8-2018

The effects of monaural and binaural cues on perceived reverberation by normal hearing and hearing-impaired listeners.

Gregory Matthew Ellis
University of Louisville

THE EFFECTS OF MONAURAL AND BINAURAL CUES ON PERCEIVED REVERBERATION BY NORMAL HEARING AND HEARING-IMPAIRED LISTENERS

By

Gregory Matthew Ellis
B.A., University of Wisconsin—Eau Claire, 2012
M.A., University of Minnesota—Twin Cities, 2014
M.S., University of Louisville, 2016

A Dissertation
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy
in Experimental Psychology

Department of Psychological and Brain Sciences
University of Louisville
Louisville, Kentucky

August 2018

THE EFFECTS OF MONAURAL AND BINAURAL CUES ON PERCEIVED REVERBERATION BY NORMAL HEARING AND HEARING-IMPAIRED LISTENERS

By

Gregory Matthew Ellis
B.A., University of Wisconsin—Eau Claire, 2012
M.A., University of Minnesota—Twin Cities, 2014
M.S., University of Louisville, 2016

A Dissertation Approved on

June 8, 2018

by the following Dissertation Committee:

____
Dissertation Director
Dr. Pavel Zahorik

____
Dr. Paul DeMarco

____
Dr. Maria Kondaurova

____
Dr. Sharon Miller

____
Dr. Christian Stilp

DEDICATION

This dissertation is dedicated to my late grandfather

Major Henry “Hank” Bartol, Ed.D.

who taught me from an early age that science is fun.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Pavel Zahorik for his bottomless patience and countless hours of mentorship. I would also like to thank the other committee members, Dr. Paul DeMarco, Dr. Maria Kondaurova, Dr. Sharon Miller, and Dr. Christian Stilp for their input and advice in developing this project and in teaching me valuable lessons over the last four years. I would also like to thank my mom, Dr. Laurie Ellis, my dad, Matthew Ellis, and my brother, Tim Ellis, for helping me stay the course and keep on track. At the most difficult points of graduate school they were all there to help keep me focused on my goal. I would also like to thank Dr. Eugene Brandewie, Dr. Paul Anderson, Dr. Chhayakant Patro, and Dr. James Shehorn for being good role models, lab mates, and friends. I would like to thank all of my other friends who supported me in big and small ways over the last four years. There are too many of them to list. I would like to thank my cat Maddie for being with me when I needed her the most. Finally, I would like to thank my moka pot and French press for endless supplies of great coffee, without which this document would never have been completed.

ABSTRACT

THE EFFECTS OF MONAURAL AND BINAURAL CUES ON PERCEIVED REVERBERATION BY NORMAL HEARING AND HEARING-IMPAIRED LISTENERS

Gregory M. Ellis

June 8th, 2018

This dissertation is a quantitative and qualitative examination of how young normal hearing and young hearing-impaired listeners perceive reverberation. A primary complaint among hearing-impaired listeners is difficulty understanding speech in noisy or reverberant environments. This work was motivated by a desire to better understand reverberation perception and processing so that this knowledge might be used to improve outcomes for hearing-impaired listeners in these environments.

This dissertation is written in six chapters. Chapter One is an introduction to the field and a review of the relevant literature. Chapter Two describes a motivating experiment from laboratory work completed before the dissertation. This experiment asked human subjects to rate the amount of reverberation they perceived in a sound relative to another sound. This experiment showed a significant effect of listening condition on how listeners made their judgments. Chapter Three follows up on this experiment, seeking a better understanding of how listeners perform the task in Chapter Two. Chapter Three shows that listeners can use limited information to make their judgments. Chapter Four compares reverberation perception in normal hearing and hearing-impaired listeners and examines the effect of speech intelligibility on

reverberation perception. This experiment finds no significant differences between cues used by normal hearing and hearing-impaired listeners when judging perceptual aspects of reverberation. Chapter Five describes and uses a quantitative model to examine the results of Chapters Two and Four. Chapter Six summarizes the data presented in the dissertation and discusses potential implications and future directions.

This work finds that the perceived amount of reverberation relies primarily on two factors: 1) the listening condition (i.e., binaural, monaural, or a listening condition in which reverberation is present only in one ear) and 2) the sum of reverberant energy present at the two ears. Listeners do not need the reverberant tail to estimate perceived amount of reverberation, meaning that listeners are able to extract information about reverberation from the ongoing signal. The precise mechanism underlying this process is not explicitly found in this work; however, a potential framework is presented in Chapter Six.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Introduction

From the barely-noticeable presence in a living room to the 9-second ring of Saint Paul's Cathedral, people—and their auditory systems—are constantly surrounded by reverberation. If we are to fully understand how the auditory system functions, researchers must consider how it is affected by reverberant environments. In general, scientific study of this has either studied the perception of reverberation directly, or the effect of reverberation on other aspects of auditory perception or listening ability.

The latter effect has been well established. Detrimental effects of reverberation on speech understanding and sound localization have been identified and studied over at least the past 50 years. These results come from studies that tightly control the physical stimuli, and therefore establish a causal relationship between the physical manipulations and the resultant perception. These studies do not, however, shed light on how the reverberation itself is perceived by the listeners.

The perception of reverberation itself is studied primarily by the concert hall acoustics literature. Studies in this field generally look at how expert listeners perceive sounds in famous concert halls and use this information to inform design of future halls. Preference ratings and specialized language are used in these studies to describe the perception of reverberation. These studies tend to be conducted in the halls they wish to study and therefore cannot tightly control the physical stimuli presented to the listeners.

Because the physical stimuli are not systematically manipulated, these studies cannot make conclusions about the causal relationship between acoustics and perception.

Stark few studies bridge the gap between these bodies of research by systematically varying the acoustics of reverberant stimuli to better understand the perceptual result. It is the goal of this dissertation to draw causal conclusions about perception through the use of tightly controlled reverberant stimuli.

Literature review

As with many areas of research, reverberation perception has an extensive list of jargon that is required to fully understand the research. This chapter aims to define these important terms. Definitions of some terms can be found in the glossary on page 112 (Appendix A). This chapter will first define reverberation as a phenomenon. Then it will examine some of the previous research on reverberation perception as it falls into the two major categories outlined above: 1) the effect reverberation has on speech perception and 2) how reverberation is perceived. Finally, the methods used in this dissertation will be described in detail. This includes virtual auditory space techniques and multidimensional scaling. These techniques play a key role in the work to be presented here.

Reverberation

When a sound is generated in a room, some of the energy propagates directly from the source (e.g., an instrument, a speaker, etc.) to the receiver (e.g., a microphone, a listener, etc.). This energy is referred to as the direct energy. Some of the energy produced by the source fans out and reflects off the walls, floor, and ceiling of the room before reaching the receiver. This energy comprises the early reflections. The remainder of the energy reflects multiple times off surfaces before reaching the receiver and produces spatially diffuse energy known as reverberation.

It is important to be able to quantify sounds in rooms if we are to study them scientifically. Fortunately, there are a number of well-defined measures that are accepted by acousticians and engineers (ISO-3382, 1997). These measures include reverberation time (T_{60}: the amount of time measured in seconds for reverberant energy to decrease in level by 60 dB), direct-to-reverberant energy ratio (DR: the level of the direct path energy divided by the level of the reverberant energy measured in dB), interaural cross-correlation (IACC: the maximum correlation (r) between the signals measured at the ears within ±2.5 ms lag time), clarity index (C_t: the ratio of energy before t to the energy after t, where t some number of milliseconds), and center time (T_c: the center point of the squared impulse response, measured in milliseconds).

The effect of reverberation and listening condition on speech understanding

Reverberation has long been known to cause problems for many functions of the auditory system. In 1950, Koenig noted in a letter to the editor of a journal that having two ears with different signals at the ears seems to help reduce—squelch—the impact of reverberation on perception (Koenig, 1950). Koenig's paper described a simple experimental set up. Two microphones were set somewhere in a moderately reverberant room. Leads from the microphones were fed into an adjacent sound-treated room, connected to a pair of headphones with a switch in between. There was a device in the sound-treated room that could change the way the leads were routed from the microphones to the headphones of a listener. The listener was able to freely switch between one of two listening modes: a dichotic mode, wherein the left lead was fed to the left earphone and the right lead was fed to the right earphone; or a diotic mode, wherein one lead fed both the left and right earphones. Koenig reported that background noises and reverberation were reduced under dichotic listening relative to diotic listening. He also reported that speech understanding was improved in the dichotic condition. Although no data were presented in this letter to the editor, Koenig's report is notable in that it was the first description of binaural squelch. Binaural squelch is defined as *the perceptual reduction of room effects when listening binaurally relative to monaural or diotic listening* (Koenig, 1950).

After the description of squelch provided by Koenig in 1950, the phenomenon received some attention for about thirty years. The first was a study performed in 1965 (Harris, 1965). A closed set speech-understanding task (materials: PAL test #8, unknown source) was used in this experiment. Three loudspeakers and two recording devices were set in a room with unreported reverberant properties. The recording devices were 12 inches apart from one another. The center loudspeaker was 12 feet away from the center of the recording devices. Flanking loudspeakers were located ±45° from the center loudspeaker and 12 feet away from the recording devices as well. Target sentences spoken by a male were played through the center loudspeaker and interfering speech was played through flanking loudspeakers. One interferer was female and the other was male. Which loudspeaker played which interferer was not specified. The target sentences with flanking interferers were recorded. The recordings were played to subjects over headphones. Subjects were asked to select the target sentence using multiple choice as a response method. The number of choices was not reported. 89 normal hearing listeners and 36 hearing-impaired (HI) listeners with asymmetric hearing losses were tested.

Listeners performed this speech-understanding task under many different listening conditions. Some of the listening conditions preserved binaural cues while some did not. The normal hearing listeners and hearing-impaired listeners saw similar improvements when they had access to binaural cues (Harris, 1965) compared to listening conditions that did not preserve binaural cues (i.e., monaural or diotic listening conditions). Harris also found that the poorer ear in the HI listeners could contribute meaningfully to speech understanding. HI listeners only scored 14.6% correct in listening conditions where they only used the poorer ear. Regardless of this impairment, the poorer ear contributed a 20.6% improvement in speech intelligibility when used in conjunction with the better ear in the stereo dichotic listening condition. Harris' results suggest that some central mechanism is able to take advantage of even a degraded acoustic signal in order to improve speech intelligibility (Harris, 1965).

Harris' study was not without its drawbacks. The report lacks details making it difficult to interpret. The ages of the subjects were not listed, the severity of the hearing loss in the HI listeners was not reported, audibility adjustments for the HI listeners (if any were made) were not discussed, the SNR at which the test was conducted was not reported, and the speech intelligibility task (PAL #8) has no citation for it (and does not seem to be used anymore), making it difficult to understand what listeners were asked to do. Regardless of these drawbacks, this study described important findings that motivated additional studies.

For example, Moncur & Dirks (1967) followed up on Harris' observation that the "worse" ear contributes more to reverberant speech understanding in binaural listening conditions than it does when alone (i.e., in monaural listening). They set an artificial head with microphones in the ear canals in a room with walls whose absorbent properties could be changed. A loudspeaker was 6 feet in front of the artificial head (0^o azimuth), and two flanking loudspeakers were set $\pm 45^o$ away on an arc with a radius of 6 feet centered on the head. The target speech materials were phonetically balanced word lists with 50 items (Egan, 1948) spoken by a male. The word lists were played out of one of the two flanking loudspeakers and recorded by the microphones in the artificial head. Some of the recordings had competing messages (one male, one female) played out of the center loudspeaker. Reverberation time (T_{60}) was changed by adjusting the absorbent properties of the walls without changing the size of the room. 48 normal hearing subjects were tested at four T_{60} values: 0.0 (i.e., anechoic), 0.9, 1.6, and 2.3 s. Half of the listeners performed the task with competing messages and the other half performed the task with no competing message. Each group was tested in all four reverberation times with three listening conditions: 1) binaural, 2) near ear only monaural, and 3) far ear only monaural.

Moncur and Dirks's findings corroborated Harris's findings. Both studies showed a clear binaural advantage in reverberant listening conditions. Both also showed that even a signal that is dominated by reverberant energy (i.e., the speech at the far ear) contributes information to the central auditory system that can be used in conjunction with a signal less affected by reverberation (i.e., the speech at the near ear) to improve speech understanding beyond the performance of either ear alone. This information can generate a 5-10% improvement in speech understanding of isolated words and sentences.

Reverberation has different effects on normal hearing (NH) and HI listeners' abilities to understand speech (Gelfand & Hochberg, 1976; Nabelek & Pickett, 1974a, 1974b). In Nabelek and Pickett's studies, listeners performed a speech understanding task using the modified rhyme test (MRT) (Bell, Kruel, & Nixon, 1972; Kruel et al., 1968) with a male speaker. The MRT is a speech intelligibility task that requires listeners to identify a monosyllabic English word by selecting it from a list of 6 possible answers (Bell et al., 1972; Kruel et al., 1968). Within each list, only the first or final consonant of the words differs (e.g., "went," "sent," "bent," "dent," "tent," "rent"). The carrier sentence "Number (item number), try (test word) now." was used. Speech was presented at 50 dB SPL for NH subjects and 60 dB SPL for HI subjects. HI listeners had, on average, a 72 dB pure-tone average bilateral, sensorineural hearing loss. Listeners (5 NH, all female; 5 HI, 2 female, 3 male) were seated in a large room (14'6" x 29'0" x 10'-14') with a sloping ceiling whose T_{60} could be changed from 0.3 to 0.6 s. HI listener wore hearing aids during the testing. Loudspeakers were set 11 feet away from the listener at $\pm 30^{\circ}$ azimuth. One loudspeaker produced eight-talker babble (half women, half men); the other produced the target speech. The target speech was produced from one loudspeaker on half the trials and from the other loudspeaker on the other half (i.e., both ears were tested equally as the near ear). Listeners were tested at different SNRs, two T_{60}s (0.3 and 0.6 s), and in monaural and binaural listening conditions. To achieve the monaural listening condition, the contralateral ear was plugged with a foam earplug and presented with a broadband noise presented at 82 dB SPL. This ensured that the speech and babble were sufficiently masked.

Nabelek and Pickett found significant effects of reverberation time (T_{60} = 0.3 or 0.6 s) and presentation mode (binaural or monaural) on speech understanding in both NH and HI listeners. For NH listeners at a given performance level (e.g., 70% correct), changing from monaural to binaural listening resulted in an effective SNR decrease of 3 dB. HI listeners received an effective SNR decrease of 1.5 dB when changing from monaural listening to binaural listening at a lower performance level (50% correct). This study demonstrates that HI listeners get a smaller benefit of binaural listening than NH listeners.

