12-2023

Spatial Release from Masking in Anechoic and Reverberant Environments

Drake Andrew Hintz
University of Nebraska-Lincoln

12-2023

Spatial Release from Masking in Anechoic and Reverberant Environments

Drake Hintz

SPATIAL RELEASE FROM MASKING IN ANECHOIC AND REVERBERANT ENVIRONMENTS

By

Drake Andrew Hintz

A THESIS

Presented to the Faculty of

The Graduate College at the University of Nebraska at Lincoln

In Partial Fulfillment of Requirements

For the Degree of Master of Science

Major: Architectural Engineering

Under the Supervision of Dr. Lily M. Wang

Lincoln Nebraska

December 2023

SPATIAL RELEASE FROM MASKING IN ANECHOIC AND REVERBERANT ENVIRONMENTS

Drake Andrew Hintz, M.S.

University of Nebraska – Lincoln, 2023

Advisor: Dr. Lily M. Wang

Listening with both ears provides children with access to binaural and monaural cues that are helpful for understanding speech in competing babbles. Specifically, when the target and masker are spatially separated, children can gain an intelligibility benefit which is known as spatial release from masking (SRM). Recent work [Peng et al., 2021 JASA] suggested that school-age children demonstrated immature SRM using binaural cues that are distorted by reverberation. In this follow-up study, we further investigate the effect of reverberant distortion on individual auditory spatial cues, namely binaural and monaural head shadow cues. We compare SRM between adults and school-age children with typical hearing using the novel measure of minimum angle of separation (MAS) between target and masker for which individual achieves a 20% intelligibility gain, in both virtually simulated anechoic and reverberant environments. MAS was measured in both binaural and monaural hearing conditions, as well symmetric versus asymmetric masker displacement to probe access to various auditory cues of interest. Preliminary results show statistically significant effect on MAS when comparing binaural and monaural conditions as well as when comparing symmetric and asymmetric masking conditions. Binaural listening seems to have a positive effect on MAS. Asymmetric masking also seems to improve MAS.

Acknowledgements

First and foremost, I would like to thank my wife, Katia who has supported me unconditionally through this journey. Throughout the research and thesis writing process, Katia has served as my sounding board, proofreader, cheerleader, and a dependable shoulder to lean on. Most importantly, she's been an incredible wife and mother to our daughter. Secondly, I would like to thank my daughter, Eleanora, who was born in the middle of my thesis writing effort. While she couldn't offer much in the way of proofreading or, especially, time management, she brought a joy into my life that I couldn't have imagined before her and now couldn't imagine being without. I love you Eleanora and I am so lucky I get to be your dad. Next, I would like to give a shoutout to my parents, Mike and Sharla. Their support through this endeavor was a continuation of the support they have always had for me through every other stage of my life. I am beyond grateful to have parents who I know will always pick up the phone and help me work through anything from this thesis to taxes to babysitting. I hope that someday Eleanora, and any of her siblings that come along, see me the same way I see you both.

I would also like to acknowledge my workplace, IP Design Group, and particularly my boss, Jessica Hiatt. They've made it remarkably easy to balance my work responsibilities with my academic pursuits. It was reassuring to know that my workplace not only endorsed my pursuit of a master's degree but would be accommodating when it came to scheduling conflicts and understanding when I needed to prioritize my thesis work.

Lastly, I would like to thank my advisors, Dr. Lily Wang and Dr. Ellen Peng. You both have been exceptionally supportive throughout every phase of this journey. My study is a portion of a larger ongoing effort being run at Boys Town National Research Hospital by Dr. Ellen Peng. Ellen taught me everything I needed to know for the study then trusted me enough to head a portion of it. I truly couldn't have done any of this without the groundwork she laid or without her guidance through it all. Dr. Lily Wang was my first contact in the world of acoustics and took me under her wing. She has been there at every step helping me along the way, from my application to the master's program to my application to graduate. Thank you both so much for your guidance and knowledge and wisdom.

I am lucky to be surrounded by such an incredible support system, those mentioned here and many others as well. Each one of you provided the encouragement and support I needed to reach this milestone. I couldn't have done it without all your faith in me, and I am so grateful to have you all in my life.

Table of Contents

List of Multimedia Objects

Tables

Figures

1 Introduction

1.1 Introduction to Work

Spatial release from masking (SRM) represents a critical component of our understanding of auditory perception. The phenomenon of SRM involves the improvement of speech comprehension when the source of the desired speech is spatially separated from the interfering sounds. The ability to discern specific sounds amidst a background of noise is vital for effective communication, especially in complex acoustic environments. Despite its significance, specific mechanisms and influential factors associated with SRM remain to be fully understood in normal-hearing children who are developing various auditory skills. Therefore, there is a need for additional research into this phenomenon.

This study delves into some intricacies of SRM, focusing on auditory cues such as the binaural interaural level difference (ILD) and interaural time difference (ITD) cues, as well as the monaural head shadow cue. It utilizes a set of conditions to compare symmetric spatial separation to asymmetric spatial separation and binaural listening to monaural listening to investigate the influence various auditory spatial cues on SRM performance. A key aspect of this research involves the innovative application of the minimum angle of separation (MAS) metric, which quantifies SRM using the smallest angular gap needed between the target speech and masking speech to achieve a substantial improvement in speech intelligibility.

Additionally, the impact of reverberation on SRM is investigated. Reverberation is a common attribute of indoor real-world acoustic environments. Understanding its impact on SRM can provide important insights for improving speech intelligibility in more realistic settings. This study also evaluates the performance differences across age groups by comparing the SRM abilities of adults and children in various listening conditions.

Ultimately, this research aims to contribute to our understanding of SRM during typical development and the factors that affect it to inform the design of assistive listening devices and acoustic environments. It is the goal of this paper to both fill gaps in our current knowledge of SRM and point the way for future research in this critical area of auditory science.

1.2 Outline of Thesis

Chapter 2 provides a comprehensive review of existing literature on SRM, exploring the key roles of various auditory cues, the influence of masking types, the application of the MAS metric, and the impact of reverberation on SRM.

Chapter 3 details the methodology used in this research to explore SRM in various listening conditions. The experimental setup, testing procedures, and data analysis techniques are described in depth to assist others in replicating the study and testing reliability of the findings.

Chapter 4 presents the findings of this study, discussing how various factors, such as reverberation, type of listening condition, and age can impact SRM. The chapter also provides an in-depth analysis of the results and their implications.

Chapter 5, the conclusion, offers a synthesis of the research findings and their implications. It summarizes the key results, highlights the potential applications of this study, and suggests potential avenues for future research.

2 Literature Review

2.1 Introduction

Understanding speech in noisy environments is a challenge faced by everyone in everyday life. SRM is a phenomenon that enables improved speech understanding by spatially separating the target speech and the interfering sounds (Litovsky, 2012). This literature review aims to provide a comprehensive understanding of SRM, focusing on the role of reverberation and its impact on ILD, ITD, and head shadow cues, as well as examining the influence of symmetric vs. asymmetric masking and monaural vs. binaural listening. Furthermore, it will discuss the metric of MAS and its application in various listening conditions as investigated in current research.

2.2 Spatial Release from Masking and Auditory Cues

SRM relies heavily on the auditory system's ability to use auditory spatial cues, particularly ITD and ILD. ITD refers to the difference in time it takes for a sound to reach each ear, while ILD is the difference in sound intensity between the two ears, which is primarily caused by the head shadow effect. These cues help the listener localize sounds and segregate speech from competing noise sources (Middlebrooks & Green, 1991).

Head shadow is a phenomenon that occurs when the listener's head obstructs the direct path of sound to the contralateral ear, creating an acoustic shadow. This shadow causes a difference in sound intensity, ILD, between the ears, which helps the listener localize the sound source (Moore, 2012). With the head shadow cue listeners can listen to the ear with a better signal-to-noise ratio of the target sounds, which is monaural listening that arises from binaural cues. Monaural head shadow and binaural ITD and ILD cues are crucial for the spatial hearing involved in SRM.

Apart from ITD, ILD and head shadows, other auditory cues, such as spectral cues, play a role in spatial hearing. These cues are essential for understanding how listeners perceive sound sources in various listening environments and how they segregate speech from competing noise sources. Spectral cues arise from the filtering properties of the listener's outer ear, specifically the shape of the pinna. These cues provide information about the elevation and front-back localization of sound sources (Musicant & Butler, 1984). Spectral cues can interact with ILD and ITD, and this interaction affects the perception of sound sources in different listening environments.

2.3 Quantifying Spatial Release from Masking

Various methods have been used to quantify SRM in previous studies (Freyman et al., 1999; Litovsky, 2005; Bronkhorst, 2015). These methods provide a means to compare the benefits of spatial separation between the target speech and the interfering sounds across different listening conditions and populations. One such metric is the MAS, which offers a specific way to quantify SRM in terms of spatial separation.

Traditionally, SRM has been measured by comparing speech intelligibility using speech reception thresholds (SRTs) between co-located and spatially separated conditions. In the co-located condition, the target speech and masker are presented from the same location. Conversely, in spatially separated conditions, the target speech and masker are presented from different directions. SRT is then measured in both conditions in an experiment where participants are read sentences and instructed to repeat what the sentence said. This allows an SRT to be measured based on the accuracy with which a participant can repeat the sentence. The difference in speech intelligibility, as measured by SRTs, between these conditions represents the SRM benefit (Peng & Litovsky, 2022).

The MAS is a metric that quantifies SRM based on the smallest angular separation between the target speech and the masking speech that yields a significant (~20%) improvement in speech intelligibility (Peng & Litovsky, 2021). By measuring the angle at which listeners can effectively isolate speech from interfering sounds, MAS provides a more detailed understanding of the spatial resolution of auditory processing. Incorporating MAS into the analysis of SRM allows researchers to examine the influence of different listening conditions and environmental factors on spatial hearing more accurately. Some of these conditions are reverberant environments, symmetric vs. asymmetric masking, and monaural vs. binaural listening. Additionally, comparing MAS values across various conditions and populations can provide valuable insights into the underlying factors of SRM. This information can inform the development of assistive listening devices, and it can steer the development of acoustic designs that promote better speech understanding for all listeners.

Reverberation, which is caused by sound reflections in an environment, can alter speech intelligibility and SRM (Culling et al., 2003; Beutelmann & Brand, 2006; Lavandier & Culling, 2010). The role of early and late reflections in SRM has been studied and shows that late reflections can reduce the benefit of SRM, especially in listeners with hearing loss (Srinivasan et al., 2017). Goupell et al. (2012) found that an additional reflection in a precedence effect experiment can negatively influence the localization of sounds. These studies suggest that reverberation may reduce the salience of spatial cues for them to be useful for listeners. Additional reflected sound reaching both ears will change perceived arrival times and levels of sound, reducing time difference and level difference benefits that auditory cues use. The research described in this thesis aims at understanding the impact of reverberation on MAS in a variety of listening conditions.

