with the male versus female interferer are likely due to within-group variability. All NH groups showed more SRM than either of the BiCI groups, suggesting that listeners with NH are better able than listeners who use BiCIs to make use of spatial cues to improve speech understanding in noise.

Children who use BiCIs have difficulty distinguishing target speech from interfering speech, even in situations with less informational masking in which different-sex talkers are used [1, 6]. This may be due to the fact that the CI does not provide spectral cues with clarity, rendering performance under same-sex talker conditions similar to that with different-sex talkers. In general, children with BiCIs demonstrated more masking (higher SRTs) than NH children with either type of stimulus, especially when the target and the interferers were spatially separated (i.e., Asymmetrical and Symmetrical conditions). It is possible that spatially separating the interferers poses a more difficult listening environment than when the stimuli are co-located. This may be partially due to the microphone characteristics of the CI and also the lack of coordination between CI devices at each ear, rendering binaural cues to be minimal or absent [7, 8].

Together, the experiments described in Chapter 2 and 3 of this dissertation showed large amounts of variability in performance on the speech-in-noise task within and between each subject group, for both children with NH and with BiCIs. We confirmed that individual performance remains stable when all participants were tested on the speech-in-noise task, using a repeated measures design (see Chapter 2). Given this, variability is likely due to integral differences between subjects, rather than within.

It was not surprising to see variability within the BiCI groups. In fact, variability is a hallmark of CI research and can be accounted for by a multitude of factors, including demographics, anatomical differences, and device limitations. It was of particular interest that NH groups showed a considerable amount of variability, given that, unlike in CI users, the peripheral auditory system of NH listeners is intact at birth [9]. Previous literature has suggested that differences between NH listeners may originate in more central mechanisms [10].

In Chapter 4, we examined masking, source segregation abilities, and executive function for a continuum of NH listeners ages 7-23 years. To investigate abilities across a large age range, we used a more complex auditory task than in previous studies (see Chapters 2 and 3). The target stimuli used in Chapter 4 for the speech-in-noise task (see also Davis *et al.* [11]) consisted of open-set sentences that varied in semantic content, either coherent or anomalous. Previous work has used simple tasks of one-word identification or closed-set stimuli. We hypothesized that the auditory stimuli used in this study would better elicit mechanisms involved in executive function due to its inherent complexity (i.e. open-set, varying in semantic content). Furthermore, we investigated SRM with different types of maskers, either nosise or speech (the latter of which was intended to create a more informational masker).

Our results revealed numerous findings. First, less masking was shown with an increase in age. Additionally, more SRM was shown with an increase in age, especially at the least favorable SNRs and in conditions that contained the speech (vs. noise) interferer. These results coincide with previous research showing that children perform worse than adults on tasks of auditory source segregation [12-14]. Furthermore, auditory components related to source segregation continue to develop throughout adolescence [15]. We also showed that semantic content has an effect when listening to speech-in-noise, such that less masking was demonstrated with the coherent vs. anomalous sentences, suggesting that semantic content provides a cue that can aid in release from masking. Amount of SRM with the coherent vs. anomalous sentences varied for each SNR, especially with the speech interferer. At the least favorable SNRs, SRM was higher for the coherent sentences, suggesting that listeners used coherent content to aid in segregation of sources beyond the cues that were available due to spatial separation. At the most favorable SNRs, SRM was often higher for the anomalous sentences, particularly for younger groups with the speech interferer. This suggests that spatial separation of sources at favorable SNRs was most beneficial when other cues, such as semantic content, were not available.

As shown in Chapters 2 and 3, performance on the speech-in-noise task described in Chapter 4 also varied between listeners. Because of this, in Chapter 4 we investigated measures of executive function that were hypothesized to have a predictive relationship with performance on the speech-in-noise task. Results showed a positive linear relationship between measures of executive function, specifically attention and memory, and SRM for only the least favorable SNR tested. These significant relationships disappeared when expressive vocabulary was accounted for, as well as at the more favorable SNRs.

The results of these studies propose findings that are worthy of further examination, as well as suggest future directions of study. First, because we found relationships between SRM and measures of executive function when no other factors were accounted for, it would be useful to further investigate this finding to gain a better understanding of how these relate in NH populations. To further explore these relationships, it may be beneficial to focus on one age group at a time, and reduce the number of variables in the study design in order to gain more power within each group. It is possible that stronger relationships were not seen within groups due to numerous within group variables (e.g. interferer type, semantic content). In the future, it will be valuable to compare these results to clinical populations (e.g., BiCI users) to help predict performance with assistive listening devices. If a set of measures that are highly predictive of performance on realistic tasks of auditory source segregation is found, it is possible that these could be used as training mechanisms that aid in improved performance in noisy environments for all populations.

I. REFERENCES

1. Misurelli, S.M. and R.Y. Litovsky, *Spatial release from masking in children with normal hearing and with bilateral cochlear implants: effect of interferer asymmetry.* J Acoust Soc Am, 2012. **132**(1): p. 380-91.

2. Hawley, M.L., R.Y. Litovsky, and J.F. Culling, *The benefit of binaural hearing in a cocktail party: effect of location and type of interferer.* J Acoust Soc Am, 2004. **115**(2): p. 833-43.

3. Brungart, D.S., *Informational and energetic masking effects in the perception of two simultaneous talkers.* J Acoust Soc Am, 2001. **109**(3): p. 1101-9.

4. Bronkhorst, A.W., *The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions.* Acustica, 2000. **86**(1): p. 117-128.

5. Culling, J.F., M.L. Hawley, and R.Y. Litovsky, *The role of head-induced interaural time and level differences in the speech reception threshold for multiple interfering sound sources.* J Acoust Soc Am, 2004. **116**(2): p. 1057-65.

6. Litovsky, R.Y., et al., *Bilateral cochlear implants in children: localization acuity measured with minimum audible angle.* Ear Hear, 2006. **27**(1): p. 43-59.

7. Litovsky, R.Y., et al., *Studies on bilateral cochlear implants at the University of Wisconsin's Binaural Hearing and Speech Laboratory.* J Am Acad Audiol, 2012. **23**(6): p. 476-94.

8. van Hoesel, R.J., *Exploring the benefits of bilateral cochlear implants.* Audiol Neurootol, 2004. **9**(4): p. 234-46.

9. Boothroyd, A., *Auditory development of the hearing child.* Scand Audiol Suppl, 1997. **46**: p. 9-16.

10. Lutfi, R.A., et al., *One factor underlies individual differences in auditory informational masking within and across age groups.* Percept Psychophys, 2003. **65**(3): p. 396-406.

11. Davis, M.H., et al., *Does semantic context benefit speech understanding through "top-down" processes? Evidence from time-resolved sparse fMRI.* J Cogn Neurosci, 2011. **23**(12): p. 3914-32.

12. Wightman, F.L. and D.J. Kistler, *Informational masking of speech in children: effects of ipsilateral and contralateral distracters.* J Acoust Soc Am, 2005. **118**(5): p. 3164-76.

13. Hall, J.W., 3rd, et al., *Spondee recognition in a two-talker masker and a speech-shaped noise masker in adults and children.* Ear Hear, 2002. **23**(2): p. 159-65.

14. Litovsky, R.Y., *Speech intelligibility and spatial release from masking in young children.* J Acoust Soc Am, 2005. **117**(5): p. 3091-9.

15. Gomes, H., et al., *The development of auditory attention in children.* Front Biosci, 2000. **5**: p. D108-20.