This smaller binaural benefit in HI listeners could be due in part to the hearing aids. When normal hearing listeners were asked to wear hearing aids with a flat amplification, their overall performance fell (Nabelek & Pickett, 1974b). To achieve a given performance level (70% correct), NH listeners required an increase in SNR of 2-3 dB when wearing hearing aids with a flat amplification relative to unaided listening. This was observed in all listening conditions. It is possible that this effect is due in part to the processing delay intrinsic to hearing aids. Binaural cues would be affected by this delay.

A small effect of age has been observed for binaural squelch (Nabelek & Robinson, 1982). In this study, listeners were divided into six groups of 10 with mean ages 10, 27, 42, 54, 64, and 72. The authors reported difficulties recruiting older listeners with minimal hearing losses (Nabelek & Robinson, 1982). The 10, 27, 42, 54, 64, and 72 year-old age groups had pure tone average (PTA) losses of 2.7, 5.9, 6.0, 10.9, 14.4, and 17.5 dB, respectively. The authors point out that these losses are similar to other studies examining speech understanding in the elderly (Jerger, 1973), however the correlation between age and PTA should be noted when interpreting the results. Listeners' speech understanding scores were measured using the MRT in different T_{60}s and in two listening conditions (monaural and binaural). The MRT was recorded in a room (7.5 x 6.1 x 3.6m) whose reverberant properties could be changed. The loudspeaker that produced the speech was set in a corner, 1m from each wall. A KEMAR manikin with microphones in its ear canals was placed 4m away from the loudspeaker along the room diagonal. Recordings were made at four T_{60}s (0.0s for monaural only; 0.4, 0.8, and 1.2s for both monaural and binaural). The prerecorded MRT was played over headphones to the listeners. In binaural listening conditions, sound was played through both earphones. In monaural listening conditions, one earphone was unplugged. No masker was presented to the contralateral ear. Because speech was presented at 70 dB SPL and no masking was used, it is possible that some of the sound "leaked" to the contralateral ear. Future studies should take precautions to avoid possible sound leakage.

The authors found main effects of age ($F(5,324) = 61.24$, $p < 0.001$), T_{60} ($F(2,324) = 84.90$, $p < 0.001$), and listening condition ($F(1,324) = 60.62$, $p < 0.001$). Mean performance for all age groups were significantly different except for the 42- and 54-year-old groups and the 64- and 72-year-old groups (no p value reported). Mean performance for T_{60} was significantly different between 0.4s and the other two groups (0.8 and 1.2s groups n.s.; no p value reported). Mean performance was significantly different between the binaural and monaural listening conditions (no p value reported). Generally speaking, speech intelligibility scores decreased with increasing T_{60} and increasing age (except for the 10 year-old age group). Binaural listening improved speech intelligibility across T_{60} and age.

The studies above show that increasing T_{60} decreases speech understanding using a variety of methods (Gelfand & Hochberg, 1976; Moncur & Dirks, 1967; Nabelek & Pickett, 1974a, 1974b; Nabelek & Robinson, 1982). That these effects were found in different labs, across decades, and using different methodologies lends power to their claims.

These studies also show that there is a binaural benefit in reverberant environments. This effect is large, even when using the ear with the *higher* SNR as a reference. This means using the ear with the higher SNR as a baseline results in the *smallest* benefit we would expect to see for NH and HI listeners.

These results align nicely with the original definition of binaural squelch: *the perceptual reduction of room effects when listening binaurally relative to monaural or diotic listening* (Koenig, 1950). There is clearly a benefit to having two ears in these difficult listening environments. That this mechanism has some binaural component is clear; however, the exact nature of this mechanism is unknown. The work summarized in Chapters 2, 3, 4, and 5 seeks to better understand the functional aspects of this mechanism through both psychophysical testing and modeling.

The perception of reverberation

Only a few major studies have examined how reverberation itself is perceived using subjective methods. In general, this was done by systematically varying physical aspects of sounds in rooms to understand how perception changes. Most of the work has been motivated by questions related specifically to concert hall acoustics, although previous work in our lab has also examined reverberation perception in this way for a smaller room more typical of an everyday listening situation. Other work has examined the effects of reverberation on perception by systematically varying the physical stimuli (Traer & McDermott, 2016; Zahorik, 2002a, 2002b), but these studies used objective methods and are therefore omitted here.

In 1981, Barron and Marshall examined how spatial impression (defined as the subjective width of a sound source) is affected by the spatial distribution of early reflections (Barron & Marshall, 1981) in a concert hall. Auditory source width (*ASW*) is now the preferred term for spatial impression (Beranek, 2004; Rumsey, 2002) and will be used in lieu of the author's original term for the phenomenon. Though Barron and Marshall report on several experiments in the original paper, only one is of interest to us here. In this experiment, listeners performed a matching task. They were instructed to make two different soundfields have the same *ASW*. On each trial, they were presented with two stimuli: a fixed standard, and a variable test stimulus. Listeners manipulated the sound level of early reflections in the test stimulus until they were satisfied that *ASW* was equal between the soundfields.

The soundfields were generated with simple impulse responses convolved with a Mozart musical excerpt and presented in the free field. The impulse responses contained the direct sound and two early reflections. In both the reference and the test stimuli, the early reflections were separated from one another by 1 ms. The first early reflection trailed the direct sound by 40 ms. In the reference stimulus, both reflections were 10 dB lower than the direct sound and the spatial properties of the reflections varied between blocks. In the test stimulus, the reflections approached the listener at ±90 degrees, and reflection level was under listener control.

Barron and Marshall found that the listeners changed their early reflection level matches based on the incidence angle of the early reflections. As the incidence angle of the early reflections in the reference stimulus departed from the interaural axis, listeners decreased the level of the early reflections in the test stimulus (Barron & Marshall, 1981). They used this information to derive the widely accepted formula for lateral energy fraction (LEF), a physical measure that underlies *ASW* (Beranek, 2004; Blauert, 1983; ISO-3382, 1997; Rumsey, 2002). In short, LEF considers that a reflection contributes energy proportional to its angle of incidence in azimuth and elevation. The closer the reflection is to the interaural axis, the more energy it contributes to *ASW*. This study represents one of the first studies that systematically manipulated physical aspects of reverberation to see how they affected reverberation perception.

A study by Bradley and Soulodre (1995) sought to tie a causal relationship between acoustic measurements and listener envelopment (*LEV*, defined as the impression that one is surrounded by a sound). Listeners performed a rating task, where they reported the degree of similarity on a 1-to-5 scale between two sound fields regarding *LEV*. Sound fields varied with regard to overall level (dBA), T_{60}, angular distribution of late reverberation, and C_{80}. Listeners made two judgments per stimulus. Manipulated BRIRs were convolved with a musical excerpt and presented in the free field.

The experiment showed significant effects of T_{60}, C_{80}, dBA, and angular distribution (Bradley & Soulodre, 1995). Based on these findings, Bradley and Soulodre derived a measurement (LG) to account for the level (dBA, C_{80}, and T_{60}) and spatial (angular distribution) aspects that underlies *LEV*.

A 2002 study by Okano examined the effects of many different physical attributes on *ASW* (Okano, 2002). They manipulated the following physical attributes: the level of the direct sound; the number of wall, ceiling, and rear reflections; the time delay of the wall, ceiling, and rear reflections; the level of the wall, ceiling, and rear reflections; and the angle of incidence of the wall, ceiling, and rear reflections. These parameters were grouped in different ways to generate 16 unique soundfields. The manipulated soundfields were used to spatialize a Mozart musical excerpt in the free field. For each of the soundfields, one set of reflections was free to vary in level.

Listeners performed a task similar to the method of constant stimuli. On each trial, the listeners heard two different soundfields. One of the soundfields was fixed from trial to trial and the other was free to vary. Listeners responded "same" if the soundfields were perceptually identical, and "different" if they were not. If listeners responded "different," they had to indicate along which dimension the soundfields were different. Then, the level of the variable reflections in the test stimulus was changed adaptively toward the fixed stimulus. This procedure was repeated until the listener responded "same." After a "same" response, the variable reflections became louder until the listener responded "different." After each such reversal, the step sizes became smaller until the listener responded "same" for the 0.5 dB step size. The track was then terminated.

They found that the perception of *ASW* is the most sensitive to energy in early lateral reflections (Okano, 2002). This study was limited, however, by a small sample size and the use of overtrained listeners. Regardless of these drawbacks, these results were consistent with Barron and Marshall (1981).

Previous work has examined the effect of listening condition on reverberation perception (Ellis, Zahorik, & Hartmann, 2016; Shore, Hartmann, Rakerd, Ellis, & Zahorik, 2016; Zahorik & Ellis, 2016). Most of this work relates to binaural squelch. Binaural squelch is a theory that predicts listeners will perceive more reverberation in a diotic or monaural listening condition than in a binaural dichotic listening condition. These studies used subjective techniques like multidimensional scaling and magnitude estimation to examine the effect of listening condition on reverberation perception.

In 2016, multidimensional scaling (MDS) was used to determine the stimulus properties underlying binaural squelch (Ellis et al., 2016). In this study, listeners made similarity ratings between pairs of stimuli. Using MDS, these similarity ratings were transformed into distances such that more similar stimuli are closer together (and less similar ones further apart) in a "stimulus space," where each dimension of the space is related to a physical property (e.g., intensity, frequency, etc.) or perceptual aspect (e.g., loudness, pitch, timbre, etc.) of the stimuli (Kruskal & Wish, 1978).

Twenty-five NH listeners participated in two MDS experiments (n = 12 in experiment 1, n = 13 in experiment 2). Virtual auditory space (VAS) techniques were used to generate stimuli in a large simulated room (14x10x5m, T_{60} = 1.8s) (Zahorik, 2009) for headphone presentation. For this experiment, simulated sources ranging from 1 to 12m were generated in front of the listener (0° azimuth). An anechoic condition at 1.4m was also generated. Three listening conditions were tested. A summed condition was created by adding the left (L) and right (R) channels. The result of this manipulation was put into both the L and R channel, making this listening condition diotic. A mirrored condition was created by presenting the R channel to both the left and right ears. Again, this manipulation made this listening condition diotic. Unaltered simulations were presented in a binaural condition. A speech stimulus spoken by a female talker ("Ready tiger go to red four now") from the coordinate response measure speech corpus (CRM) was used as the signal (Bolia, Nelson, Ericson, & Simpson, 2000). Speech stimuli were presented to the listeners at the simulated distances for each listening condition. In Experiment 1, no mirrored listening condition was used. In Experiment 2, all three listening conditions were used.

In both experiments, listeners provided two similarity ratings for all possible pairs of stimuli. The similarity ratings were analyzed using individual differences scaling MDS (INDSCAL), an MDS algorithm that preserves individual differences, allowing them to be analyzed (Carroll & Chang, 1970). The binaural stimuli were rated by all listeners as being very different from the mirrored and summed stimuli along a dimension strongly related to broadband interaural cross-correlation (IACC) (Ellis et al., 2016). This suggests that IACC underlies the differential perception of binaural and mirrored/summed stimuli.

Another study using subjective techniques to examine binaural squelch was presented at a conference (Shore et al., 2016). Binaural recordings of four Harvard sentences ("IEEE recommended practice for speech quality measurements," 1969) were made in a small room. Recordings were made at two distances (2m and 3m), with or without a foam head. Binaural and diotic listening conditions were included in addition to a single anechoic recording. Diotic listening conditions were generated by presenting the left channel to both ears. Every possible combination of distance, foam head, and listening condition was included, resulting in 8 different recordings plus 1 anechoic condition for each sentence. The 9 recordings of each sentence were grouped by sentence and put into 4 playlists on an iPod or in Windows Media Player. Files were randomly named one of the letters “A” through “I,” and a key identifying the sentences was kept for decoding later. Listeners were asked to sort the sentences within each playlist in order of increasing perceived room effect, defined explicitly to the listeners as “reverberation, coloration, or echo.” Nonparameteric tests were run to determine the effect of listening condition (binaural squelch), distance, and presence/absence of foam head. Significant effects of listening condition and distance were observed (Shore et al., 2016). Listeners perceived more room effect when listening diotically and for shorter distances (Shore et al., 2016). Interactions between distance, listening condition, and presence/absence of foam head were not analyzed.

These studies show significant effects of certain acoustic properties (i.e., the spatial location of reflections, spatial distribution of late energy, and listening condition) on how reverberant sounds are perceived. These studies indicate that reverberation perception can be changed, and causality established, by directly manipulating physical aspects of reverberation.

Methods background

VAS techniques

Virtual auditory space (VAS) techniques are a powerful tool available to psychoacousticians. In short, they allow for an experimenter to simulate a sound in an arbitrary room, *r*, at a given location relative to the listener, *x*, over headphone listening in a sound booth. This allows researchers to ask research questions that would otherwise be impossible due to logistics (e.g., juxtaposing two very different rooms on two adjacent trials, creating a virtual copy of a room that is unavailable due to expense or scheduling conflicts, etc.) or unreal listening conditions (e.g., an anechoic signal in one ear and a reverberant signal in the other, mismatched rooms across the ears, etc.).

There are many methods available to simulate acoustics in a virtual environment. The one most relevant for our discussion here is the image model. In its original conception, it is limited to rectangular rooms (Allen & Berkley, 1979). This is acceptable for our purposes here because we use exclusively rectangular rooms. The image model works by essentially creating mirror images of an omnidirectional point source in adjacent modeled rooms, then drawing vectors between the receiver and the mirrored sources (Figure 1).

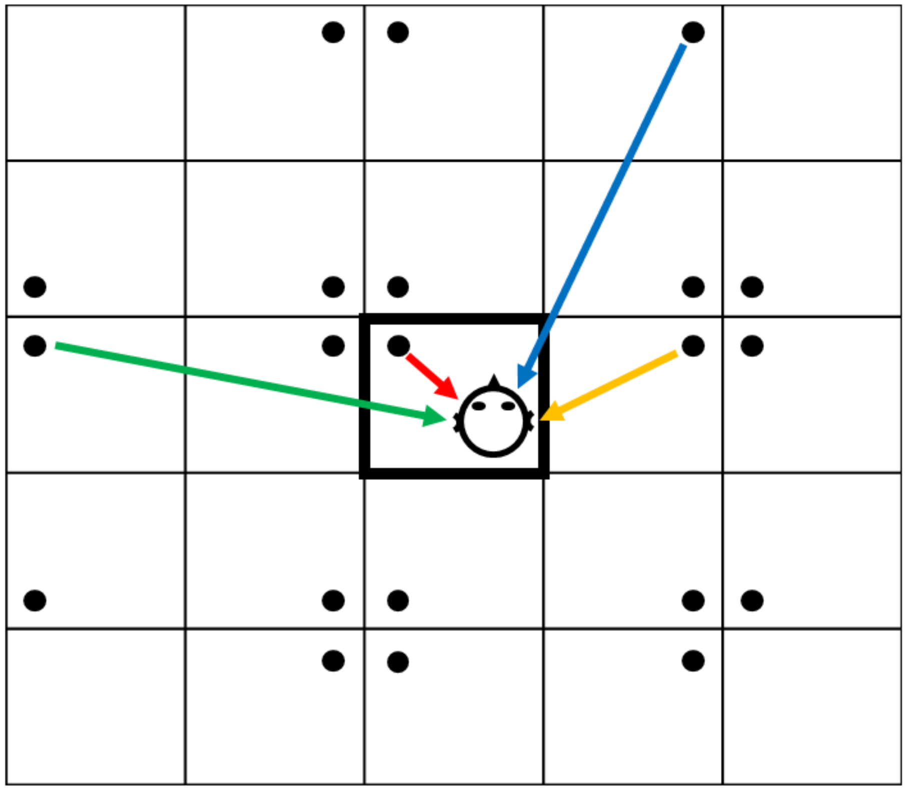


Figure 1. Simplified diagram of an image model. Modeled listening condition is in the boxed room in the center of the diagram. Direct sound in red, first order reflection in orange, second order reflection in green, third order reflection in blue. Each of the black points represents an omnidirectional point source. The human head represents a receiver.