Symmetric and asymmetric masking refers to the spatial configurations of the masking voices relative to the target sound source and the listener. Symmetric masking involves positioning the maskers at equal but opposite angles from the listener in relation to the target sound, creating a balanced sound field. This setup means that the masking noise impacts the listener equally from both sides. Conversely, asymmetric masking places both maskers on the same side of the listener. This results in an imbalanced sound field, with the masking noise predominantly coming from one side. The effects of symmetric and asymmetric masking on SRM have been investigated and it has been observed that asymmetrical masking can yield larger SRM (Misurelli & Litovsky, 2012). This is likely due to the monaural head shadow cue being available for listeners in asymmetric masking conditions and not in symmetric masking conditions.

Several studies (Litovsky, 2005; Johnstone and Litovsky, 2006; Litovsky et al., 2006; Garadat and Litovsky, 2007) reported that in a population of children, SRM was affected by interferer asymmetry with greater masking release in the presence of asymmetric maskers. Binaural sensitivity and release from speech-on-speech masking have been studied in listeners with and without hearing loss (Baltzell et al., 2020). Furthermore, Peng and Litovsky (2021) explored the role of ILD, head shadow, and binaural redundancy in SRM among school-aged children, comparing how these factors contribute to binaural hearing advantages. In this thesis, MAS data are compared between symmetric and asymmetric masking conditions using monaural and binaural listening in both anechoic and reverberant environments, for adults and children.

2.4 Experimental Designs and Methodologies

The studies included in this literature review employed various experimental designs and methodologies to investigate SRM. The experimental designs include studies using sound booth laboratories with near-anechoic conditions (Best et al., 2012) and studies using simulated anechoic and reverberant environments (Biberger & Ewert, 2019). It has been observed that simulated rooms and real rooms have a close correspondence in a variety of perceived acoustical parameters (Wendt et al., 2014). While controlled laboratory studies can isolate specific factors influencing SRM, more ecologically valid studies may provide insights into how these factors interact in real-world settings. The current research bridges this gap by testing MAS in simulated listening conditions, both reverberant and anechoic environments, allowing for a more comprehensive understanding of how reverberation impacts SRM.

The literature reveals some inconsistencies regarding the effects of reverberation on SRM. For instance, while some studies report a negative impact of reverberation on SRM (Lavandier & Culling, 2010; Biberger & Ewert, 2019; Peng et al., 2021), others suggest that early reflections may not always impair SRM (Srinivasan et al., 2017). The current research addresses these inconsistencies by quantifying SRM using MAS and exploring the impact of reverberation on MAS in various listening conditions, contributing valuable insights to the field.

Despite the valuable insights provided by previous studies on spatial release from masking, several limitations must be acknowledged. These limitations may affect the generalizability and applicability of the findings in real-world settings. This section outlines some of the common limitations encountered in SRM research.

Many studies on SRM have had relatively small sample sizes, which can limit the statistical power of the findings and increase the risk of Type II errors (i.e., failing to detect a true effect). Additionally, some studies may suffer from selection bias, as participants are often recruited from a limited pool, such as university students or individuals with specific types of hearing loss. This can limit the generalizability of the results to other populations.

While previous studies have focused on the role of ITD and ILD in SRM, other auditory cues, such as spectral cues and monaural localization cues (Wightman & Kistler, 1997), have received less attention. A more comprehensive understanding of SRM requires a thorough investigation of all auditory cues involved in spatial hearing and their interactions in different listening environments.

SRM research has employed a wide range of experimental designs and methodologies, which can make it challenging to directly compare findings across various studies. The use of different stimuli, maskers, and listening conditions can lead to discrepancies in the reported effects of various factors on SRM. Standardizing experiment protocols and statistical analysis techniques could help overcome this limitation and allow for more consistent conclusions about the underlying factors of SRM.

2.5 Summary

This literature review has provided an overview of SRM and its relationship with auditory cues, reverberation, symmetric vs. asymmetric masking, and monaural vs. binaural listening. The metric of MAS offers a unique approach to quantify SRM, allowing for a deeper understanding of spatial hearing in different listening conditions. The current research aims to investigate the influence of reverberation on MAS in symmetric and asymmetric masking conditions and examine the differences between monaural and binaural listening in both anechoic and reverberant environments.

3 Methodology

The purpose of this study was to test SRM in normal hearing adults and normal hearing children in different listening conditions. Spatial release from masking was tested through measuring MAS in children and adults for comparison. In this study, we investigated the effect of reverberant distortion on auditory spatial cues, namely, ITD, ILD and binaural and monaural head shadow cues.

3.1 Facilities

3.1.1 Boys Town National Research Hospital AV Lab

All data collection was done at the Boys Town National Research Hospital AV Lab in Omaha, Nebraska. The AV lab is a rectangular room with dimensions of approximately 10 feet tall x 20 feet deep x 15 feet wide (**Error! Not a valid bookmark self-reference.** and Figure 3.2). It is a carpeted room (1/8" pile carpet) with bass traps in the corners of the room and acoustical absorption pads covering 95-99% of the walls and ceiling. The room's background noise level was measured to be about 30 dBA, and the reverberation time was measured as 0.34 seconds at 500Hz, 0.44 seconds at 1000Hz, and 0.55 seconds at 2000Hz. Background noise level measurements were conducted using the NIOSH SLM app, and reverberation time measurements were performed using a series of claps and the ClapReverb app on an iPhone 13 Pro Max.

Figure 3.1 - Boys Town National Research Hospital AV Lab Studio

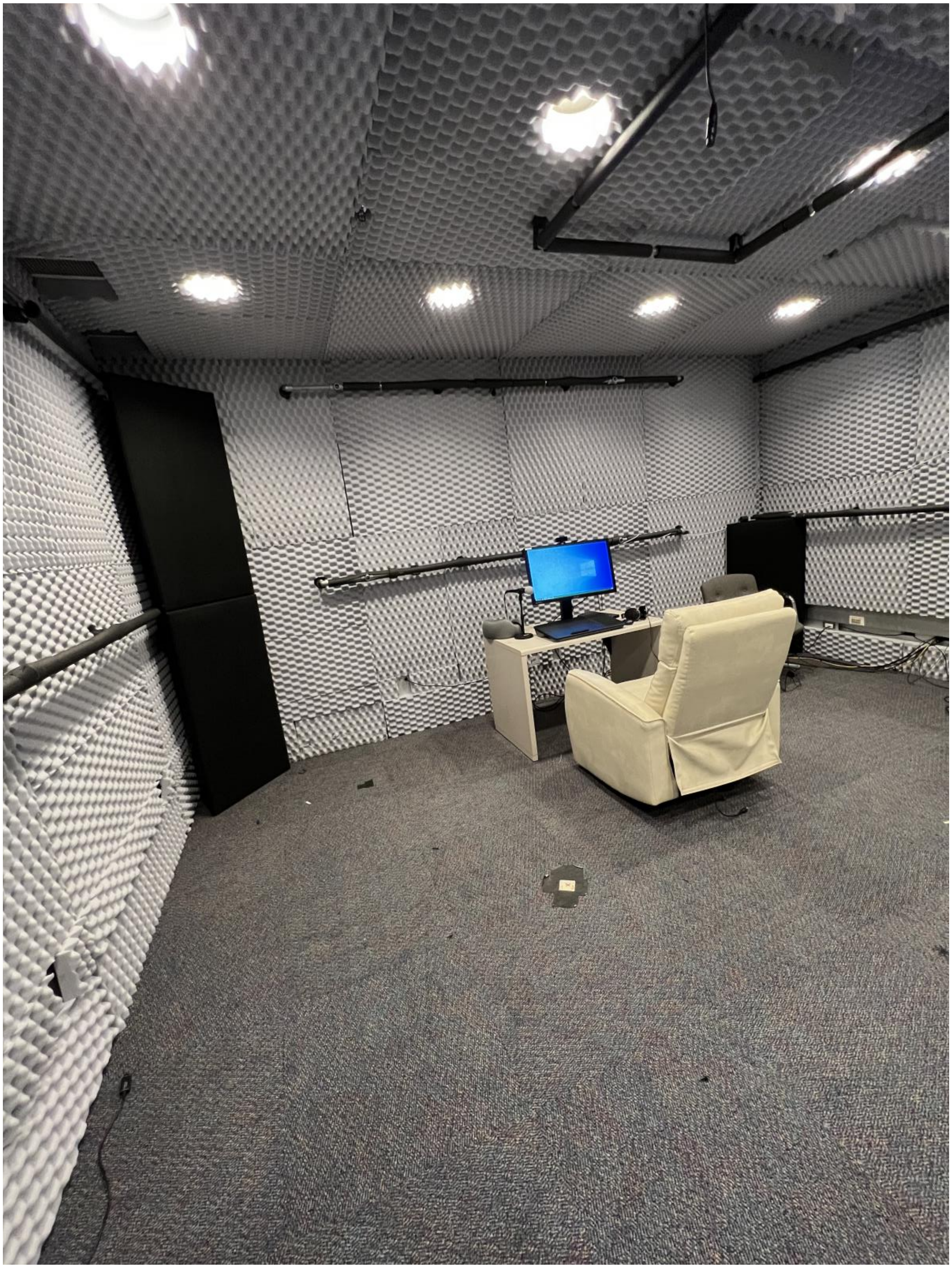


Figure 3.2 – Boys Town National Research Hospital AV Lab Studio

3.2 Experimental Methods

A total of 23 adults and 15 children participated in this study. Each participant's visit lasted approximately 2-2.5 hours, and breaks were provided to prevent mental fatigue. The adults had a mandatory break around the halfway point, while the children were given two breaks at the 1/3 and 2/3 checkpoints. Additional breaks were given if the test administrator noticed signs of fatigue, loss of motivation, or inattention. Adult participants signed electronic consent forms, and consent was obtained for the children via electronic consent forms signed by a parent or guardian on behalf of their children in addition to the child participants completing electronic assent forms. Participants received $15 per hour for their participation in the study. As the study involved research on human subjects, it was approved by the Institutional Review Board (IRB) of Boys Town National Research Hospital.

3.2.1 Hearing Screening

Before commencing the study, all participants underwent a hearing screening using a Grason-Stadler GSI AudioStar Pro™ type audiometer. The purpose of the hearing screening was to ensure participants met the required hearing levels for inclusion in the study. During the screening, participants wore over-the-ear headphones and were presented with pulsed sine wave tones at varying frequencies in one ear at a time. Participants pressed a clicker to indicate when they heard a sound. The screening covered octave bands ranging from 125 Hz to 8000 Hz, starting at 20 decibels hearing level. If the participant could hear at 20 decibels, the next octave band was tested. If the participant failed to hear at 20 decibels, the signal was increased to 25 decibels and further testing was conducted. Participants who could not hear the signal at 25 decibels did not meet the qualifications for the study. Only one participant failed to hear a signal at 25 decibels and was therefore excluded from the study. Figure 3.3 shows a sample of the hearing screening document.

3.2.2 Covid Screening

At the beginning of each visit, participants underwent a Covid screening process. The participants, or the participants' guardians in the case of the children participants, were verbally asked a series of questions about Covid-related symptoms and situations, and their answers determined whether they were allowed to participate in the study on that day. Throughout the study, no participants needed to reschedule due to the Covid screening. The Covid screening document given to participants is provided in Figure 3.4.

Figure 3.3 – Hearing Screening Document

Instruction: Results need to be filed & recorded on REDcap

Tester Initials: _______

Study ID: S___________ Participant ID: __________ Screening Date: ___________

RIGHT Ear:

	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
Screened at 25 dB HL							
Screened at __ dB HL							

LEFT Ear:

	125 Hz	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
Screened at 25 dB HL							
Screened at __ dB HL							

Did participant pass hearing screen? ____Yes ____ No

Notes:

Figure 3.4 – Covid Screening Questionnaire

Functional Hearing Laboratory

COVID-19 Screener for Research Participants

(1) In the past 10 days, have you been tested positive for COVID-19?
(2) Have you been in close contact with someone who has tested positive for COVID-19? (Close contact means you've been within 6 feet of someone who is infected for at least 15 minutes)
(3) Do you currently or have you in the past 14 days had any new COVID-19 symptoms? • Fever > 100.4 F • Sore throat • Cough • Nasal congestion or runny nose • Loss of taste or smell • Shortness of breath • Nausea, vomiting, or diarrhea

If you answer "Yes" to any of the above questions, we kindly ask you to reschedule the test appointment to a later date.

3.2.3 Target Sentences and Masker Stories

In this study, we utilized open-set sentences from the Australian Speech in Noise Test (AuSTIN) as the target speech (Dawson et al., 2013). The AuSTIN corpus consists of 1280 Bamford-Kowal-Bench-like sentences suitable for speech intelligibility testing with children as young as 6 years old. Open-set sentences were chosen to accurately reflect standard verbal communication and align with the developmental effects of speech perception in noise better than word recognition methods (Cameron et al., 2006; Corbin et al., 2016; Griffin et al., 2019). A subset of 596 AuSTIN sentences, each containing three keywords, was used for keyword-based scoring. The AuSTIN sentences were recorded by a 26-year-old female native American English speaker from the Midwest with a minimal regional accent. The average fundamental frequency of the target talker was 195 Hz. The AuSTIN sentences used in this work were slightly modified for American English speakers and re-recorded with a female speaker from the Upper Midwest region with minimal regional accent (Peng & Litovsky, 2021). The sentences used in for target speech contained words that were well within the vocabulary of all participants, adults and children. However, it's worth noting that we did not specifically re-confirm the psychometric functions of the re-recorded AuSTIN sentences for American English-speaking children. Nevertheless, the deliberate choice of vocabulary ensured that all participants, regardless of age, could readily comprehend the content.

The masker consisted of a two-talker dialogue featuring short science stories (TIME USA, n.d.) suitable for first-graders. Using a two-talker masker limits the ability to "glimpse" words during silent gaps and effectively masks the target speech (Buss et al., 2017). The masker dialogue was recorded by a second female talker, also from the

Midwest, aged 22, with a minimal regional accent. The average fundamental frequency of the masker talker was 251 Hz, slightly higher than that of the target female talker. Both talkers had a similar speech rate of 3.5 syllables per second. Employing a two-talker, same-sex masker maximized informational masking, which allowed listeners to rely mostly on auditory spatial cues for unmasking (Brungart et al., 2001; Freyman et al., 2001; Johnstone & Litovsky, 2006; Cameron & Dillon, 2007; Misurelli & Litovsky, 2015; Leibold et al., 2018; Leibold et al., 2020). The masking speech's dialogue contained scientific jargon and technical terminology that was explained in the stories being told. Both adults and children would likely be able to understand the dialogue if they listened to it.

3.2.4 Testing Procedure

During the main part of the study, participants sat in a chair at a desk with a computer in the Boys Town National Research Hospital's AV lab studio. They wore Sennheiser HD 650 open-back over-the-ear headphones, and the door was closed to prevent interference from external sounds. The test administrator used a computer linked to the one in the lab studio to control the experiment. The test administrator then provided a brief explanation of the study to the participant through a microphone connected to the participant's headphones.

3.2.4.1 Testing Process

The study consisted of 14 experimental "runs," each containing approximately 10-30 "trials," designed to test either SRT or MAS under various listening conditions. The study presented a pre-recorded female voice as the target voice speaking sentences to the participant. The participant's task was to repeat the sentence exactly as they believed they had heard it, regardless of its understandability and including any nonsense syllables the participant believed they had heard. Each sentence consisted of five or six words with three keywords.

We considered a “trial” to be one instance of a sentence being played for the participant and the participant repeating back what they believed they heard. The test administrator would then compare what the participant repeated to what the sentence said and count the number of keywords the participant got correct to determine whether the trial was completed successfully or unsuccessfully. A trial was considered unsuccessful if the participant repeated zero or one key word correctly, and the trial was considered successful if the participant repeated two or three keywords correctly. Contractions, pronouns, prepositions, and other articles in a sentence were generally not considered keywords. For example, if a participant repeated, "THE lady crossed A road," but the actual sentence was, "A lady crossed THE road," it would still be considered correct, with the keywords being "lady," "crossed," and "road."

After two runs testing SRT in quiet conditions, masking voices were added to the sounds presented to the participant. The maskers were additional voices that overlapped with the target speech. The masking voices were female voices of a similar frequency to that of the target voice. Two masking voices were always used simultaneously, each telling a pre-recorded story. It is important to note that the two masking voices told different stories to prevent duplicate auditory signals from being perceived as louder due to multiple sources. The participant’s task was to attempt to listen to the target voice reading the sentence while the masking voices provided competing babble.

3.2.4.2 Test Set Up

During an SRT run, the target level started at 60 decibels, and, if maskers were used, their level was set at 55 decibels. If a participant was unsuccessful in a trial, the target level increased, and if a participant was successful in a trial, the target level decreased. This process is called the "one up one down" method and was employed during SRT trials to determine when to adjust the target level. The "one up one down" method results in an SRT value representing the signal to noise ratio (SNR) at which participants achieved 50% speech comprehension. The run concluded when the direction of change in target voice level switched (e.g., decreasing level changing to increasing level or vice versa) seven times. This change in direction was called a reversal.

Similarly, MAS trials concluded after seven reversals. However, in MAS runs, it was the angle of the maskers that changed, not the level of the target. As when testing for an SRT with masking, the masker level in MAS runs was set at 55 decibels. In MAS runs, the target level was set at the SRT value found in the previous run. In MAS runs, a "two up one down" method was used to adjust the masker positions. This meant that two consecutive correct trials led to the maskers moving closer to zero degrees, while a single failed trial caused the maskers to move further away from zero degrees. This method results in a MAS value that represents an SNR where the participant achieved 70.7% speech comprehension of the target voice for each condition. It should be noted that the MAS runs needed to find an angle that yielded greater than 50% speech comprehension due to using an SRT value representing a 50% comprehension point as the baseline. If we had attempted to find a MAS representing 50% comprehension, the resulting value would have been 0 degrees regardless of the listening conditions. By employing the "two-up one down" method, we tested for an angle that yielded 70.7% speech comprehension, which exceeded the required 50%, ensuring that the obtained angle for MAS would be greater than 0 degrees.

The study was conducted using a custom MATLAB program to control the trial-to-trial procedures. This program allowed for adjustments in level, positioning, starting angles for target and maskers, as well as selecting which sentence would be played for each trial. The SRT runs had the magnitude of the change in level at each of the seven reversals set to be as follows: 8 decibels, 4 decibels, 2 decibels, 2 decibels, 2 decibels, 2 decibels, and 2 decibels. The MAS runs had the magnitude of the change in masker angles at each of the seven reversals set to be as follows: 40 degrees, 20 degrees, 10 degrees, 5 degrees, 5 degrees, 5 degrees, and 5 degrees.

After completing a run, the Matlab program analyzed the data to determine the 50% speech comprehension point for SRT runs and the 70.7% speech comprehension point for MAS runs. These values were recorded by hand in a study log for each participant and later compiled into an Excel spreadsheet containing the data for all participants.

The study was designed to test for MAS under eight different listening conditions. To do this we first needed the SRT for each corresponding listening condition. A comprehensive list of each run type with its listening condition along with what the run was testing for (SRT/MAS) is provided in Table 3.1 below:

Table 3.1 - All Run Types and the Metric Being Tested

Listening Condition	Environment	Masker Configuration	Metric Tested

Binaural Without Maskers	Simulated Anechoic Field	N/A	SRT
	Simulated Reverberant Field	N/A	SRT
Binaural With Maskers	Simulated Anechoic Field	Collocated at Target	SRT
		Symmetric	MAS
		Asymmetric	MAS
	Simulated Reverberant Field	Collocated at Target	SRT
		Symmetric	MAS
		Asymmetric	MAS
Monaural With Maskers	Simulated Anechoic Field	Collocated at Target	SRT
		Symmetric	MAS
		Asymmetric	MAS
	Simulated Reverberant Field	Collocated at Target	SRT
		Symmetric	MAS
		Asymmetric	MAS

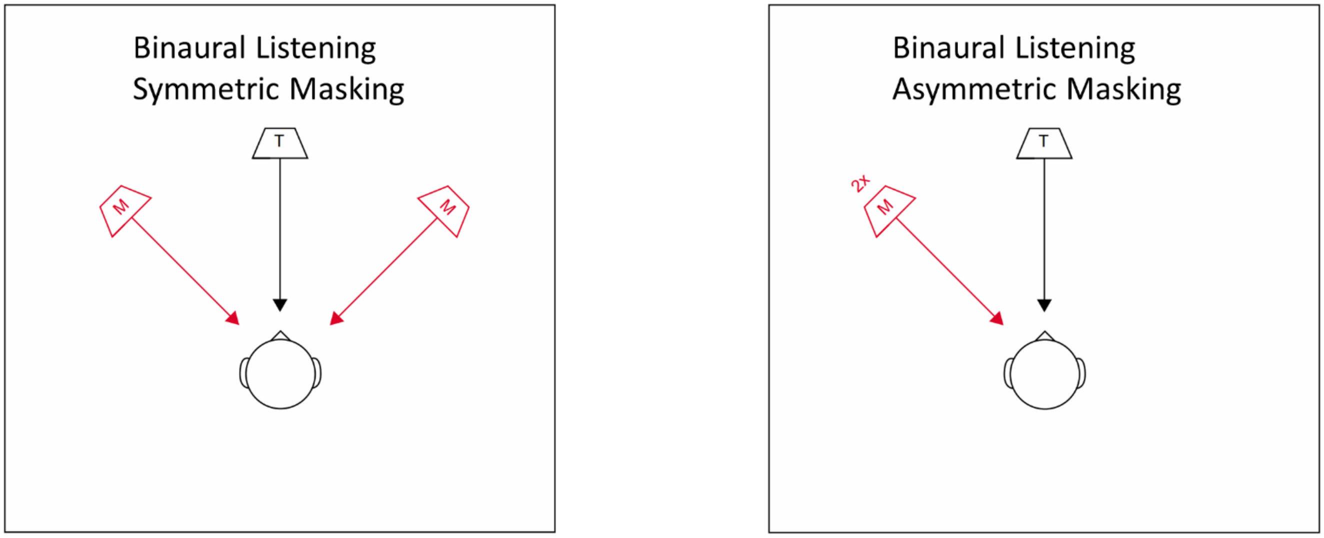
3.2.4.3 Listening Conditions

In the binaural listening conditions, audio signals were delivered to both the left and right ears. The monaural listening conditions used audio signals presented to only one ear - specifically the right ear due to the "right ear advantage" phenomenon which states listeners report stimuli more correctly from the right ear than the left (Tanaka et al., 2021).