The image model will perfectly represent walls with 100% reflectance, however this is not valid in the real world or in the simulations presented here. A coefficient (α) is used to change the absorbance of the walls in the model. α varies between 1 and 0, where 1 is an infinitely absorbent surface (i.e., no reflection), and 0 is an infinitely nonabsorbent surface (i.e., always generates a reflection).

It is computationally expensive to calculate large numbers of reflections. The model used in the experiments described here uses the image model for the first 500 reflections then models all energy after the 500^{th} reflection (referred to as "late reverberation") as noise in frequency bands that decays in octave bands as described by Sabine in the late 1890s. It has been shown that this technique produces simulated rooms perceptually similar to their real-world counterparts (Zahorik, 2009).

The direct sound, reflections and late reverberation are combined in a waveform to create a mathematical representation of the room and the location of the source and receiver in that room. This is the room impulse response (RIR). If the receiver has two inputs on the side of a modeled head, then this is referred to as binaural room impulse response (BRIR).

To realistically reproduce this sound over headphone listening, the shape of the listener's shoulders, head, and ears needs to be measured (Wightman & Kistler, 1989a). This is done by placing a loudspeaker in a desired location in an anechoic chamber and probe-tip microphones in the listeners' ears. A signal with known time and frequency properties (typically a maximum length sequence) is played out of the loudspeaker and recorded by the microphones. The effects of the listeners' shoulders, head, and ears can be captured by finding the differences between the known signal and the recorded signal.

The response captured by the microphones is referred to as the head-related transfer function (HRTF).

Finally, a headphone correction filter needs to be applied. This is done by measuring the frequency response of the headphones and generating a filter to correct for them. The HRTF is used to adjust the properties of the BRIR such that they match a specific listener, then the headphone correction filter is used to correct for the frequency response of the headphones. This generates a realistic stimulus that has been used to study localization (Brungart, Cohen, Zion, & Romigh, 2017; Hassager, Wiinberg, & Dau, 2017), distance perception (Anderson & Zahorik, 2014; R. Gilkey & T. R. Anderson, 1997; R. H. Gilkey & T. R. Anderson, 1997; Hassager et al., 2017; Zahorik, 2002a, 2002b), and many other psychophysical phenomena (Hassager et al., 2017; Reinhart, Souza, Srinivasan, & Gallun, 2016; Wightman & Kistler, 1989b; Zahorik et al., 2011) in virtual spaces.

Multidimensional scaling

Multidimensional scaling (MDS) will be used to analyze data in Experiment 2 (Chapter 4), so it is important to understand what it is and why it was chosen here. MDS is an analysis method that uses similarities between each pair in a set of stimuli to generate a perceptual map of that stimulus set. Distance between stimuli is proportional to perceptual similarity, such that stimuli that are perceived as similar are located nearer to one another and stimuli that are perceptually different are located further from one another in the perceptual map. The goal of this technique is to describe the perceptual map (hereafter referred to as the stimulus space) and extract what cues listeners were likely using when performing the similarity judgments. Physical measures of the stimuli (e.g., level, location, etc.) and perceptual measures of the stimuli (e.g., loudness, pitch, etc.) are typically used to interpret the stimulus space.

The power of this technique comes from the fact that listeners are never told explicitly what to do in the task other than to make similarity judgments. Listeners are therefore encouraged to use a multidimensional approach to the task. Not all psychoacoustic techniques allow for this—many demand listeners to use one dimension of a sound to perform a task (e.g., detection tasks, discrimination tasks, magnitude estimation, etc.). While these techniques offer a considerable amount of experimental control, they are not necessarily able to capture the multifaceted nature of human hearing.

Let us go through a concrete example of what an MDS algorithm does. Consider driving distance between U.S. cities. If the distances between every possible pair of 10 U.S. cities were subjected to an MDS analysis, the analysis would reproduce a two-dimensional map with accurate relative distances between each city. Driving distance is analogous to perceptual similarity between stimuli—stimuli that are more distant from one another are more different.

Since human experience is more complicated than driving distance, most perceptual studies are not two-dimensional solutions. The researcher must determine which n-dimensional solution is the best fit for the data. The goal is to maximize variance explained by the space (R^2) and to minimize how much the ratings needed to be changed to fit the space (Stress) without overfitting the data. To find this balance, a scree plot is typically examined, and the knee-point is found in the R^2 and Stress functions. This is the point where increasing the number of dimensions no longer contributes to a meaningfully higher R^2 or meaningfully lower Stress.

Once the appropriate number of dimensions are determined, the researcher can then proceed to interpret the axes of the space (Kruskal & Wish, 1978). This is done by performing a multiple regression where the coordinates in the space serve to predict a physical or perceptual value associated with the stimuli. The resulting regression weights can be used to plot a line through the space that explains the most variance for that physical or perceptual value (Kruskal & Wish, 1978). This aids in interpreting the space.

MDS, and techniques like it, have been used to study many multidimensional phenomena in hearing. The most notable is a study conducted on timbre perception (Grey, 1977). In this study, Grey had listeners rate the similarity between many synthesized musical instruments all producing the same note. Because they were synthesized, the different samples were normalized for pitch, loudness, and duration. Listeners performed similarity ratings between every pair of the synthesized stimuli and a specific MDS algorithm (INDSCAL, Carrol and Chang, 1970) was used for the analysis. This study found evidence that timbre perception is driven by spectral energy distribution, synchronicity in the envelopes of higher frequency harmonics, and temporal properties of inharmonic energy in the attack (onset) of the tone. More recently, the room acoustics literature has used MDS and techniques like it (e.g., principle cluster components analysis, cluster analysis, etc.) to determine the perceptual aspects underlying preference of different concert halls (Lokki, Patynen, Kuusinen, & Tervo, 2012; Lokki, Patynen, Kuusinen, Vertanen, & Tervo, 2011).

Motivation

This dissertation seeks to better understand reverberation perception using experimental designs that allow causal inferences between physical stimuli and perception to be drawn. The effects of monaural and binaural listening will be examined through manipulation of listening condition. These manipulations are motivated by robust preliminary findings related to binaural squelch. The physical amount of reverberation will also be manipulated to measure the influence of this acoustical property on perception. This manipulation follows studies that have similarly changed acoustical properties of sounds to determine their effects on perception. Listeners will be asked to judge how much reverberation they perceive, or to compare sound properties between two reverberant sounds. These studies will lead to a better understanding of how the human auditory system processes reverberant sound.

Chapter 2 will examine the effects of listening condition and physical amount of reverberation on how much reverberation listeners perceive. This experiment serves to test binaural squelch as a hypothesis and motivates the work described in subsequent chapters. Chapter 3 builds on the conclusions of Chapter 2, seeking to understand the role the reverberant tail plays in reverberation perception. Chapter 4 seeks to understand similarities and differences in reverberation perception between age-matched young normal hearing and young hearing-impaired listeners using MDS methods. Chapter 5 describes a model that was used to predict and explain the results of Chapters 2 and 4. Chapter 6 is a general discussion of the results of Chapters 2 through 5 and includes general conclusions.

CHAPTER II

MOTIVATING EXPERIMENTS: THE EFFECTS OF LISTENING CONDITION AND PHYSICAL LEVEL OF REVERBERATION ON PERCEIVED REVERBERATION IN NORMAL HEARING LISTENERS

The purpose of Chapter 2 is to explain the motivation behind the experiments described in Chapters 3 and 4. The stimuli are similar across the three chapters and the methods used in Chapter 2 are also used in Chapter 3. The work described in Chapter 2 was initially motivated by an attempt to examine binaural squelch. Based on the definition of binaural squelch, we would predict if the listener is in a monaural or diotic listening condition, then he/she will perceive more reverberation when compared to a binaural listening condition.

Perceived amount of reverberation: The effect of listening condition

The motivation behind these studies was to test a slightly more specific version of the hypothesis stated above. These studies are further motivated by the assumption that there is some neural mechanism underlying binaural squelch. If this mechanism requires binaural information about the *reverberant time-portion of the signal only*, then removing reverberation from one ear should produce the same results as a truly monaural listening condition (i.e., we can present the direct sound binaurally but present the reverberation monaurally). To serve as a control, a listening condition in which reverberation was removed from both ears was included. This listening condition should produce less perceived reverberation as the reverberation level decreases.

Methods

Subjects

These data were collected in three experiments. The data for these experiments were collected before this dissertation was proposed. As such, they will be referred to collectively as "the motivating experiments" throughout the document. All participants in the motivating experiments were undergraduates from the University of Louisville. 10 listeners (9 F, 1 M, ages 18 to 24, M = 20.3 years) participated in motivating experiment 1 (M1). 20 listeners (14 F, 6 M, ages 18 to 26, M = 20.5 years) participated in motivating experiment 2 (M2). 25 listeners (17 F, 7 M, 1 did not specify gender, ages 18 to 26, M = 20.3 years) participated in motivating experiment 3 (M3). All participants had normal hearing status as indicated by self-report. All methods were approved by the University of Louisville IRB. Listeners received course credit for their time.

Stimuli

Virtual auditory space (VAS) techniques using simulated binaural room impulse responses (BRIRs) were implemented to create all stimuli tested in this experiment. The BRIRs were computed using methods described by Zahorik (2009) in which non-individualized head-related transfer functions (HRTFs) were used to simulate the direct-path sound and 500 early reflections estimated using an image model (Allen & Berkley, 1979). Late reverberation in the BRIR was simulated using interaurally decorrelated noise that decayed per the Sabine equation in octave bands from 125 to 4000 Hz. Zahorik

(2009) demonstrated that this room modeling technique yielded results that were highly similar perceptually to those experienced in a real room.

Here, the simulated room was rectangular with dimensions of 10 x 14 x 5 m (L x W x H) and moderately hard surfaces. The broadband reverberation time of this room was 1.8 s. The sound source was located 3 m to the listener's right along the interaural axis (+90 degrees).

To create experimental BRIRs, reverberant sound energy was manipulated by scaling selected time portions of the BRIR in such a way that the direct-path sound was unmodified. The direct-path sound was identified as the first 2.5 ms of the BRIR. This value was chosen because it is the approximate length of an anechoic head-related impulse response (Wightman & Kistler, 1989a). This approach separating direct from reverberant energy is similar to that used in previous work (Zahorik, 2002b). Time portions of the BRIR after this point contained only the early reflections and late reverberation. This later portion of the BRIR was changed by *s* dB where *s* is a level listed in Table 1. Note that some of the stimuli in experiment 3 had reverberation level increased. These data are not discussed here, but are relevant for the modeling done in Chapter 5. The level manipulations changed the amount of reverberant energy relative to the unchanged stimulus. A change of 0 dB was equivalent to the unaltered room simulation. A change of -∞ dB was equivalent to a sound in anechoic space. Energy decay curves of a few experimental BRIRs are shown in Figure 2.

Table 1

Listening Conditions, Reverberation Level Values, and Number of Ratings for Magnitude Scaling Experiment

Experiment	Listening Conditions	Reverberation level	Judgments
M1	Ipsilateral hybrid	-∞, -21, -18, -15, -12, -9, -6, -3, 0	5
	Symmetric		
M2	Contralateral hybrid	-∞, -21, -18, -15, -12, -9, -6, -3, 0	4
	Contralateral mono		
	Ipsilateral mono		
M3	Ipsilateral hybrid	-∞, -18, -12, -6, 0, 3, 6, 9, 12, 15, 18, 21, 24	4
	Symmetric		

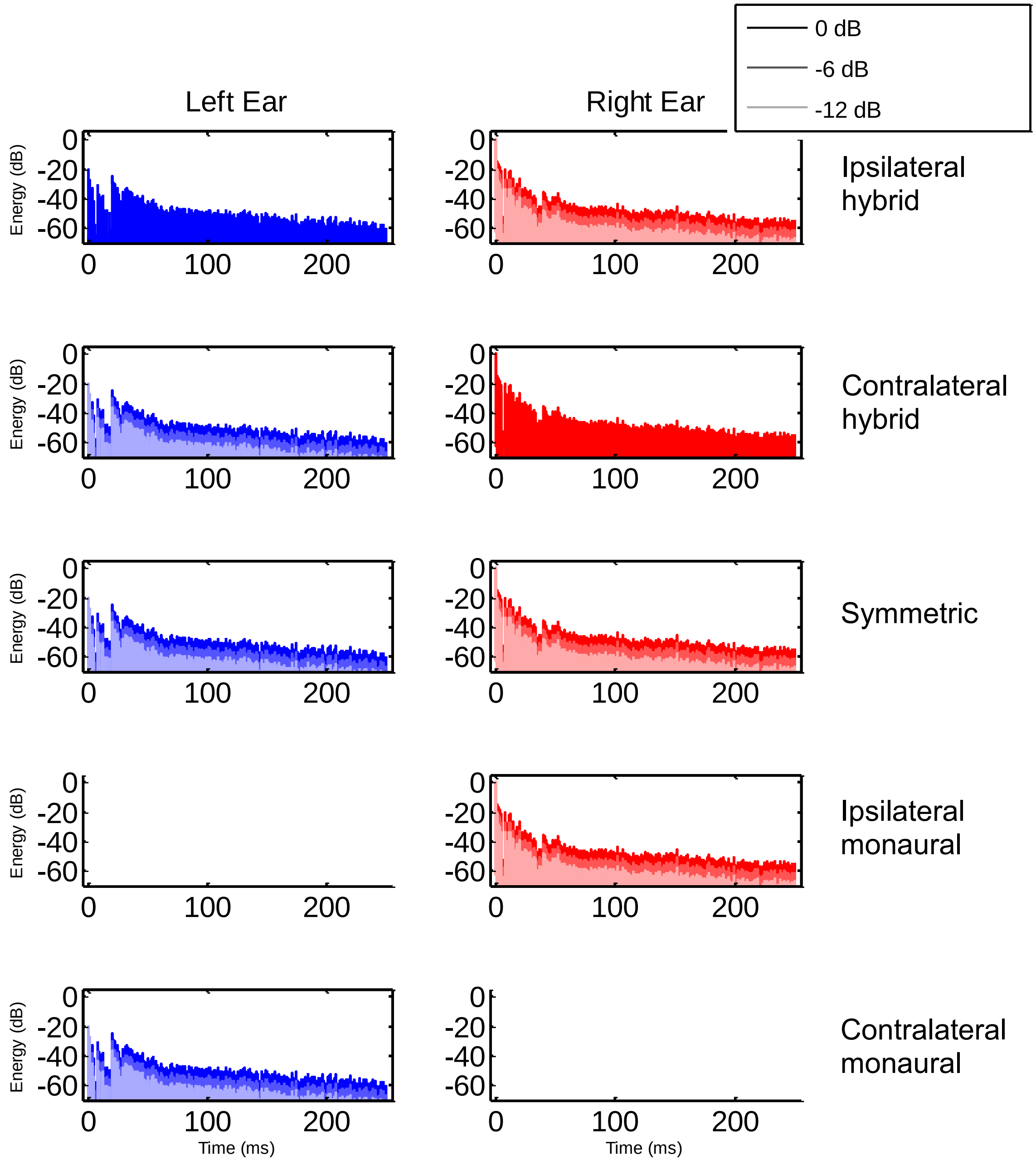


Figure 2. Energy decay curves of the first 250 ms of a few experimental binaural room impulse responses (BRIRs). X-axis is time and y-axis is dB FS. Shade of the waveform represents change in reverberation level (lighter shades indicate less reverberation present, i.e., larger changes in level). It can be seen here that the early reflections and late reverberation were changed while the direct path was left unaltered.

Five different listening conditions were tested. These conditions were generated by changing the amplitude of the early reflections and late reverberation present in the left and right channels of the BRIR in different ways. “Hybrid” listening conditions are those where the reverberation in either the left or the right channel was changed while leaving the other ear unchanged. In the ipsilateral hybrid listening condition, the ear nearer the source was changed (i.e., the ipsilateral ear). In the contralateral hybrid listening condition, the ear further from the source was changed (i.e., the contralateral ear). In the symmetric listening condition, both ears were changed equally. In the monaural listening conditions, one channel was changed while no signal was sent to the other ear.

The source signal was a single sentence (“Ready tiger, go to red four now.”) from the CRM corpus (Bolia, Nelson, Ericson, & Simpson, 2000), spoken by a female talker (talker #06). After convolution with the BRIR, level was adjusted across stimuli such that the 0 dB reverberation level stimulus had a peak presentation level of 65 dB SPL. The peak levels of all stimuli were the same. Overall levels of the other stimuli were lower than 65 dB SPL due to the level manipulation.

Procedure

Listeners performed a magnitude estimation task with a standard using a custom-built GUI developed in MATLAB (Mathworks, Inc., Natick, MA). This magnitude estimation task was similar to methods used in previous perceptual magnitude estimation tasks (Gescheider, 1985; Stevens, 1957). On each trial, listeners rated the amount of reverberation perceived in a test stimulus relative to a standard. Listeners were told that the standard stimulus had a reverberation of 100 arbitrary units. They were instructed to rate the amount of perceived reverberation in the test stimulus relative to the standard stimulus using a ratio scale (i.e., 200 is a doubling; 50 is a halving, etc.). Listeners could use any positive number desired, including decimals.

On each trial, participants could listen to the test and standard stimuli as many times as they desired. Text visible on the GUI reminded participants that the standard stimulus had a reverberation of 100 units and that they were to rate the amount of reverberation perceived in the test stimulus. Participants entered their responses using a keyboard and mouse. The task was self-paced and took less than one hour to complete.

Participants performed the task in a number of blocks equal to the "judgments" column in Table 1. Participants completed these blocks back-to-back. Stimuli were randomized across listening condition and scale value within each block using pseudorandom number generation in MATLAB.

Analysis

Data were pooled across blocks, experiments M1-M3, and participant and analyzed for each listening condition. Stimuli in which the reverberation level was increased were omitted from this analysis. A power function of the form $y = kx^a$ (where x is amount of reverberation removed from the signal and y is reported amount of perceived reverberation) was fit to the geometric means of the data for each listening condition using a least-squares method. This is the same methodology used in previous experiments tying physical measurements like intensity to perceptual phenomena like loudness (Stevens, 1957).

Results

Reverberation level was a significant predictor of perceived reverberation for the symmetric (a = 0.3665, $t(6)$ = 13.4529, p < 0.001), ipsilateral monaural (a = 0.3686, $t(6)$ = 16.0671, p < 0.001), and contralateral monaural (a = 0.2825, $t(6)$ = 10.8551, p < 0.001) listening conditions. Surprisingly, reverberation level was also a significant predictor of perceived reverberation in the contralateral hybrid (a = 0.0155, $t(6)$ = 4.8824, p = 0.0028) listening condition. Relative reverberation level was not a significant predictor of perceived reverberation for the ipsilateral hybrid (a = -0.0098, $t(6)$ = -0.8669, p = 0.4193). Aggregate results are shown in Table 2 and plotted in Figure 3.

Table 2.

Results of Magnitude Scaling Experiments

Listening condition	***k***	***a***	***p***
Ipsilateral hybrid	101.59	-0.0098	0.4193
Contralateral hybrid	101.86	0.0155	0.0028
Symmetric	121.62	0.3665	< 0.001
Ipsilateral monaural	102.62	0.3686	< 0.001
Contralateral monaural	89.44	0.2825	< 0.001

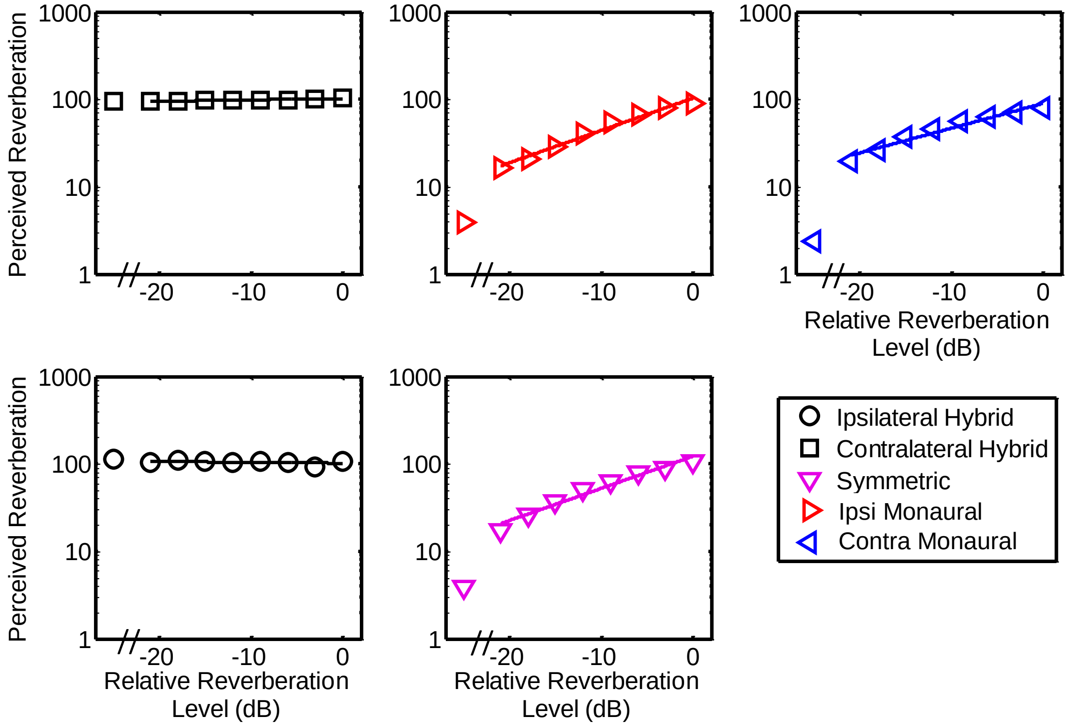


Figure 3. Aggregate results of the magnitude scaling experiments. Lines are power function fits through the geometric means. Each plot represents the results from a different listening condition. Starting from the top left they are: contralateral hybrid, right monaural, left monaural, ipsilateral hybrid, and symmetric.

Discussion

Listeners report hearing less reverberation when reverberation is removed from both ears equally or when removed from the only available ear under monaural listening conditions. In these listening conditions, the amount of reverberation they report hearing is proportional to the amount of reverberation removed following a power law. This is similar to many other relations between perceptual and physical measures: loudness and amplitude, brightness and intensity, etc. (Stevens, 1957).

Listeners do not report hearing less reverberation when reverberation is removed from only one ear so long as the other ear is left unchanged. The contralateral hybrid listening condition was observed to have a slope significantly different from 0; however, the effect is small and not meaningfully significant. This static response to reverberation in the hybrid listening conditions is true even if the signal in the changed ear is replaced with an anechoic signal (i.e., *all* reverberation is removed). The hybrid anechoic case was also tested by Sivonen, Alanko, Gamper, Raummukainen, and Pulkki (2011). They found using a 2AFC task that listeners could not discriminate perceived amount of reverberation between dichotic speech and a condition like our anechoic hybrid listening conditions (their "monaural reverb" condition). To generate their monaural reverb listening condition, they summed the reverberant energy at the two ears without adjusting level for the summation of the signals. This summation would produce a 3 dB increase in overall reverberant level. In our study, reverberation was left unaltered in the unchanged ear. Regardless of this 3-dB difference in relative reverberation level, different experimental design, and different source positions (Sivonen et al.: 0°; Present study: 90°), our results align nicely. In both experiments, perceived amount of reverberation is not perceptually different between the hybrid anechoic condition and unaltered dichotic listening.

CHAPTER III

EXPERIMENT 1: PERCEIVED AMOUNT OF REVERBERATION WHEN REVERBERANT TAILS ARE REMOVED

The reverberant tail of a sound in a room is defined as the energy that is present once the source stops producing energy. It is comprised of the early reflections and late reverberation. When asked to judge the amount of reverberation he/she perceives in a sound, a shrewd listener could wait for this reverberant tail and use that as an opportunity to get a clean “look” at the reverberation. In this way, it is possible listeners that participated in the motivating experiments used the reverberant tail of the stimuli to make judgments on perceived reverberation. To test if listeners were using only the reverberant tail as a cue to perceived reverberation, we removed it. If listeners were completely dependent on this cue, they would not be able to perform the task at all. If they were using it preferentially but not completely dependent on it, it would change their results. If their results did not change at all, then we can conclude that the reverberant tail was not necessary for listeners to perform the task. We would then conclude that listeners can use other cues in the ongoing stimulus to make judgments about perceived reverberation. That is, these ongoing cues would be sufficient to perform the task.

Methods

Subjects

23 undergraduates from the University of Louisville with normal hearing per self-report participated in Experiment 1 (14 female and 9 male, ages 18 to 35.6, *M* = 20.3 years, *SD* = 4.13 years). All methods were approved by the University of Louisville IRB. Listeners received course credit for their time.

Stimuli

The stimuli for this experiment were largely generated the same way as those in the motivating experiments (Chapter 2). In this experiment, two different listening conditions were tested. These conditions were generated by changing the early reflections and late reverberation present in the left and right channels of the BRIR by either: 1) changing reverberation in both ears equally ("symmetric"), or 2) changing reverberation only in the right ear and leaving the left ear unaltered ("ipsilateral hybrid").

Again, the source signal was a single sentence ("Ready tiger, go to red four now.") from the CRM corpus (Bolia, Nelson, Ericson, & Simpson, 2000), spoken by a female talker (talker #06). After convolution with the BRIR, level was adjusted across stimuli such that the 0 dB reverberation level stimulus had a peak presentation level of 65 dB SPL. The peak levels of all stimuli were the same. Overall levels of the other stimuli were lower than 65 dB SPL due to the level manipulation.

One extra step was performed compared to the methods in the motivating experiments (Chapter 2): the reverberant tails were removed from the convolved stimuli. This was done by finding the last non-zero sample in the anechoic stimulus and noting it.

Every sample after this in the reverberant stimuli was set to 0, resulting in stimuli with ongoing reverberation, but no reverberant tail.

In total, 17 stimuli were tested: 1 unaltered stimulus, 8 symmetric stimuli, and 8 ipsilateral hybrid stimuli. None of the stimuli tested had reverberant tails.

Procedure

The procedure was like that described in the motivating experiments (Chapter 2). Participants performed the task in five blocks of 17 trials, one for each of the stimuli (8 ipsilateral hybrid, 8 symmetric, 1 standard). Participants completed 5 blocks back-to-back for a total of 85 trials. Stimuli were randomized within each block using pseudorandom number generation in MATLAB.

Analysis

Data collected from individual participants were pooled across blocks and analyzed. A power function of the form $y = kx^a$ (where x is amount of reverberation removed from the signal and y is reported amount of perceived reverberation) was fit to the geometric means of the data for both listening conditions using a least-squares method. Data were then pooled across participants and analyzed, again fitting a power function to the geometric means.

Results

Aggregate results will be discussed first. For the symmetric condition, physical amount of reverberation in the signal was a significant predictor of perceived reverberation (a = 0.3817, $t(6)$ = 23.2853, $p < 0.001$). The power function fit to the geometric means explained a significant proportion of the variance ($F(1,6)$ = 542.2040, R^2 = 0.9891, $p < 0.001$). For the ipsilateral hybrid condition, physical amount of reverberation was not a significant predictor of perceived reverberation (a = -0.0110, $t(6)$ = -1.8769, p = 0.1096). The power function fit to the geometric means did not explain a significant proportion of the variance ($F(1,6)$ = 3.5226, R^2 = 0.3699, p = 0.1096). Fit parameters for individual subjects and summary statistics are reported in Table 3. Geometric means and lines of best fit for aggregate data are plotted in Figure 4.

Table 3.