Anechoic listening conditions provided an unaltered signal to the participant’s headphones without any reverberation, simulating an anechoic chamber. Reverberant listening conditions applied a reverberation time of 0.6 seconds to the audio signals, aiming to simulate a real-world classroom environment. The reverberation time of 0.6 seconds was chosen to adhere to the recommended reverberation time criteria for classrooms under 10,000 ft^3 as outlined by ANSI/ASA (ANSI/ASA S12.60-2010). The reverberant environment used in the study was created using a sample classroom that was modeled using ODEON and designed to have an average reverberation time of 0.6 seconds.

In symmetric MAS runs, the maskers started at -70 degrees and +70 degrees, with 0 degrees representing the direction directly in front of the participant at ear level, and +/-90 degrees indicating straight to the participant's right or left at ear level, respectively. The decision to start the maskers at +/-70 degrees was based on an observation made during the initial testing of adult participants. When the maskers were placed at +/- 90 degrees, many participants were completely unable to distinguish between the maskers and the target. This led to the modification of the initial position to avoid such issues. In asymmetric MAS runs, the maskers started collocating at -70 degrees. As the run progressed, both maskers would move together to each new position. See Figure 3.5 and Figure 3.6 for a visual representation of the various masking configurations for each of the listening conditions along with the auditory cues each configuration provides the subject.

Figure 3.5 - Masking Configuration for the Binaural Listening Conditions



Binaural Listening – Symmetric Masking

<u>Cues Accessible</u>

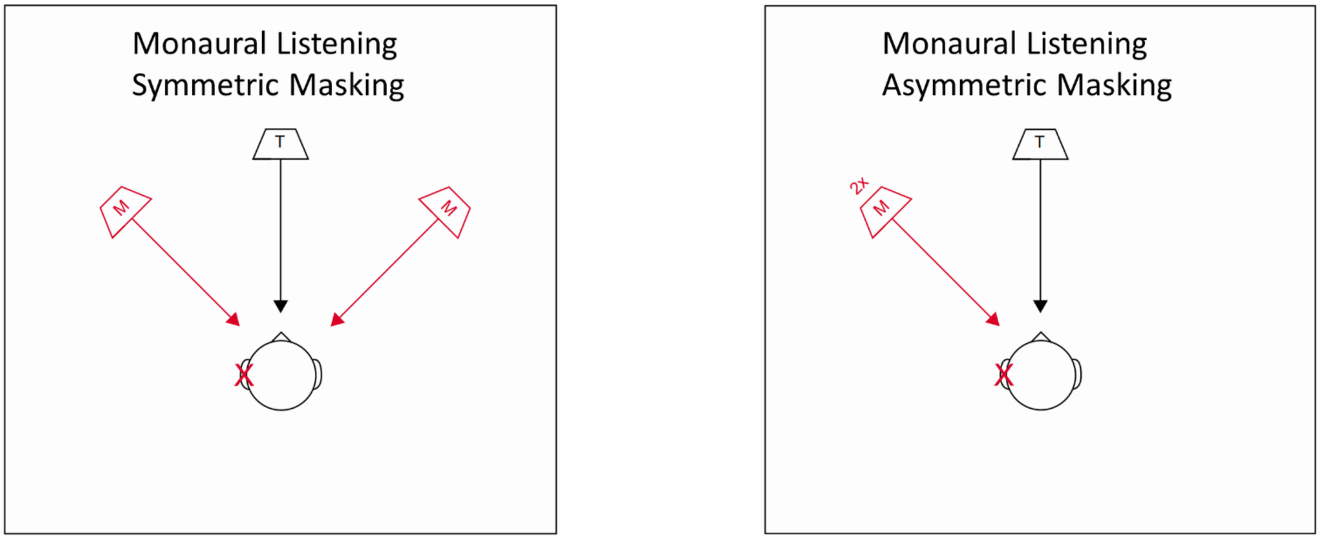
- Binaural
  - ITD
  - ILD
- Monaural
  - None

Binaural Listening – Asymmetric Masking

<u>Cues Accessible</u>

- Binaural
  - ITD
  - ILD
- Monaural
  - Head Shadow

Figure 3.6 - Masking Configuration for the Monaural Listening Conditions



Cues Accessible (Monaural Listening, Symmetric Masking)

- Binaural
  - None
- Monaural
  - Partial Head Shadow

Cues Accessible (Monaural Listening, Asymmetric Masking)

- Binaural
  - None
- Monaural
  - Head Shadow

3.2.4.4 Testing Order

In the quiet conditions, only the SRT was recorded since it was not possible to obtain a "quiet" MAS without any maskers present. Therefore, the SRT represented the level at which participants achieved 50% speech comprehension without any competing auditory influences.

During testing, the first two runs consisted of binaural anechoic quiet and binaural reverberant quiet listening conditions testing for SRT, the order of which was alternated between participants. In these quiet conditions, no maskers were present, and the participant only heard the target speaker. As there were no maskers in these listening conditions, there was also no testing for MAS. These runs served to establish a baseline SRT for each participant in both anechoic and reverberant listening conditions as well as to familiarize a participant to the testing procedure before complicating the test with the presence of maskers.

The remaining twelve runs were grouped into four sets with three runs each. The first run of the three tested SRT with maskers collocated at the target position. This was used to set a level for the target speaker in the subsequent two runs for each set. The next two runs tested MAS, one with symmetric masking and the other with asymmetric masking. Each of the four sets of runs was done with a different listening condition. As outlined above, these conditions were Binaural Listening in a Simulated Anechoic Field, Binaural Listening in a Simulated Reverberant Field, Monaural Listening in a Simulated Anechoic Field, and Monaural Listening in a Simulated Reverberant Field. The order in which participants received the four sets of three runs was determined using the Latin Square method to ensure random assignment and prevent any listening order biases. The order of symmetric and asymmetric masking within each set of three runs alternated to eliminate potential data errors related to the order of masking conditions. Combining the Latin Square method and alternating masking symmetry order yielded four possible testing orders before repeating the sequence for any participant. This meant participants 1 and 5 would have the same listening condition order and participants 2 and 6 would have the same listening condition order, etc.

To further prevent biases from testing methods, the sentences the participants listened to for each trial were started at different points in the sentence order. This means that, although participants 1 and 5 experienced the same listening condition order, the sentences for each run were not necessarily the same. This was done by the test administrator manually picking a random starting point in the database of sentences.

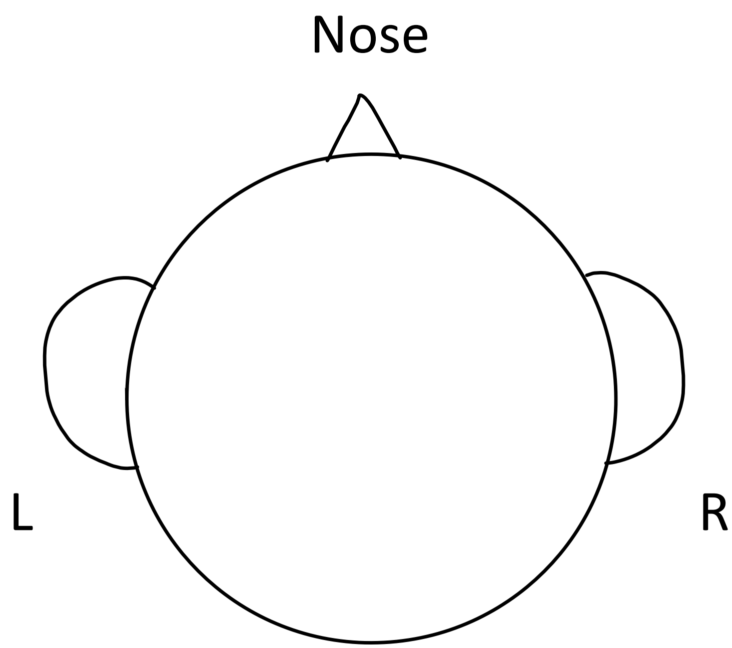
Additionally, participants required a different number of sentences to complete a run depending on their individual performance in the tasks. This means that even if two participants were given the same listening condition order and happened to be started at the same point in the database of sentences, they would quickly be receiving different sentences for each run due to previous runs using up a differing number of sentences.

3.2.5 Subjective Hearing Location for Children

Following each set of runs for each condition, the child participants were asked to mark on a sheet of paper where they perceived the target and maskers' voices to be located in relation to their head in the respective listening condition. They were then instructed to mark a "T" for the target and two "M"s for the maskers on a sheet of paper showing an aerial view of their head. The graphic representation is shown in Figure 3.7.

Figure 3.7 - Subjective Target/Masker Location Graphic

Subject ID:______________



Condition

isReverb:_________
Hdph Chann:_____

4 Results and Discussion

In this chapter, we present the findings of the study, which include a compilation of data and a detailed analysis of the results. Specifically, we examine the performance of participants in the study, as well as the statistical analysis and demographic considerations that were considered. Overall, this chapter provides a comprehensive overview of the data collected and the key insights gained from the analysis.

4.1 Demographics

The study's participant sample was divided into two age groups: children and adults. The child group consisted of 15 participants, ranging in age from 10 to 15 years old, with an average age of 12.4 years and a standard deviation of 1.5. Of the child participants, 13 were female and two were male. The adult group consisted of 23 participants, ranging in age from 19 to 29 years old, with an average age of 23.2 years and a standard deviation of 2.3. Of the adult participants, seven were female and 16 were male. A target sample size of at least 15 was set for each group, adults and children. This sample size was achieved in both groups and exceeded in the adult group. This sample size was determined based off other similar studies (Hawley & Litovsky, 2004; Peng et al., 2021).

4.2 MAS in Anechoic vs Reverberant Environments

To evaluate spatial hearing abilities, four different listening conditions were used in both reverberant and non-reverberant settings. The findings have been visually represented in three separate graphs: one for all participants (Figure 4.1), another for adult participants only (Figure 4.2), and a third for child participants only (Figure 4.3).

Figure 4.1 - MAS Performance Across Different Listening Conditions: All Participants

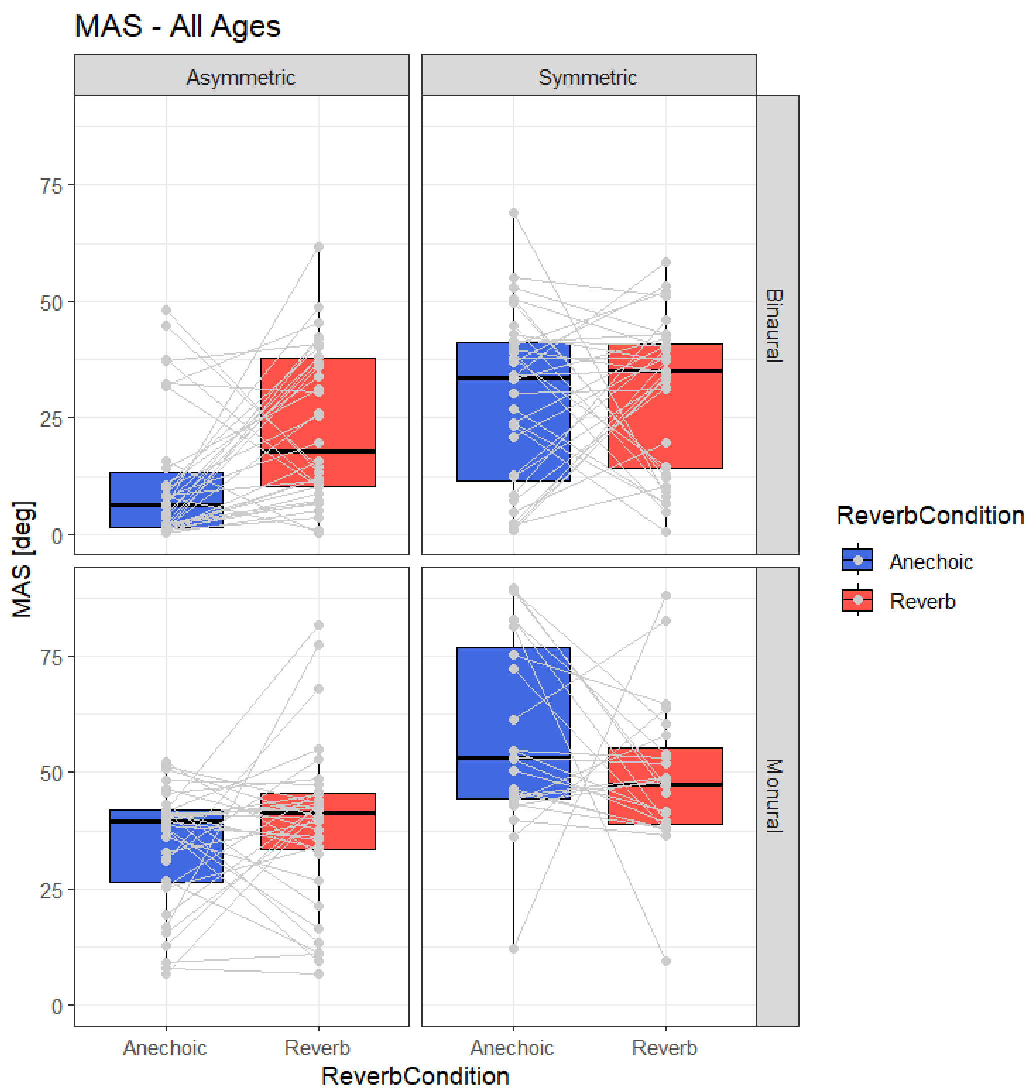


Figure 4.2 - MAS Performance Across Different Listening Conditions: Adults Only

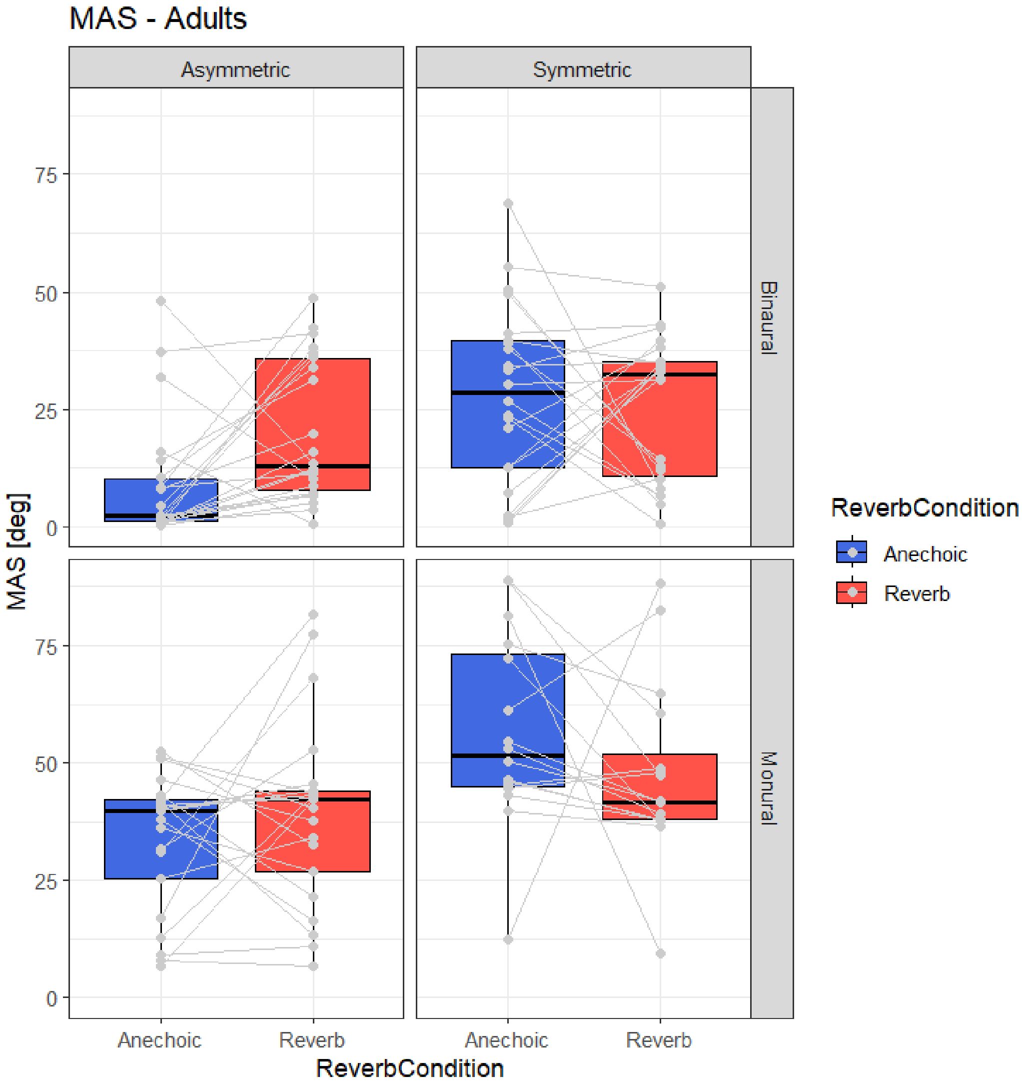
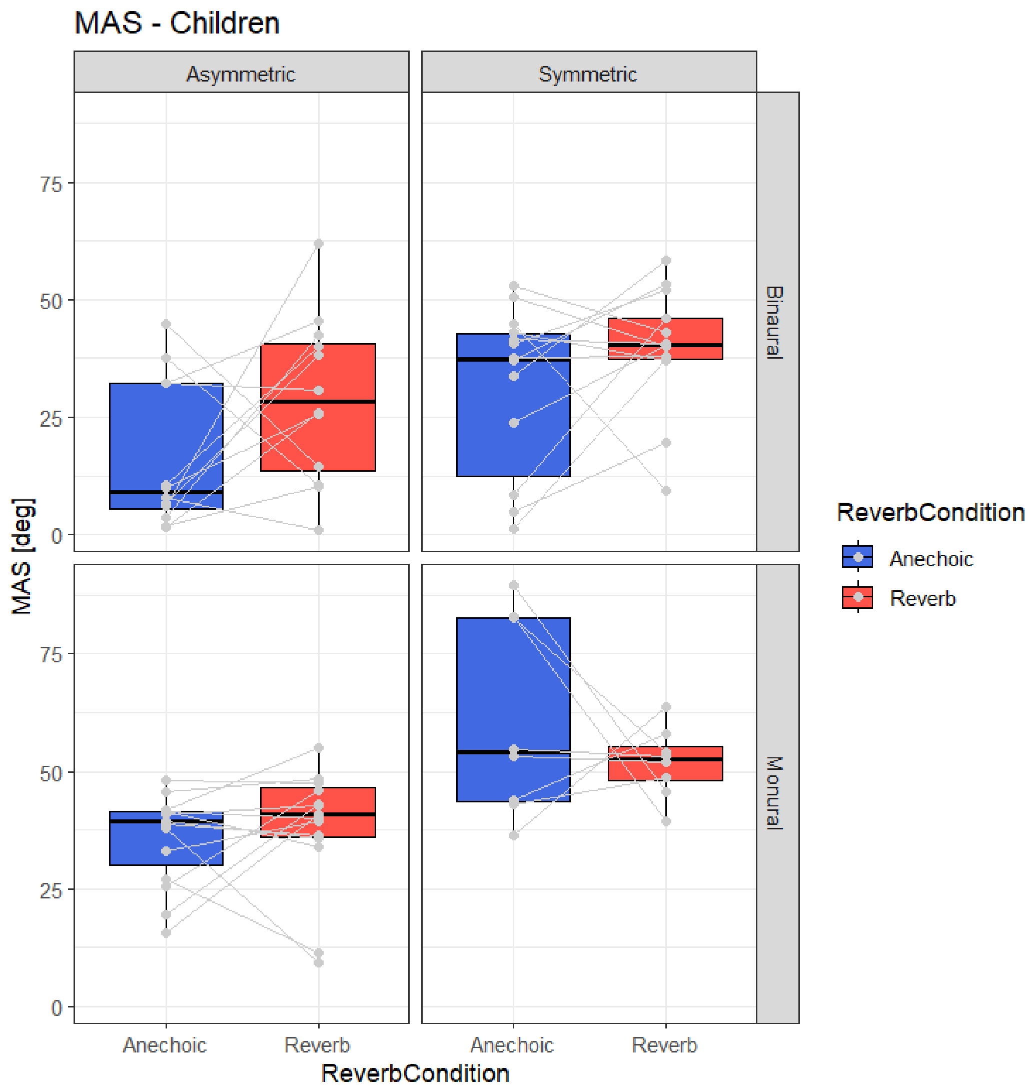


Figure 4.3 - MAS Performance Across Different Listening Conditions: Children Only



To generate these graphs and conduct the statistical analysis, data were trimmed so that only paired data would be included. So, for a participant's data to be included in a condition, a valid result was required from that participant for both the anechoic and reverberant versions of the listening condition. As with any study, there was invalid data and data loss. In our case, for a data point to be valid, we asserted its anechoic/reverberant counter parts had to be valid. This was done to better observe the impact reverberation has on the various listening conditions. It would not be beneficial to include data points that did not include both conditions for comparison. When data loss is reported in the following sections, it means a participant failed or gave an inconclusive result for one or both of the anechoic and reverberant conditions. A failed or inconclusive result could happen if the resulting calculated MAS was less than 0°, greater than 90°, or the test had an error. These things happened when the participant was reaching a point of aural fatigue, or simply had an inability to complete the task due to task difficulty. Of the total 152 MAS data pairs, 130 could be used for the analysis, resulting in ~86% data validity throughout the study.