Individual Subject Power Function Fits and Variance Explained
*: $p < 0.05$
**: $p < 0.01$
***: $p < 0.001$

Subject	Hybrid			Symmetric		
	a	***k***	R^2	***a***	***k***	R^2
QPN	-0.02	105.08	0.03	0.35	112.21	0.64 ***
QPO	0.02	101.93	0.07	0.98	188.71	0.79 ***
QPP	-0.02	100.05	0.42 ***	0.14	105.01	0.78 ***
QPQ	-0.01	101.59	0.00	0.16	91.73	0.62 ***
QPR	0.12	101.04	0.43 ***	0.87	152.58	0.87 ***
QPS	0.02	96.92	0.12 *	0.33	121.22	0.69 ***
QPT	0.02	96.11	0.02	0.97	189.91	0.77 ***
QPU	-0.01	100.12	0.01	0.22	117.04	0.68 ***
QPV	0.02	107.67	0.02	0.49	106.56	0.85 ***
QPW	-0.07	110.60	0.26 ***	0.24	114.23	0.67 ***
QPX	-0.07	106.18	0.24 *	-0.15	115.27	0.39 ***
QPY	0.02	121.21	0.05	0.14	126.29	0.73 ***
QPZ	0.01	102.29	0.14 *	0.15	99.55	0.75 ***
QQA	-0.03	116.69	0.11 *	0.19	125.24	0.45 ***
QQB	-0.06	96.95	0.07	0.23	111.54	0.76 ***
QQC	-0.05	104.11	0.16 *	0.37	119.11	0.83 ***
QQD	0.03	93.14	0.17 **	0.96	72.91	0.61 ***
QQE	0.11	91.59	0.15 *	0.99	176.91	0.86 ***
QQF	-0.01	125.20	0.00	0.16	113.31	0.28 ***
QQG	-0.13	102.89	0.69 ***	0.12	102.25	0.62 ***
QQH	-0.04	104.89	0.23 *	0.30	123.12	0.83 ***
QQI	-0.07	115.09	0.19 **	0.30	117.07	0.83 ***
QQJ	-0.05	108.66	0.15 *	0.27	110.38	0.78 ***
Mean	**-0.01**	**104.78**		**0.38**	**122.27**	
SD	**0.05**	**8.31**		**0.32**	**28.26**	

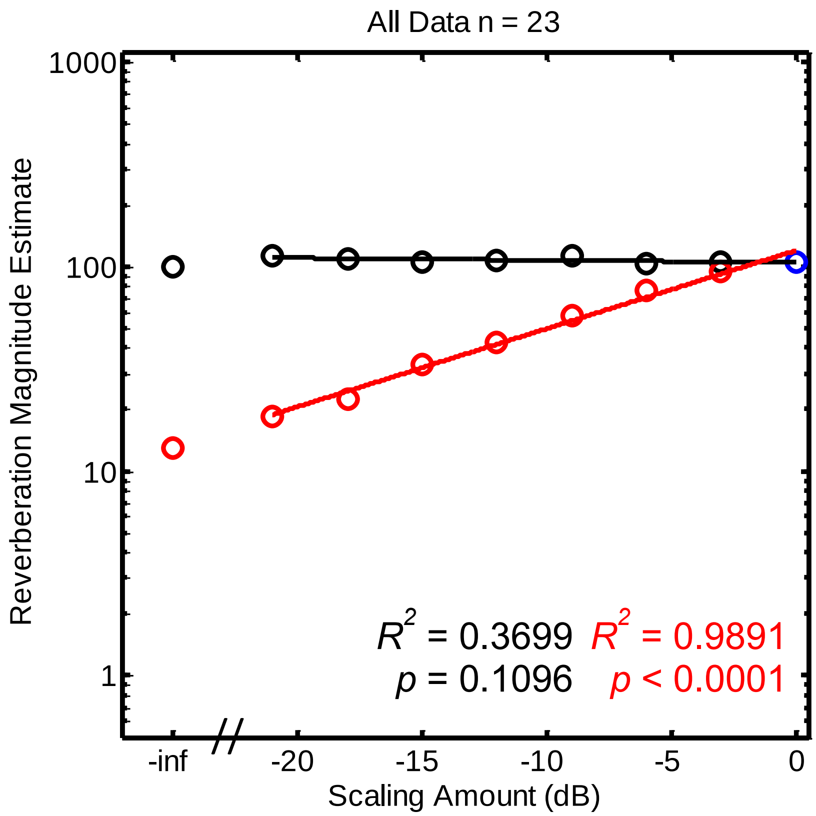


Figure 4. Aggregate results of experiment 1 (Chapter 3). Circles represent geometric means across subject and trial. Lines represent power function least square regression lines. Black plot features represent ipsilateral hybrid listening condition. Red plot features represent symmetric listening condition. R^2 and *p*-values reported for these fits in their respective colors.

Discussion

Chapter 2 demonstrated that listeners perceive less reverberation when reverberant energy is reduced in both ears but do not perceive less reverberation when reverberant energy is reduced in only one ear. The current experiment tested the hypothesis that listeners based their judgments solely on the reverberant tails of the stimuli. If this were the case, removing the tails would alter listeners' abilities to judge the amount of reverberation they perceived. The results of the present experiment were consistent with the explanation that listeners do not need reverberant tails to judge perceived amount of reverberation. This result aligns well with a previous study that removed the reverberant tail and found that listeners were still able to discriminate between reverberant stimuli (Sivonen et al., 2011). These results suggest that listeners can use only the ongoing reverberation to make judgments about reverberation.

This result disagrees with previous studies that have proposed the importance of the reverberant tail for the perception of sounds in rooms. The reverberant tail has been posited to underlie a room compensation mechanism (Watkins, 2005; Watkins & Raimond, 2013). In these studies, listeners were asked to indicate whether they perceived an ambiguous speech token to be the word "sir" or "stir" using a temporal cue. Before the test word, listeners heard a short context phrase. The reverberant properties of the test word and the context were changed in different experimental conditions. VAS techniques were used to spatialize the test word and the context and to manipulate the reverberant properties of each by changing the distance between the virtual source and receiver. When both the context and test word were "near" (0.32 m) to the listener, listeners reported hearing mostly "stir" for the ambiguous tokens. This was presumably because there was not enough reverberant energy to mask the temporal cue that distinguishes "stir" from "sir." When the context was "near" to the listener (0.32 m) and the test word was "far" (10 m), listeners reported hearing mostly "sir" sounds for the ambiguous tokens. This change is presumably because the reverberation in the "far" condition masks the temporal cue differentiating "sir" from "stir." When both the context and test word were "far," however, listeners again reported hearing predominantly "stir" for the ambiguous tokens. Watkins explains this recovered sensitivity to the temporal cue through a reverberation-compensation mechanism (Watkins, 2005; Watkins & Raimond, 2013). He specifically claims that the reverberant tail is important for this process (Watkins & Raimond, 2013), however this paper has some design and statistical flaws.

In terms of study design, all signals in the 2013 study were presented monaurally. These signals had the largest effects in his previous work (Watkins, 2005) and could have served to bias the results toward larger effect sizes. Monaural signals are also not ecologically valid and may be perceived differently than binaural signals when considering reverberation (Koenig, 1950; Koenig, Allen, Berkley, & Curtis, 1977; Zurek, 1979).

Statistically, Watkins does not report using ANOVA when his data would have been well-served by this test. A factorial ANOVA with precursor condition and test word as factors and category boundary as the dependent variable would suffice. Instead, Watkins calculates the difference between near and far test word category boundary within each precursor condition and runs *t*-tests on these difference scores. Running *t*-tests on difference scores is particularly problematic when there does not appear to be any effect of listening condition on the "near" test words. By calculating a difference score between “far” and “near” test words, Watkins is essentially subtracting a constant from each of the measures before running the *t*-tests. These statistical and design flaws cast doubts on the reliability of Watkins’s results.

A paper by Nielsen and Dau (2010) followed up on the results of Watkins’ 2005 study using the same materials Watkins did. Nielsen and Dau used a number of dry signals and showed that no reverberation compensation mechanism was necessary to explain Watkins’ results (Nielsen & Dau, 2010). They found that amplitude modulated noise provided a recovery of “stir” responses like that of the “far/far” test condition. They propose that amplitude modulation-based forward masking at least partially explains the results of Watkins’ original study citing previous work that has shown this effect (Wojtczak & Viemeister, 2005).

The present study is limited in that it addresses only one aspect of reverberation perception and does this from one spatial location. It would be valuable to continue this work with other perceptual aspects of reverberation perception (e.g., *ASW*, *LEV*, etc.), and in other spatial locations (e.g., distances, locations in azimuth, etc.).

Conclusions

Experiment 1 demonstrates that listeners can extract reverberation level information from an ongoing reverberant speech signal located to the side of the head. This result is consistent with past work indicating that ongoing reverberation is sufficient to make judgments on perceptual properties of reverberation.

CHAPTER IV

EXPERIMENT 2: MULTIDIMENSIONAL SCALING OF REVERBERANT SOUND IN YOUNG NORMAL HEARING AND YOUNG HEARING-IMPAIRED LISTENERS

Hearing-impaired listeners struggle with listening in reverberant environments. For this reason, we wanted to examine similarities and differences between normal hearing and hearing-impaired reverberation perception. To better control for the effects of hearing impairment separate from age, we used age-matched young normal hearing (YNH) and young hearing-impaired (YHI) listeners. An open-ended experimental design—multidimensional scaling—was used to avoid studying only one aspect of reverberation perception. This allowed YNH and YHI listeners to use whatever cues they found most salient to perform the task.

Methods

Subjects

15 young normal hearing (YNH) listeners (Age = 25.2 ± 3.9 years; Gender: 11 female, 4 male) recruited via word of mouth participated in Experiment 2. Normal hearing status was defined as having pure-tone audiometric thresholds below 20 dB HL for all frequencies between 250 Hz and 8000 Hz spaced in octaves. Normal hearing status was verified through audiometric screening performed by the author.

11 young hearing-impaired (YHI) listeners (Age = 25.2 ± 4.4 years; Gender: 8 female, 3 male) recruited from the Heuser Hearing Institute (HHI) database participated

in Experiment 2. Of these 11 listeners, two were lost to attrition and a third did not complete the time-forward experimental block. The data available for these listeners were analyzed. Inclusion criteria were chosen to age-match to normal hearing listeners, to ensure that the NAL-R gain rule would provide adequate gain for these listeners (Byrne & Dillon, 1986), that the listeners did not report tinnitus, and that the listeners did not report symptoms consistent with Meniere's disease. The etiologies for these listeners varied, but all listeners had a sensorineural loss. Some had an additional conductive loss in the low or mid-frequencies. After searching on the HHI database, 70 listeners met these criteria. All 70 were contacted via phone call or email. Hearing status was confirmed with pure-tone audiometry for all HI listeners. Pure-tone audiometric thresholds for YNH and YHI listeners are plotted in Figure 5

All participants were compensated $20 per hour for their participation in the study. All methods were approved by the University of Louisville and HHI IRBs.

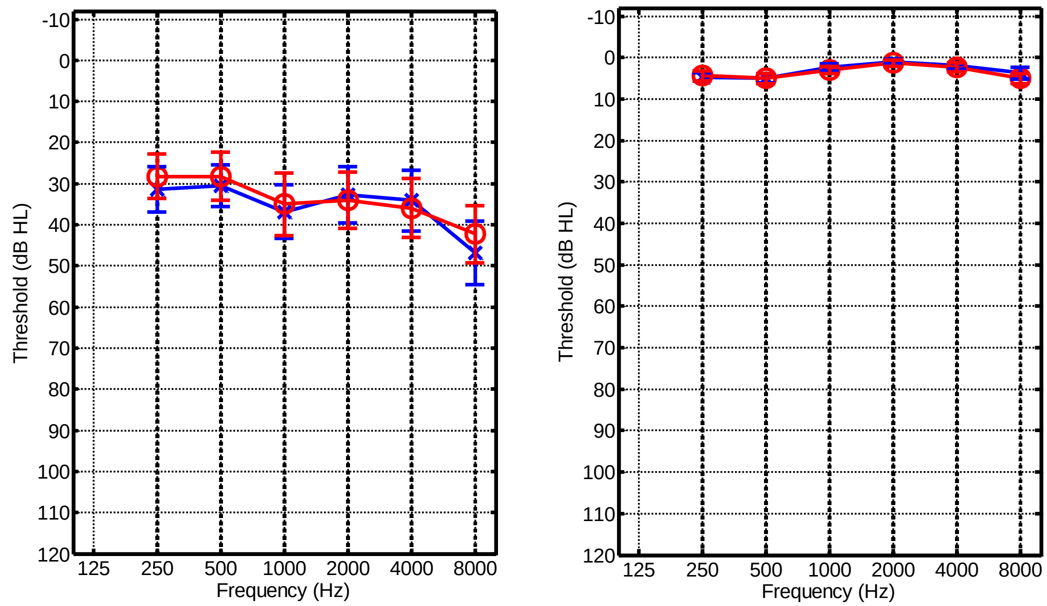


Figure 5. Audiograms for the 11 young hearing impaired listeners (left) and the 15 young normal hearing listeners (right) that participated in Experiment 2. Group means represented with circles, error bars are standard deviations. Red lines and features are right ear thresholds. Blue lines and features are left ear thresholds. Individual audiograms can be found in Appendix B.

Stimuli

Stimuli were generated by changing reverberation level like in Experiment 1 (Chapter 3). The unaltered BRIR was the same as that used in Experiment 1 (Chapter 3). The direct energy was defined as the first 2.5 ms after the first zero crossing in the BRIR. Everything after this point was changed to generate experimental BRIRs. Reverberation level values were -12, -6, 0, or +6 dB in one of five listening conditions: 1) ipsilateral hybrid, 2) contralateral hybrid, 3) symmetric, 4) ipsilateral monaural or 5) contralateral monaural. The unaltered BRIR was also included. Since the 0 dB reverberation level conditions in the ipsilateral hybrid, contralateral hybrid, and symmetric listening conditions are identical to the unaltered BRIR, these three stimuli were excluded. With 4 level values, 5 listening conditions, and 3 conditions excluded due to redundancy, 17 experimental BRIRs and one unaltered BRIR were generated. In addition to manipulating reverberation level, interaural cross-correlation was systematically altered.

To manipulate interaural cross-correlation, the VAS techniques used in our lab were combined with an older method used to manipulate the correlation between two noise sources. The room modeling software uses a pair of decorrelated noise generators to generate late reverberation. A pair of decorrelated noise generators were also used in a study from the binaural masking level difference (BMLD) literature to manipulate interaural correlation (Dolan & Robinson, 1967; Robinson & Jeffress, 1963). By using the technique outlined in the BMLD papers in conjunction with the virtual pair of noise generators used in the room model, we can generate any desired correlation between the reverberant portion of the channels in the impulse responses.

The BMLD technique can be illustrated by walking through the following scenario. Suppose we have two independent noises, N_1 and N_2 in the left and right channels of a headset. This will result in decorrelated noise in the two channels ($r \approx 0.00$, this is also the typical scenario in the room modeling software). Using the equation below, we can manipulate the correlation between the noises (r).

$$r = \sqrt{\frac{a^2 * RMS_1^2}{a^2 * RMS_1^2 + RMS_2^2}}$$

Where RMS_1 is the RMS of N_1, RMS_2 is the RMS of N_2, and a is the amplitude of the noise present in both channels (Dolan & Robinson, 1967; Robinson & Jeffress, 1963). We can solve for a and get an equation that will allow us to generate r:

$$a = \sqrt{\frac{r^2 * RMS_2^2}{RMS_1^2 - r^2 * RMS_1^2}}$$

This equation allows for the approximate online generation of any desired r. Once a is known, appropriately mixing the noises is trivial. By mixing the noises in the room modeling software, the correlation of the reverberant portion of the impulse response can be varied systematically.