4.2.1 Binaural Listening with Asymmetric Masking

This section presents the results and analysis of the binaural listening with asymmetric masking testing condition of the study, including performance outcomes, statistical analyses, and a demographically faceted analysis. In this condition, participants had access to ILD, ITD, and head shadow cues. With all three cues present, participants were able to complete tasks relatively easily, resulting in very little data loss. Specifically, 22 out of 23 adults completed this section successfully and 12 out of 15 children completed this section successfully.

The study found that when reverberation was present in the binaural listening with asymmetric masking condition, participants had a harder time separating the masking and target sounds. This resulted in a higher MAS in the reverberant condition than in the anechoic condition. The difference in MAS was statistically significant, $t(33) = -3.13$, $p =$

0.002. Specifically, the average MAS was 23 degrees with reverberation and 12 degrees without reverberation.

The statistical significance was present in the adult sample and not in the child sample with $t(21) = -2.64$, $p = 0.008$ for the adult sample and $t(11) = -1.70$, $p = 0.059$ for the child sample. The adult sample had an average MAS of 20 degrees in the reverberant condition and 9 degrees in the anechoic condition. Whereas the child sample had an average MAS of 29 degrees with reverberation and 16 degrees without reverberation. This leads us to believe the statistical significance in the overall test sample comes from the adult sample, though the child sample shows roughly the same change in MAS with the addition of reverberation (~5 degrees). The trend observed in the adult sample is still present in the child sample – just not to a statistically significant extent.

Notably, the average MAS achieved by the child sample was larger in both reverberant and anechoic conditions than the average MAS achieved by adults. This suggests that adults have better spatial hearing overall while listening binaurally with asymmetric masking. However, the impact of reverberation on the difference in MAS between the anechoic and reverberant conditions was similar in both child and adult samples. Therefore, while adults may have better spatial hearing abilities overall, the impact of reverberation on their performance appears to be similar in adults and children.

Taken together, these findings suggest that the addition of reverberation weakens one or more of the ITD, ILD or head shadow auditory cues used for spatial hearing in binaural listening with asymmetric masking. Additionally, the impact of reverberation on these cues is similar in both adult and child samples.

4.2.2 Monaural Listening with Asymmetric Masking

This section presents the results and analysis of the monaural listening with asymmetric masking testing condition of the study, including performance outcomes, statistical analyses, and a demographically faceted analysis. Due to the monaural nature of this task the participants did not have access to ILD or ITD cues and relied completely on the head shadow cue. Despite the limited access to auditory cues, participants were able to successfully complete this section. Specifically, 21 out of 23 adults successfully completed this section, while 15 out of 15 children successfully completed this section.

In the monaural listening with asymmetric masking condition, the total sample had an average MAS of 39 degrees with reverberation and 35 degrees without. The statistical analysis showed a non-significant difference, $t(35) = -1.29$, $p = 0.102$. Similarly, adults had an average MAS of 39 degrees with reverberation and 34 degrees without. Adults had a non-significant difference in this listening condition, $t(20) = -1.10$, $p = 0.142$. The child sample had an average MAS of 39 degrees with reverberation and 36 degrees without reverberation. The children had a non-significant difference in this listening condition as well, $t(14) = -0.66$, $p = 0.260$.

These results suggest that, unlike in the binaural listening with asymmetric masking condition, the presence of reverberation did not have a significant impact on the monaural spatial hearing abilities of participants with asymmetric masking. While the average MAS was still higher in the reverberant condition for all samples, like in the binaural listening with asymmetric masking, the difference was not statistically significant. It is also worth noting that the average MAS achieved by children was lower than that of adults in both the reverberant and anechoic conditions. This is the opposite of what was seen in the binaural listening with asymmetric masking condition, though with the differences being small, this could be due to sampling variability.

Overall, these findings suggest that while the presence of reverberation can negatively impact spatial hearing abilities in binaural listening with asymmetric masking, it may not have a significant impact in monaural listening with asymmetric masking. Interestingly, the MAS was lower without reverberation than with it in all cases in this condition, but the statistics do not show that pattern to be pronounced enough to be significant in the monaural listening with asymmetric masking condition. As the head shadow cue is the only cue present in the monaural condition, the data suggest that the effect reverberation plays on this cue may be small. Alternatively, the lack of a statistically significant change could be an artifact of this condition having so little access to auditory cues. Increasing the difficulty of an already difficult task could diminish the magnitude of the overall change we would expect to see.

4.2.3 Binaural Listening with Symmetric Masking

This section presents the results and analysis of the binaural listening with symmetric masking testing condition of the study, including performance outcomes, statistical analyses, and a demographically faceted analysis. As this task used binaural listening, participants had access to both binaural auditory cues, ITD and ILD. The symmetry of the maskers negated any access the participant would have to head shadow. In this portion of the study 22 out of 23 adults and 14 out of 15 children were able to successfully complete the test without data loss.

The results of the study for binaural listening with symmetric masking showed that the total sample had an average MAS of 31 degrees with reverberation and 29 degrees without. In this listening condition there was no significant difference when comparing MAS in reverberant and anechoic fields, $t(35) = -0.51$, $p = 0.611$. The results for the adult sample showed that they had an average MAS of 25 degrees with reverberation and 28 degrees without, $t(21) = 0.45$, $p = 0.659$. There was no significant difference in the MAS between the two conditions for the adult sample as well. Interestingly, the child sample had an average MAS of 40 degrees with reverberation and 31 degrees without, $t(13) = -1.67$, $p = 0.120$. This indicates that there was a much more pronounced difference with the child sample than the adult sample. This is the opposite of what was found in the other listening conditions.

Overall, the findings suggest that reverberation in the binaural listening with symmetric masking does not significantly affect the MAS and thus the ITD and ILD auditory cues present are not significantly affected by reverberation. Though the results are not statistically significant, it is interesting that the anechoic condition had a higher MAS than the reverberant condition. It is especially curious as this was the case in both the adult and child samples. The child sample had a higher MAS in both conditions than the adult sample did as well. This is consistent with what has been observed so far.

4.2.4 Monaural Listening with Symmetric Masking

This section presents the results and analysis of the monaural listening with symmetric masking testing condition of the study, including performance outcomes, statistical analyses, and a demographically faceted analysis. Monaural listening does not give the participant access to the binaural cues of ITL or ILD. The symmetric nature of the maskers would typically also block any access to head shadow cue a participant would have access to; however, as this is a monaural listening condition, the symmetry of the maskers would likely not produce that same effect. In a monaural listening condition, the symmetric maskers would give a weighted effect to the head shadow. In this portion of the study, we considered the auditory cue to which participants had access to be only a "partial head shadow." As the availability of auditory cues participants had access to in this portion of the study was very small, this proved to be a challenging task to complete, which resulted in more data loss for this section than other sections. In this portion of the study 15 out of 23 adults and 8 out of 15 children were able to complete the study without data loss.

The results of the study for monaural listening with symmetric masking showed that the total sample had an average MAS of 49 degrees with reverberation and 58 degrees without. There was no significant difference in this listening condition when comparing MAS in reverberant and anechoic fields, $t(23) = 1.50$, $p = 0.927$. The results for the adult sample showed an average MAS of 48 degrees with reverberation and 56 degrees without, $t(15) = 1.15$, $p = 0.867$. There was no significant difference in the MAS between the two conditions for the adult sample as well. Similarly, the child sample had an average MAS of 52 degrees with reverberation and 61 degrees without, $t(7) = 0.94$, $p = 0.811$. This indicates that there was no significant difference in the MAS between the two conditions for children.

Overall, the findings suggest that monaural listening with symmetric masking does not significantly affect the MAS. It is not unexpected; there was no significant difference in this listening condition as the participants had a very weak auditory cue to begin with. If reverberation lessens the participants access to a head shadow, it would be difficult to see the change when they only had access to a partial head shadow before reverberation was added.

4.3 SRT Values Achieved Across Listening Conditions

Presented in this section are the average SRT values and their standard deviations from each of the different types of runs study subjects participated in. As SRT values were not directly relevant to the research questions being asked but rather a steppingstone to obtain the MAS values, SRT values were simply averaged for reporting rather than statistically analyzed. The data trimming procedure used for the MAS data analysis was applied to the SRT analysis. This procedure entailed omitting the corresponding reverberant or anechoic run's data if a participant failed to provide a valid data point for either run in any listening condition. Consequently, the SRT used in these omitted runs did not contribute to the values presented here.