Three interaural coherence-manipulated impulse responses were generated using the technique above. The target coherences were 0 (the standard case), 0.60, and 1.00. These 3 interaural-coherence manipulated impulse responses in addition to the 18 impulse responses in which reverberation level was altered gave us 21 total impulse responses.

These 21 BRIRs were then convolved with a phrase ("Ready Tiger") from the CRM corpus spoken by a female talker (Bolia et al., 2000). This was a portion of the sentence used in Experiment 1 (Chapter 3).

It has been suggested that speech intelligibility and reverberation perception may interact (Personal Communication w/ Tapio Lokki, unpublished sound quality data from our lab). To test for this, time-reverse speech was tested as well. Time-reversed speech is spectrally identical to time-forward speech but is unintelligible, and therefore serves as a good control for intelligibility. The time-forward phrase used for the first stimulus set ("Ready Tiger") was time reversed, then convolved with all BRIRs.

Before presentation to the YHI listeners, an NAL-R (Byrne & Dillon, 1986) gain rule was applied via MATLAB. NAL-R is a linear gain rule that restores audibility at a comfortable listening level for mild-to-moderate hearing losses (Byrne & Dillon, 1986). YNH listeners received no gain prescription, though the signals were passed through the same software to avoid potential signal processing differences. In addition to the NAL-R gain rule, a headphone correction filter was used to correct for the Sennheiser HD-200 headphones used in the booth. After the different corrections and gains were applied, the stimuli were presented at 65 dB SL.

Acoustic Measurements

Acoustic measurements were taken to help interpret the results of the multidimensional scaling (MDS) analysis. Measurements were taken in octave-wide frequency bands centered on 125, 250, 500, 1000, 2000, and 4000 Hz. Octave-wide bands were acquired by passing the impulse response through a 3rd order bandpass filter with the appropriate cutoff frequencies. In addition to the octave-wide bands, a broadband measure was calculated by passing the impulse responses through a 4th order lowpass Butterworth filter with a cutoff frequency at 4 kHz. The acoustic variables were those measured according to ISO-3382 (T60, C80, C50, Tc, IACC and DR) and overall power. Acoustic measurements were calculated for both the left and the right ear for all variables except IACC, which is a binaural measure. Taken together, these measurements provided 91 variables to explain the space (13 measures * 7 frequency bands). Due to the considerable number of predictors, principal components analysis was used to reduce the 91 variables to a fewer number of orthogonal predictors. Three components were found to explain most of the variance ($R^2 = 0.92$, see Figure 6).

Correlations showed that component 1 was strongly related to reverberation level ($r = 0.94$, $p < 0.001$) and broadband IACC ($r = -0.83$, $p < 0.001$). These values are inherently related to one another based on the way IACC is calculated. IACC takes the full impulse response in to account—this includes the highly correlated direct sound and decorrelated reverberant energy. Therefore, as reverberation level is increased (i.e., reverberation level is reduced), the relatively decorrelated portion of the signal is slowly removed, increasing the contribution of the highly correlated portion of the signal (i.e., the direct sound). This causes overall IACC to increase as more reverberation is removed from the signal. Component 1 captures this aspect of the stimuli and will be referred to as "reverberation level" from here on.

Component 2 was strongly correlated with every broadband monaural reverberation acoustic measurement (Table 4). The simplest way to interpret this component is as "any monaural measure of reverberant properties" and will be referred to as "monaural reverberation measurements" from here on

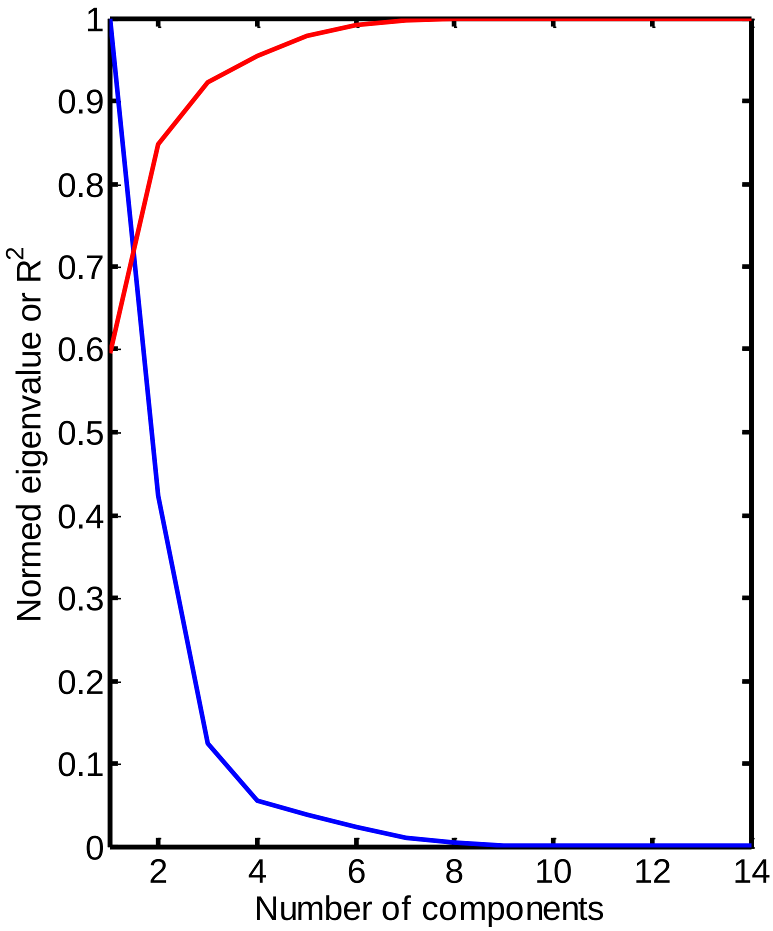


Figure 6. Scree plot for principal components analysis. R^2 plotted in red. Normed eigenvalue plotted in blue. Knee point appears to be around 3 components.

.

Table 4

Correlations Between Component 2 and Acoustic Measurements

Measure	***r***	***p***
T_{60} Left	**0.49971**	**0.021**
T_{60} Right	**-0.65872**	**0.001**
C_{80} Left	**-0.50388**	**0.020**
C_{80} Right	**0.66177**	**0.001**
C_{50} Left	**-0.50278**	**0.020**
C_{50} Right	**0.66223**	**< 0.001**
IACC	-0.1513	0.513
Tc Left	**0.48409**	**0.026**
Tc Right	**-0.60443**	**< 0.001**
DR Left	**-0.50678**	**0.019**
DR Right	**0.66619**	**< 0.001**
Total power Left	0.3461	0.124
Total power Right	-0.089512	0.700

Component 3 was strongly correlated to broadband power in the right ear (r = 0.96, $p < 0.001$). This component is self-explanatory and will be referred to as "broadband right ear power".

Based on the results of the PCA, we were able to reduce the number of acoustic measurements. This reduced space of orthogonal acoustic measurements was used to interpret the stimulus spaces.

Procedure

Listeners were brought to the third floor of Heuser Hearing Institute (HHI) to fill out paperwork. Listeners gave their informed consent, filled out a personal information worksheet, and filled out a W-2 form for payment. After this was done, all listeners were brought to a sound booth on the second floor of HHI for audiometry. YNH listeners were screened at 20 dB HL, and the YHI listeners had their audiograms confirmed by an audiologist. Listeners were then brought back to the third floor to start the experiment. Paperwork and audiometry took approximately 0.5 hours to complete.

The experiment was run in two phases: a familiarization phase and a rating phase. The familiarization phase was designed to expose listeners to the range of stimuli, allowing them to make informed similarity judgments (Carroll & Chang, 1970; Kruskal & Wish, 1978; Schiffman, Reynolds, & Young, 1981). Listeners were given instructions that briefed them on this process before beginning the familiarization phase. After the instructions, listeners were given a list of the 21 stimuli they would be rating. They were instructed to listen to every sound in the list at least once, and to make sure they knew how similar and how different the sounds could be from one another. Once they were satisfied that they were familiar with the list, they exited the booth and took a break before the rating phase.

During the rating phase, listeners were asked to judge the similarity between pairs of stimuli using custom software with a 100-point slider in an interface developed in MATLAB (Mathworks, Inc., Natick, MA). Stimulus order was randomized using pseudo-random number generation in MATLAB. The slider was anchored on the left side by the label "Exact Same" and on the right by "Completely Different." Every time a listener heard a pair of stimuli they were in the same order (i.e., for a given pair of sounds A and B, only the order A | B was tested). No trials in which a stimulus was repeated were presented. All listeners were given the same instructions before beginning the task, asking them to judge the similarity of the stimuli using whatever criteria they felt were most appropriate. It was stressed that there were no right or wrong answers. The rating phase took place in two blocks consisting of 210 trials per block. Listeners were prompted to leave the booth after the first block was over and could choose to do so if they wished. After taking the optional break, they returned to the booth and finished the second block. It took approximately 1.5 hours total to run the familiarization and rating phases.

The familiarization and rating phases were performed the same way for both the time forward and time-reversed speech stimuli. These two conditions were blocked separately. The order in which listeners participated in the time forward and time-reverse speech blocks was counterbalanced to control for order effects. Overall, it took listeners approximately 4 – 4.5 hours to complete the paperwork and both experimental conditions.

Analysis

Individual differences multidimensional scaling (INDSCAL)

Data were analyzed using INDSCAL in IBM's SPSS version 25. The following parameters were used for all INDSCAL solutions. The rating data were treated as ordinal data. Any ties between stimulus ratings were broken with an upper limit of 3000 ties. A solution for a given dimension was found within 30 iterations. When an iteration did not reduce stress by more than 0.005, the iterations were stopped, and that solution was complete. Solutions between 2 and 6 dimensions were examined.

YNH and YHI listeners were pooled for the time-forward speech stimuli. YNH and YHI listeners were also pooled for the time-reverse speech stimuli. Data from one YHI listener were omitted from the time-reverse analysis because they only responded using "exact same" or "completely different" for this condition. The remaining listeners were pooled to test whether the groups were different from one another for either stimulus set. INDCSAL requires the groups to be pooled then analyzed to determine whether the groups are different. Three-dimensional solutions explained large proportions of the variance in both datasets (Time-forward: R^2 = 0.6635; stress = 0.2321; Time-reverse: R^2 = 0.6125; stress = 0.2477) and were chosen for this reason and to facilitate stimulus space interpretation. Higher-order solutions did not add to explanatory power. Scree plots for the solutions are found in Figure 7.

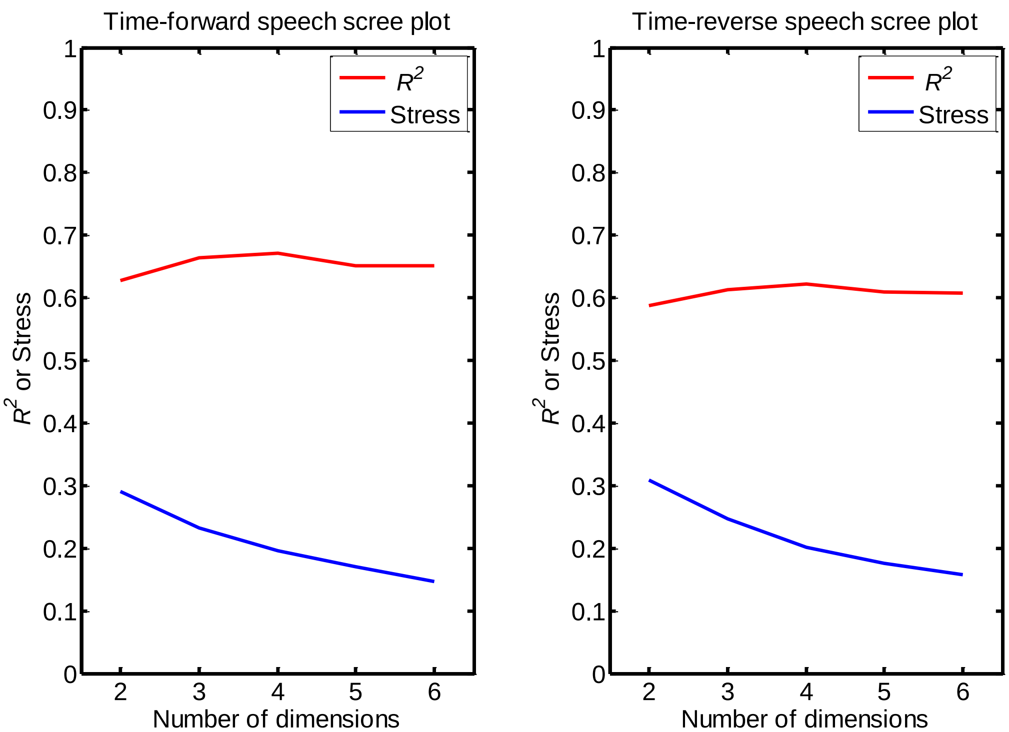


Figure 7. Scree plots for multidimensional scaling solution. Solutions with between two and six dimensions were considered. The red line represents R^2. The blue line represents Stress. A three dimensions solution was chosen because it appeared to be a knee point and because higher-order solutions did not offer more explanatory power.

Subject spaces

The subject space was examined to determine whether grouping the YNH and YHI listeners was appropriate. Subject spaces for the 3-dimensional solution can be found in Figure 8.

An analysis of angular variance (ANAVA) was run to determine whether the groups were different along any of the three axes ("Circular data analysis," 2018; Schiffman et al., 1981). ANAVA is analogous to ANOVA, except it is used on data in a three-dimensional space. The interpretation of this result is the same as that for ANOVA . Group means are replaced by the mean vector for each group and will be reported as a vector [*x y z*]. These values can be interpreted as the group's mean weights on dimensions 1, 2, and 3 for *x*, *y*, and *z*, respectively. Standard deviation is replaced by circular standard deviation (*v*) and can be interpreted in a similar fashion as regular standard deviation.

No statistically significant effect of hearing status was found ($F(1,24) = 3.73$, $p = 0.07$) between the YNH ([0.7062 0.4855 0.4577], $v = 0.2404$) and YHI listeners ([0.6082 0.6080 0.4903], $v = 0.1425$). See Table 5. Referencing Figure 8, listeners ZGT and ZGZ appear to be outliers and may be contributing to increased variability of the YNH listeners, and therefore obscuring potential group differences. If their data are omitted, the effect of hearing status is still not statistically significant, however ($F(1,22) = 2.1826$, $p = 0.1538$).

The same procedure was followed for the time-reversed speech (Figure 9). Again, no statistically significant effect of hearing status was found ($F(1,21) = 3.66$, $p = 0.07$) between the YNH ([0.7148 0.5069 0.4627], $v = 0.1347$) and YHI listeners ([0.6583 0.4815 0.5624], $v = 0.1369$). See Table 6. Unlike in the time forward subject space, there were no systematic outliers. Listeners ZHA, ZGK, and ZGQ were slightly different from the other YNH listeners, but listener RBE was different from the YHI listeners. Therefore, no secondary analysis was run omitting these outliers.