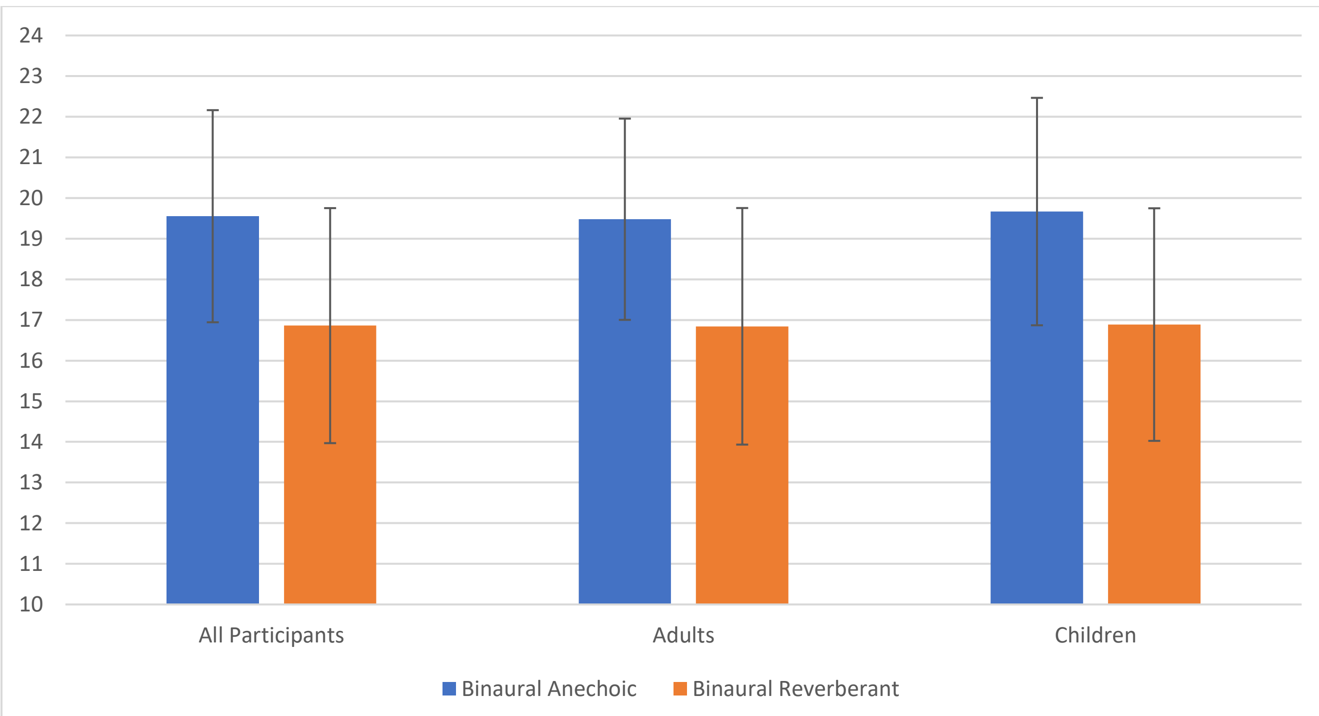
4.3.1 SRT in the Quiet Conditions

The initial tests conducted were always quiet condition SRT runs, which determined the SRT values a participant could achieve when listening to target speech without the presence of maskers. The average SRT values and standard deviations for different participant groups are presented in Table 4.1 and Figure 4.4.

Table 4.1 - Average SRT Values in the Quiet Conditions

Quiet SRT (dB)	All Participants		Adults		Children	
	Average	Std Dev	Average	Std Dev	Average	Std Dev
Binaural Listening in a Simulated Anechoic Field	19.6	2.6	19.5	2.5	19.7	2.8
Binaural Listening in a Simulated Reverberant Field	16.9	2.9	16.8	2.9	16.9	2.9

Figure 4.4 - Average SRT Values in the Quiet Conditions



As depicted in Table 4.1 and Figure 4.4 above, the reverberant SRT values were approximately 3 dB lower than the anechoic SRT values. Given that reverberant conditions introduced additional sound to the participant, simulating real-life reflected sound, it logically follows that a participant could achieve 50% speech comprehension at a lower level. The difference of ~3 dB is likely attributed to the chosen reverberation time, 0.6s, for this study. Future research exploring how this number varies with different reverberation times could yield interesting insights.

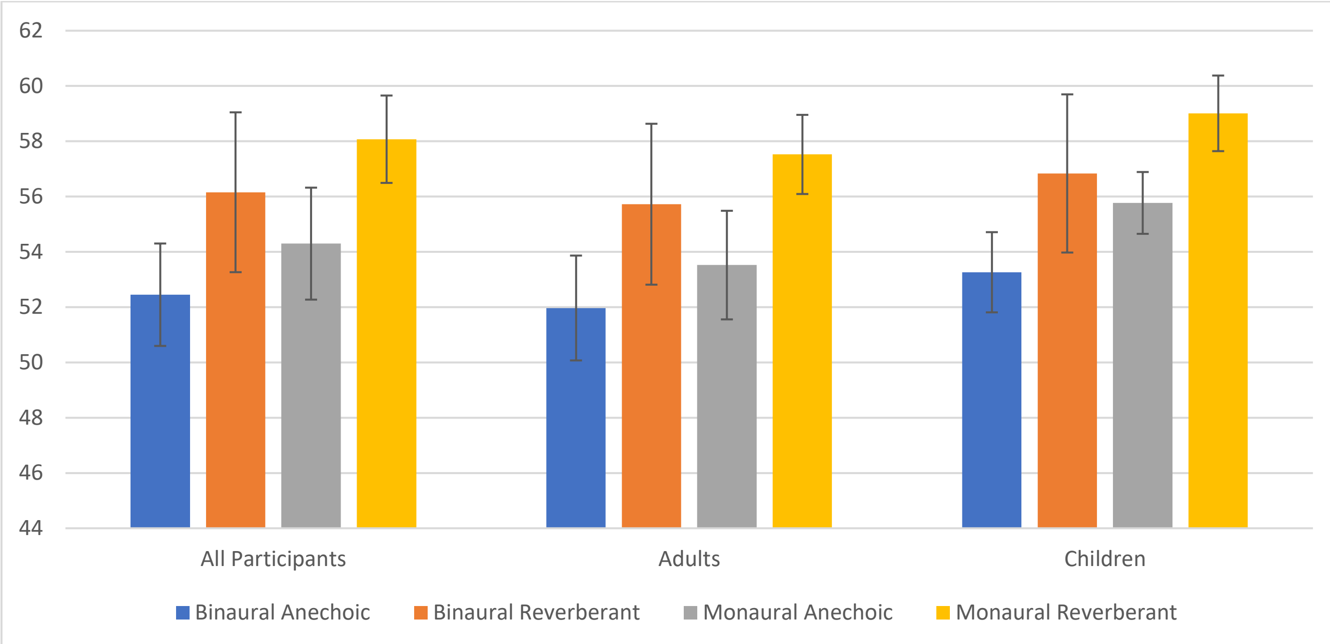
4.4 SRT in the Masked Conditions

Table 4.2 and Figure 4.5 below presents the average SRT values and standard deviations for different participant groups in conditions with collocated masking.

Table 4.2 - Average SRT Values in the Asymmetric Masking Conditions

Masked SRT (dB)	All Participants		Adults		Children	
	Average	Std Dev	Average	Std Dev	Average	Std Dev
Binaural Listening in a Simulated Anechoic Field	52.4	1.9	52.0	1.9	53.3	1.5
Binaural Listening in a Simulated Reverberant Field	56.2	1.6	55.7	1.6	56.8	1.4
Monaural Listening in a Simulated Anechoic Field	54.3	2.0	53.5	2.0	55.8	1.1
Monaural Listening in a Simulated Reverberant Field	58.1	1.6	57.5	1.4	59.0	1.4

Figure 4.5 - Average SRT Values in the Asymmetric Masking Conditions



As can be seen in

Table 4.2 and Figure 4.5 above, the reverberant SRT values were consistently higher than their anechoic counterparts, with a difference of ~2-4 dB. Interestingly, this is the inverse of what was observed in the quiet conditions displayed in Table 4.1 and Figure 4.4 . This suggests that the introduction of competing babble transforms the reverberation from a benefit to a hindrance when reducing the target volume. Intuitively, this makes sense as reverberant maskers could more easily obscure the target voice which would make it harder for a participant to discern what the target is saying.

In

Table 4.2 and Figure 4.5 we can also see that child participants had an average SRT larger than that of the adults by ~1-2 dB across all masked listening conditions. This could be due to immaturity in speech-in-noise listening skills among children. It could also be an artifact of the test itself as the smallest step size in change of the target speech level was 2 dB, on the order of the group difference seen here. Notably, in the quiet conditions, this effect was considerably smaller, <1 dB, suggesting that the presence of maskers may have amplified the difference in SRT values between adults and children.

5 Conclusion

This thesis aimed to explore the complex and important phenomenon of SRM during typical development among children and adults with normal hearing. The research focused on the role of auditory spatial cues under reverberation on SRM, through the application of the metric of MAS.

The study's findings have shed new light on the role of various factors in SRM. Firstly, the results showed that the effects of reverberation on SRM depend on the specific listening condition and the cues available to the listeners. Participants had more difficulty in reverberant conditions compared to anechoic conditions, resulting in larger measured MAS angles. This was particularly true in binaural listening scenarios with asymmetric masking. However, the presence of reverberation did not significantly affect performance in monaural listening scenarios with asymmetric masking.

The research also found that adults, ages 19 to 29, generally outperformed children, ages 10 to 15, in SRM tasks, although the differences were small. This suggests that spatial hearing abilities continue to develop beyond the age range 10 to 15 years studied in this work, reinforcing the need for further research into how these abilities evolve for individuals with normal hearing across all stages of aural development.

The application of MAS as a metric proved effective for quantifying SRM and offers a promising tool for future research in this area. By measuring the smallest angular separation between target and masking speech that yields a significant improvement in speech intelligibility, the MAS metric provides an insightful view of SRM that can help refine our understanding of this complex phenomenon.

Understanding the impact of reverberation on SRM and auditory cues has direct implications for the development of assistive listening devices, such as hearing aids and cochlear implants. As the development of these devices often rely on algorithms to enhance speech intelligibility in noisy environments, SRM research can inform the outcome of these algorithms by taking into consideration the effects of reverberation. The use of MAS as a metric provides a beneficial tool for researchers to investigate SRM and its underlying factors in various listening conditions A deeper understanding of SRM can help engineers and designers create more effective technologies that increase access to spatial auditory cues improving spatial hearing, particularly in challenging listening conditions. This is especially significant for individuals with hearing impairments and those who rely heavily on assistive listening devices.

Future research is recommended to further delve into the roles of different auditory cues in SRM and to expand on the MAS metric. Additionally, more detailed investigation into the effects of reverberation and masking conditions on SRM across different age groups could provide more insights into the development and function of spatial hearing abilities.

In conclusion, this research represents a significant step forward in our understanding of SRM and the complex interplay of factors that influence it. As we continue to explore this field, we move closer to improving communication capabilities and the quality of life for individuals living in a noisy world.

References

American National Standards Institute/Acoustical Society of America. (2010). Acoustical Performance Criteria, Design Requirements, and Guidelines for Schools (S12.60-2010/Part 1)

Baltzell, L. S., Swaminathan, J., Cho, A. Y., Lavandier, M., & Best, V. (2020). Binaural sensitivity and release from speech-on-speech masking in listeners with and without hearing loss. *The Journal of the Acoustical Society of America*, *147*(3), 1546–1561. https://doi.org/10.1121/10.0000812

Best, V., Marrone, N., Mason, C. R., & Kidd, G. (2012). The influence of non-spatial factors on measures of spatial release from masking. *The Journal of the Acoustical Society of America*, *131*(4), 3103–3110. https://doi.org/10.1121/1.3693656

Best, V., Mason, C. R., & Kidd, G. (2011). Spatial release from masking in normally hearing and hearing-impaired listeners as a function of the temporal overlap of competing talkers. *The Journal of the Acoustical Society of America*, *129*(3), 1616–1625. https://doi.org/10.1121/1.3533733

Beutelmann, R., & Brand, T. (2006). Prediction of speech intelligibility in spatial noise and reverberation for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *120*(1), 331–342. https://doi.org/10.1121/1.2202888

Biberger, T., & Ewert, S. D. (2019). The effect of room acoustical parameters on speech reception thresholds and spatial release from masking. *The Journal of the Acoustical Society of America*, *146*(4), 2188–2200. https://doi.org/10.1121/1.5126694

Bronkhorst, A. W. (2015). The cocktail-party problem revisited: Early processing and selection of multi-talker speech. *Attention, Perception, & Psychophysics*, *77*(5), 1465–1487. https://doi.org/10.3758/s13414-015-0882-9

Bronkhorst, A. W., & Plomp, R. (n.d.). *The effect of head-induced interaural time and level differences on speech intelligibility in noise*.

Brown, A. D. (n.d.). The Precedence Effect in Sound Localization. *Journal of the Association for Research in Otolaryngology*.

Chandler, D. W., & Grantham, D. W. (1992). Minimum audible movement angle in the horizontal plane as a function of stimulus frequency and bandwidth, source azimuth, and velocity. *The Journal of the Acoustical Society of America*, *91*(3), 1624–1636. https://doi.org/10.1121/1.402443

Ching, T. Y. C., van Wanrooy, E., Dillon, H., & Carter, L. (2011). Spatial release from masking in normal-hearing children and children who use hearing aids. *The Journal of the Acoustical Society of America*, *129*(1), 368–375. https://doi.org/10.1121/1.3523295

Culling, J. F., Hodder, K. I., & Toh, C. Y. (2003). Effects of reverberation on perceptual segregation of competing voices. *The Journal of the Acoustical Society of America*, *114*(5), 2871. https://doi.org/10.1121/1.1616922

Dawson, P. W., Hersbach, A. A., Swanson, B. A. (2013). An adaptive Australian Sentence Test in Noise (AuSTIN). Ear Hear

Dietz, M., Ewert, S. D., & Hohmann, V. (2012). Lateralization based on interaural differences in the second-order amplitude modulator. *The Journal of the Acoustical Society of America*, *131*(1), 398–408. https://doi.org/10.1121/1.3662078

Dieudonné, B., & Francart, T. (2019). Redundant Information Is Sometimes More Beneficial Than Spatial Information to Understand Speech in Noise. *Ear & Hearing*, *40*(3), 545–554. https://doi.org/10.1097/AUD.0000000000000660

Edmonds, B. A., & Culling, J. F. (2006). The spatial unmasking of speech: Evidence for better-ear listening. *The Journal of the Acoustical Society of America*, *120*(3), 1539–1545. https://doi.org/10.1121/1.2228573

Ellen Peng, Z., & Litovsky, R. Y. (2021). The Role of Interaural Differences, Head Shadow, and Binaural Redundancy in Binaural Intelligibility Benefits Among School-Aged Children. *Trends in Hearing*, *25*, 233121652110453. https://doi.org/10.1177/23312165211045313

Freyman, R. L., Helfer, K. S., McCall, D. D., & Clifton, R. K. (1999). The role of perceived spatial separation in the unmasking of speech. *The Journal of the Acoustical Society of America*, *106*(6), 3578–3588. https://doi.org/10.1121/1.428211

Gallun, F. J., Diedesch, A. C., Kampel, S. D., & Jakien, K. M. (2013). Independent impacts of age and hearing loss on spatial release in a complex auditory environment. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00252

Garadat, S. N., Litovsky, R. Y., Yu, G., & Zeng, F.-G. (2010). Effects of simulated spectral holes on speech intelligibility and spatial release from masking under binaural and monaural listening. *The Journal of the Acoustical Society of America*, *127*(2), 977–989. https://doi.org/10.1121/1.3273897

Goupell, M. J., Gaskins, C. R., Shader, M. J., Walter, E. P., Anderson, S., & Gordon-Salant, S. (2017). Age-Related Differences in the Processing of Temporal Envelope and Spectral

Cues in a Speech Segment. *Ear & Hearing*, *38*(6), e335–e342. https://doi.org/10.1097/AUD.0000000000000447

Goupell, M. J., Yu, G., & Litovsky, R. Y. (2012). The effect of an additional reflection in a precedence effect experiment. *The Journal of the Acoustical Society of America*, *131*(4), 2958–2967. https://doi.org/10.1121/1.3689849

Hawley, M. L., & Litovsky, R. Y. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferera). *J. Acoust. Soc. Am.*, *115*(2).

Johnstone, P. M., & Litovsky, R. Y. (2006). Effect of masker type and age on speech intelligibility and spatial release from masking in children and adults. *The Journal of the Acoustical Society of America*, *120*(4), 2177–2189. https://doi.org/10.1121/1.2225416

Lavandier, M., & Culling, J. F. (2010). Prediction of binaural speech intelligibility against noise in rooms. *The Journal of the Acoustical Society of America*, *127*(1), 387–399. https://doi.org/10.1121/1.3268612

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*(8), 1279–1292. https://doi.org/10.3758/BF03194543

Litovsky, R. Y. (2005). Speech intelligibility and spatial release from masking in young children. *The Journal of the Acoustical Society of America*, *117*(5), 3091–3099. https://doi.org/10.1121/1.1873913

Litovsky, R. Y. (2012). Spatial Release from Masking. *Acoustics Today*, *8*(2), 18. https://doi.org/10.1121/1.4729575

Litovsky, R. Y., Johnstone, P. M., & Godar, S. P. (2006). Benefits of bilateral cochlear implants and/or hearing aids in children: Beneficios de los implantes cocleares bilaterales y/o auxiliares auditivos en niños. *International Journal of Audiology*, *45*(sup1), 78–91. https://doi.org/10.1080/14992020600782956

Litovsky, R. Y., Parkinson, A., & Arcaroli, J. (2009). Spatial Hearing and Speech Intelligibility in Bilateral Cochlear Implant Users. *Ear & Hearing*, *30*(4), 419–431. https://doi.org/10.1097/AUD.0b013e3181a165be

Marrone, N., Mason, C. R., & Kidd, G. (2008). Tuning in the spatial dimension: Evidence from a masked speech identification task. *The Journal of the Acoustical Society of America*, *124*(2), 1146–1158. https://doi.org/10.1121/1.2945710

Middlebrooks, J. C., & Green, D. M. (1991). Sound Localization by Human Listeners. *Annual Review of Psychology*, *42*(1), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031

Misurelli, S. M., & Litovsky, R. Y. (2012). Spatial release from masking in children with normal hearing and with bilateral cochlear implants: Effect of interferer asymmetry. *The Journal of the Acoustical Society of America*, *132*(1), 380–391. https://doi.org/10.1121/1.4725760

Musicant, A. D., & Butler, R. A. (1984). The influence of pinnae-based spectral cues on sound localization. *The Journal of the Acoustical Society of America*, *75*(4), 1195–1200. https://doi.org/10.1121/1.390770

Peng, Z. E., & Litovsky, R. Y. (2022). Novel Approaches to Measure Spatial Release From Masking in Children With Bilateral Cochlear Implants. *Ear & Hearing*, *43*(1), 101–114. https://doi.org/10.1097/AUD.0000000000001080

Peng, Z. E., Pausch, F., & Fels, J. (2021). Spatial release from masking in reverberation for school-age children. *The Journal of the Acoustical Society of America*, *150*(5), 3263–3274. https://doi.org/10.1121/10.0006752

Plomp, R. (1976). "Binaural and monaural speech-intelligibility of connected discourse in reverberation as a function of azimuth of a single competing sound source (speech or noise)," Acustica 34, 201–211.

Rakerd, B., & Hartmann, W. M. (1986). Localization of sound in rooms, III: Onset and duration effects. *The Journal of the Acoustical Society of America*, *80*(6), 1695–1706. https://doi.org/10.1121/1.394282

Rakerd, B., & Hartmann, W. M. (2010). Localization of sound in rooms. V. Binaural coherence and human sensitivity to interaural time differences in noise. *The Journal of the Acoustical Society of America*, *128*(5), 3052–3063. https://doi.org/10.1121/1.3493447

Rennies, J., Brand, T., & Kollmeier, B. (2011). Prediction of the influence of reverberation on binaural speech intelligibility in noise and in quiet. *The Journal of the Acoustical Society of America*, *130*(5), 2999–3012. https://doi.org/10.1121/1.3641368

Schlauch, R. S., & Rose, R. M. (1990). Two-, three-, and four-interval forced-choice staircase procedures: Estimator bias and efficiency. *The Journal of the Acoustical Society of America*, *88*(2), 732–740. https://doi.org/10.1121/1.399776

Srinivasan, N. K., Stansell, M., & Gallun, F. J. (2017). The role of early and late reflections on spatial release from masking: Effects of age and hearing lossa). *The Journal of the Acoustical Society of America*, *141*(3), EL185–EL191. https://doi.org/10.1121/1.4973837

Stecker, G. C., & Bibee, J. M. (2014). Nonuniform temporal weighting of interaural time differences in 500 Hz tones. *The Journal of the Acoustical Society of America*, *135*(6), 3541–3547. https://doi.org/10.1121/1.4876179

Tanaka, K., Ross, B., Kuriki, S., Harashima, T., Obuchi, C., & Okamoto, H. (2021). Neurophysiological Evaluation of Right-Ear Advantage During Dichotic Listening. *Frontiers in Psychology*, *12*, 696263. https://doi.org/10.3389/fpsyg.2021.696263

TIME USA (n.d.). TIME for Kids. Retrieved November 6, 2020, from https://www.timeforkids.com/.

Wendt, T., Ewert, S. D., and van de Par, S. (2014). "A computationally efficient and perceptually-plausible algorithm for binaural room impulse response simulation," J. Audio Eng. Soc. 62, 748–766

Wightman, F. L., & Kistler, D. J. (1997). Monaural sound localization revisited. *The Journal of the Acoustical Society of America*, *101*(2), 1050–1063. https://doi.org/10.1121/1.418029

Yang, W Bradley, J. S. (2009). Effects of room acoustics on the intelligibility of speech in classrooms for young children. The Journal of the Acoustical Society of America, 125(2), 922-933. https://doi.org/10.1121/1.3058528

Yost, W. A. (1997). The cocktail party problem: Forty years later. In *Binaural and spatial hearing in real and virtual environments.* (pp. 329–347). Lawrence Erlbaum Associates, Inc.