Taken together, these results suggest that YNH and YHI listeners do not adopt radically different strategies for judging similarities in reverberation characteristics between pairs of sounds.

Table 5

Effect of Hearing Status on Perception of Time Forward Reverberant Speech

	SS	*df*	MS	*F*	*p*
Within	0.0822	1	0.0822	3.7311	0.0653
Between	0.5383	24	0.0224		
Total	0.6206	25			

Table 6

Effect of Hearing Status on Perception of Time-Reversed Reverberant Speech

	SS	*df*	MS	*F*	*p*
Within	0.0363	1	0.0363	3.6574	0.0696
Between	0.2101	21	0.01		
Total	0.2464	22			

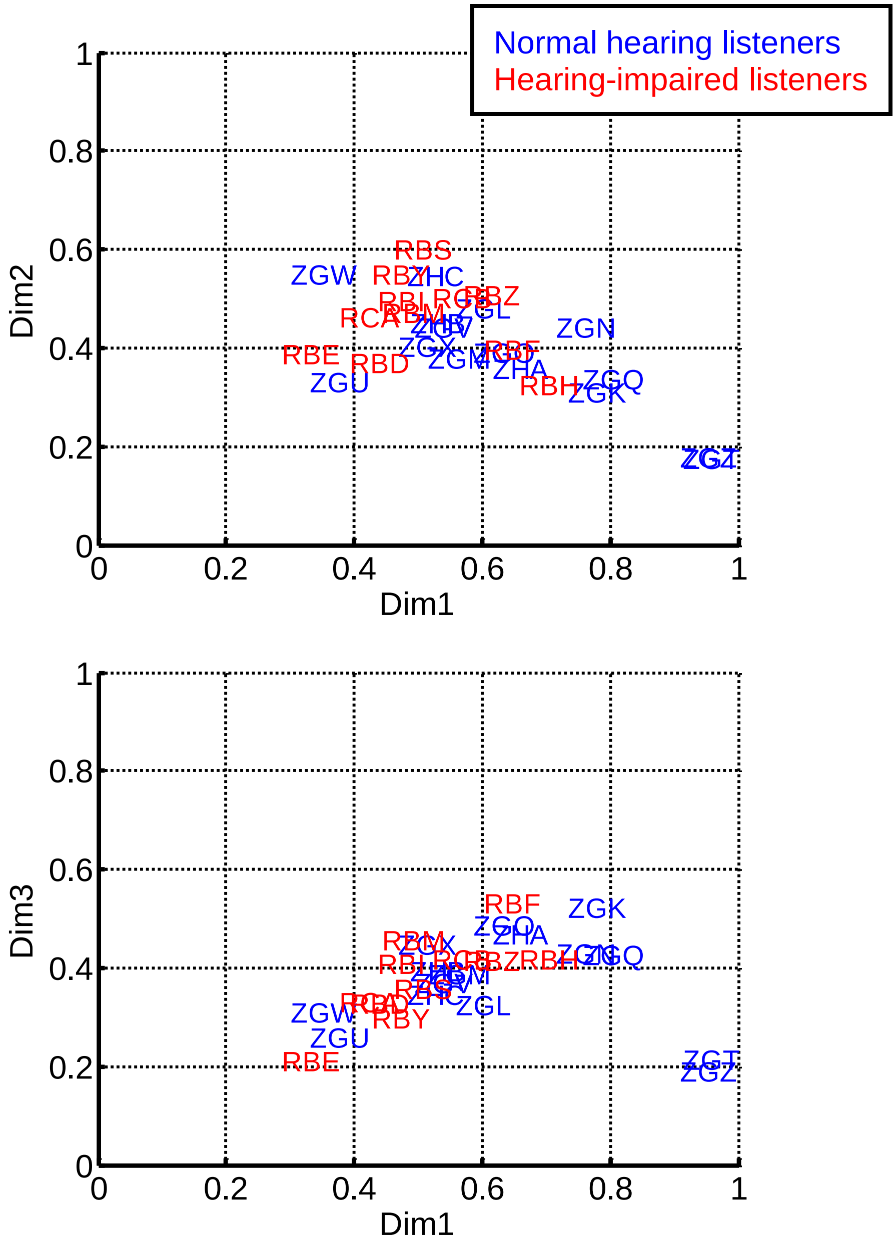


Figure 8. Projections of the 3-dimensional subject space for time-forward speech data. YHI subjects are plotted in red and YNH subjects are plotted in blue. These plots show the relative importance of each dimension for a given listener. Weights are proportional to the importance that a given dimension played in a subject's similarity ratings. The differences between the YHI and YNH group are not significant indicating good agreement across groups. See text for details.

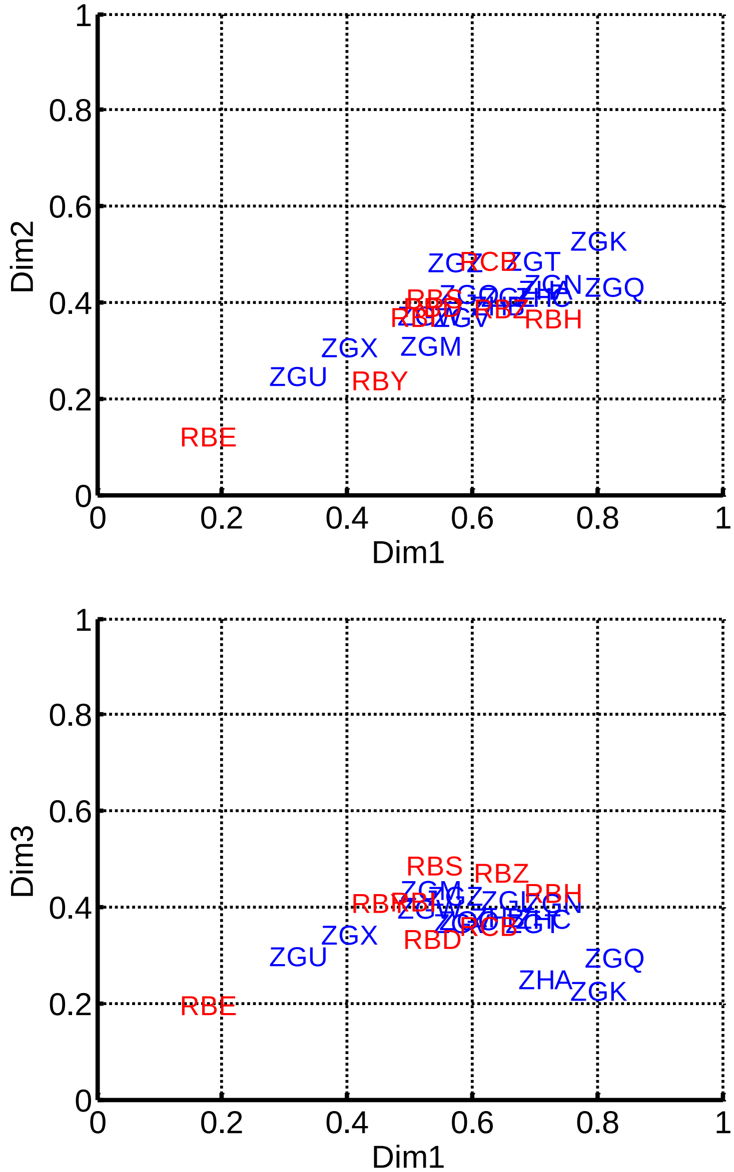
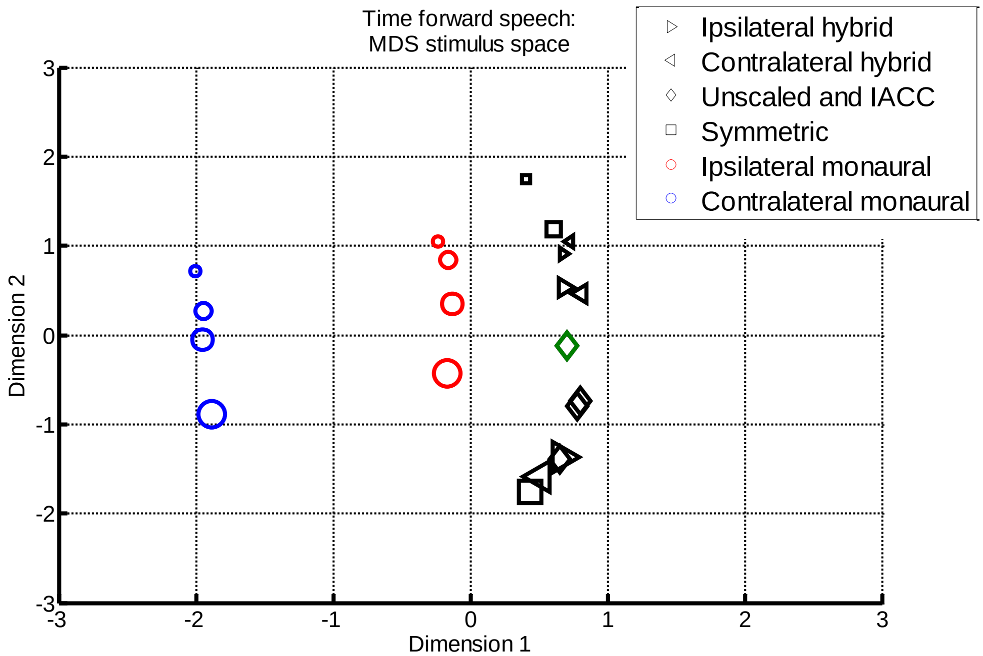


Figure 9. Projections of the 3-dimensional subject space for time-reverse speech data. YHI subjects are plotted in red and YNH subjects are plotted in blue. The differences between the YHI and YNH group are not significant indicating good agreement across groups. See text for details.

Stimulus spaces

A primary goal of MDS analysis is determining what properties listeners use to make similarity judgments. This is done by examining and analyzing the stimulus space. Projections of the time-forward stimulus space are plotted in Figure 10. Projections of the time-reverse stimulus space are plotted in Figure 11.



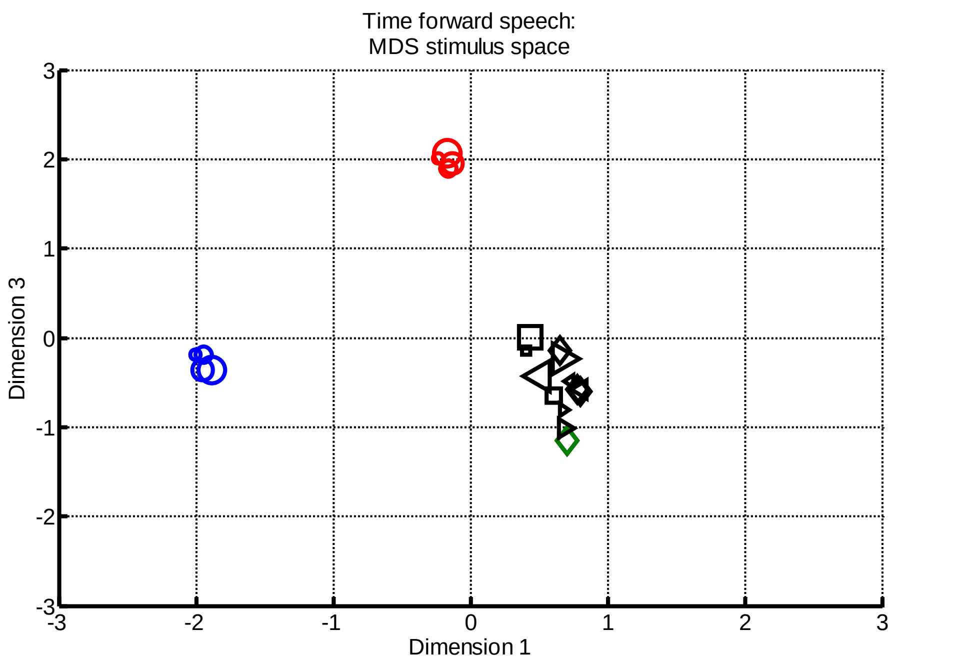


Figure 10. Stimulus space for the time forward stimuli. Size represents reverberation level. From smallest to largest: -12 dB, -6 dB, 0 dB, +6 dB. Color: red, ipsilateral monaural; blue, contralateral monaural; black, binaural. Shape: listening condition. See legend.

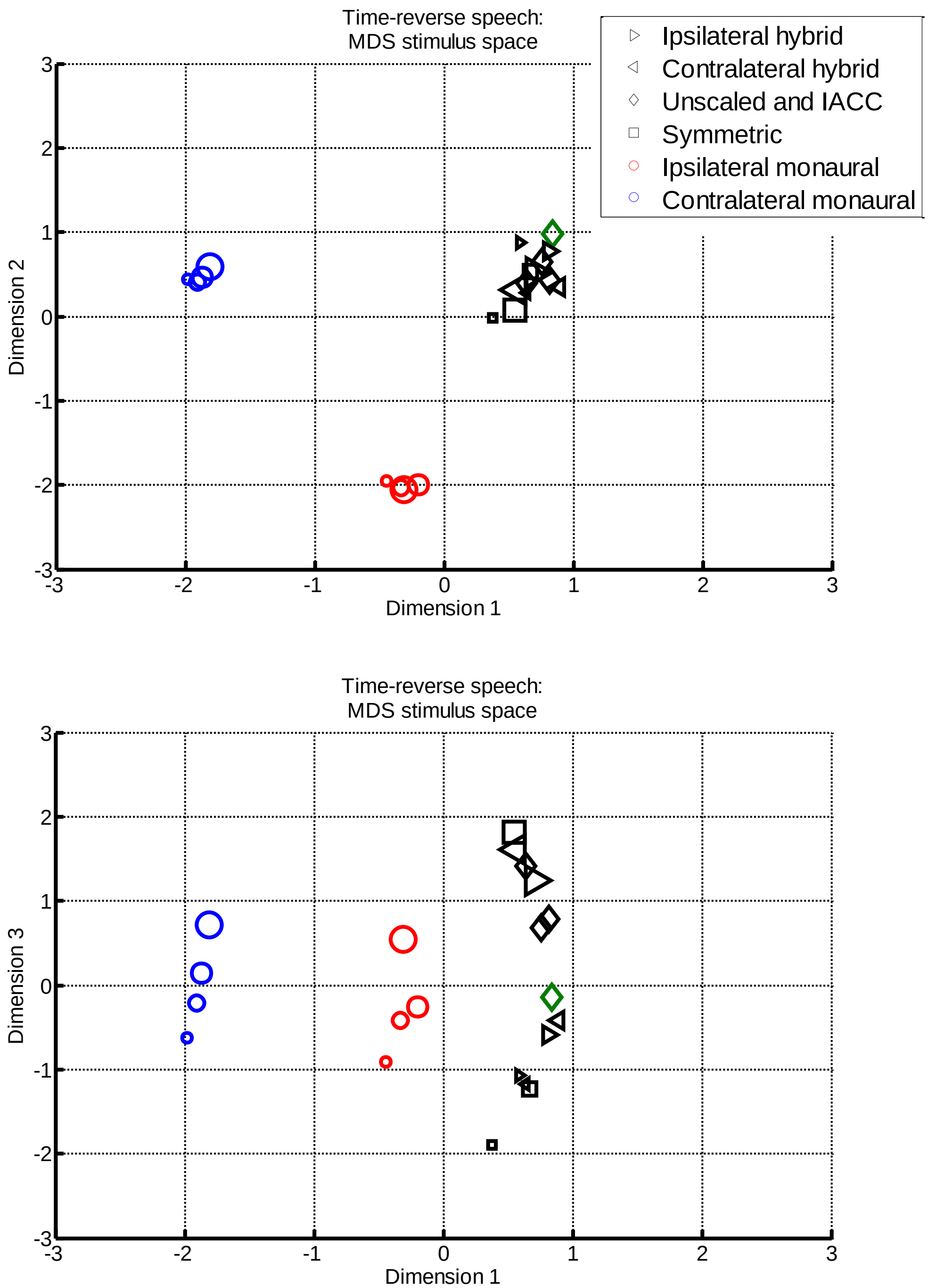


Figure 11. Stimulus space for the time-reverse stimuli. Size, color, and shape represent the same parameters as in Figure 10.

Regression analysis results

Regression analyses were used to interpret the dimensions of the stimulus space. Kruskal and Wish (1978) recommend using multiple linear regression with each dimension of the stimulus space as an independent variable and the physical or perceptual parameter of interest as a dependent variable. The contribution of each dimension of the stimulus space can then be analyzed while controlling for the others. The three components arrived at in the PCA section were used as dependent variables in all analyses: reverberation level, monaural reverberation measurements, and broadband right ear power.

Time forward speech

A multiple linear regression using the dimensions of the solution space to predict reverberation level explained a significant proportion of the variance ($F(3,17) = 47.3751$, $R^2 = 0.8932$, $p < 0.001$). Dimension 1 was not a significant predictor of reverberation level ($\beta = 0.0481$, $t(17) = 0.2107$, $p = 0.8356$). Dimension 2 was a significant predictor of reverberation level ($\beta = -2.6842$, $t(17) = -11.8833$, $p < 0.001$). Dimension 3 was not a significant predictor of reverberation level ($\beta = 0.2940$, $t(17) = 1.2821$, $p = 0.2170$). A projection of the space showing dimension 2 predicting reverberation level is plotted in Figure 12.

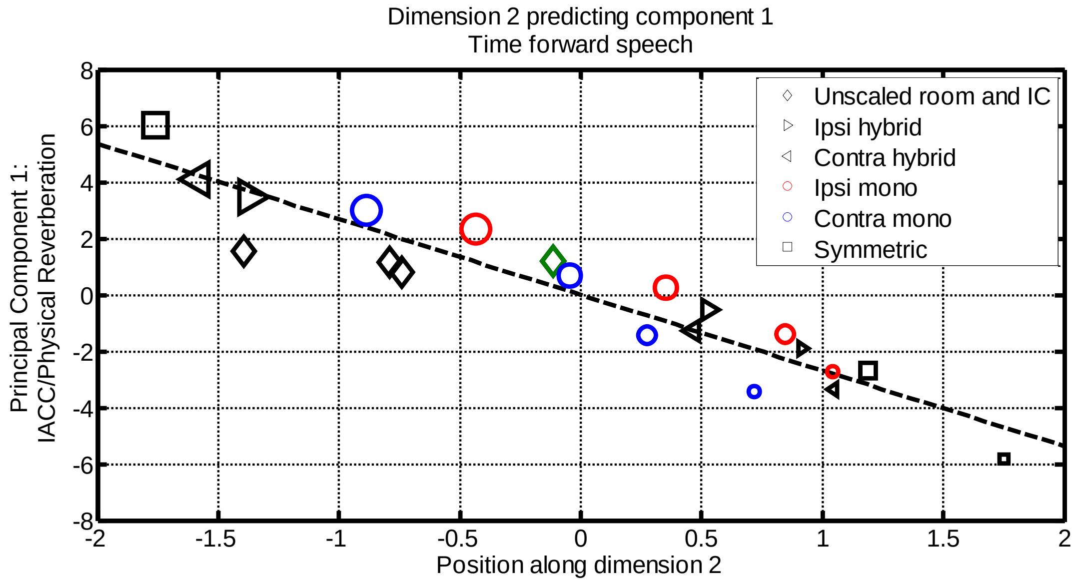


Figure 12. Position along dimension 2 as a predictor of reverberation level. Symbols use the same parameters as Figure 10. The green diamond represents the unaltered stimulus (the standard from the motivating experiment in Chapter 2). Note that the ipsilateral hybrid and contralateral hybrid stimuli are nearer to the unaltered stimulus relative to the symmetric listening conditions of the same size. This indicates that listeners perceive these hybrid listening conditions as more similar to the standard than the symmetric listening conditions. Dotted line is the least squares linear regression line. $R^2 = 0.8932$, $p < 0.001$.

A multiple linear regression using the solution space to predict monaural reverberation measurements did not explain a significant proportion of the variance ($F(3,17) = 0.0977$, $R^2 = 0.0170$, $p = 0.9602$). None of the dimensions were significant predictors of monaural reverberation measurements (Dimension 1: $\beta = -0.0216$, $t(17) = -0.0481$, $p = 0.9622$; Dimension 2: $\beta = 0.1786$, $t(17) = 0.4009$, $p = 0.6935$; Dimension 3: $\beta = -0.1852$, $t(17) = -0.4093$, $p = 0.6874$).

A multiple linear regression using the solution space to predict broadband right ear power explained a significant proportion of the variance ($F(3,17) = 138.9620$, $R^2 = 0.9608$, $p < 0.001$). Dimension 1 was a significant predictor of broadband right ear power ($\beta = 0.9623$, $t(17) = 19.7146$, $p < 0.001$). Dimension 2 was not a significant predictor of broadband right ear power ($\beta = 0.0874$, $t(17) = 1.8082$, $p = 0.0883$). Dimension 3 was a significant predictor of broadband right ear power ($\beta = 0.4237$, $t(17) = 8.6312$, $p < 0.001$). Projections of the space showing dimension 1 predicting broadband right ear power and dimension 1 predicting component 1 are plotted in Figure 13.

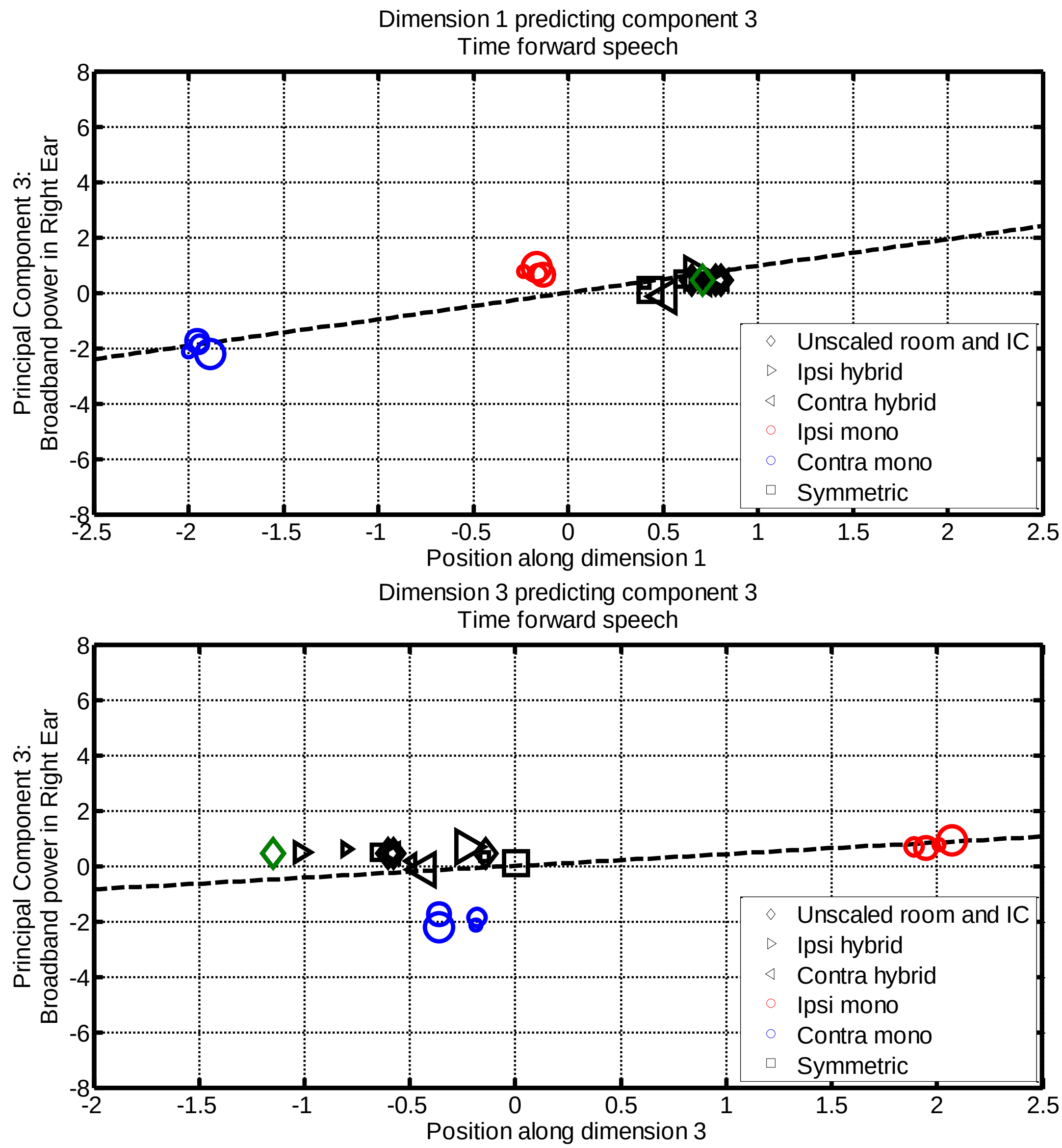


Figure 13. Position along dimensions 1 and 3 as predictors of broadband right ear power. Symbols use the same parameters as Figure 10. Dotted line is the least squares linear regression line. $R^2 = 0.9608$, $p < 0.001$.

Time-reverse speech

A multiple linear regression using the solution space to predict reverberation level explained a significant proportion of the variance ($F(3,17) = 40.579$, $R^2 = 0.8775$, $p < 0.001$). Dimension 1 was not a significant predictor of reverberation level ($\beta = 0.1363$, $t(17) = 0.5606$, $p = 0.5824$). Dimension 2 was not a significant predictor of reverberation level ($\beta = -0.0772$, $t(17) = -0.3167$, $p = 0.7554$). Dimension 3 was a significant predictor of reverberation level ($\beta = 2.6352$, $t(17) = 10.936$, $p < 0.001$). A projection of the space showing dimension 3 predicting reverberation level is plotted in Figure 14.

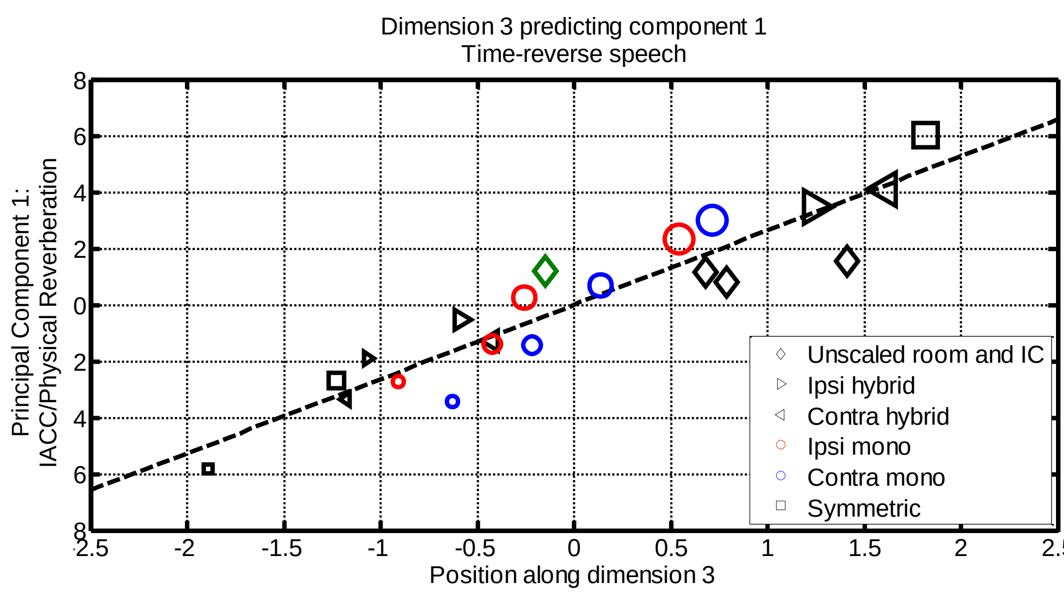


Figure 14. Position along dimension 3 as a predictor of principal component 1. Symbols use the same parameters as Figure 10. Note the same compression of the hybrid listening conditions observed in Figure 12. Dotted line is the least squares linear regression line. $R^2 = 0.8775$, $p < 0.001$.

A multiple linear regression using the solution space to predict monaural reverberation measurements did not explain a significant proportion of the variance ($F(3,17) = 0.0856$, $R^2 = 0.0149$, $p = 0.9670$). None of the dimensions were significant predictors of monaural reverberation measurements (Dimension 1: $\beta = -0.0160$, $t(17) = -0.0357$, $p = 0.9719$; Dimension 2: $\beta = 0.2940$, $t(17) = 1.2821$, $p = 0.7357$; Dimension 3: $\beta = -0.1789$, $t(17) = -0.4027$, $p = 0.6921$).

A multiple linear regression using the solution space to predict broadband right ear power explained a significant proportion of the variance ($F(3,17) = 105.98$, $R^2 = 0.9492$, $p < 0.001$). Dimension 1 was a significant predictor of broadband right ear power ($\beta = 0.9238$, $t(17) = 16.706$, $p < 0.001$). Dimension 2 was a significant predictor of broadband right ear power ($\beta = -0.4836$, $t(17) = -8.7207$, $p < 0.001$). Dimension 3 was not a significant predictor of broadband right ear power ($\beta = -0.0787$, $t(17) = -1.4349$, $p = 0.1695$). Projections of the space showing dimensions 1 and 2 predicting component 1 is plotted in Figure 15.

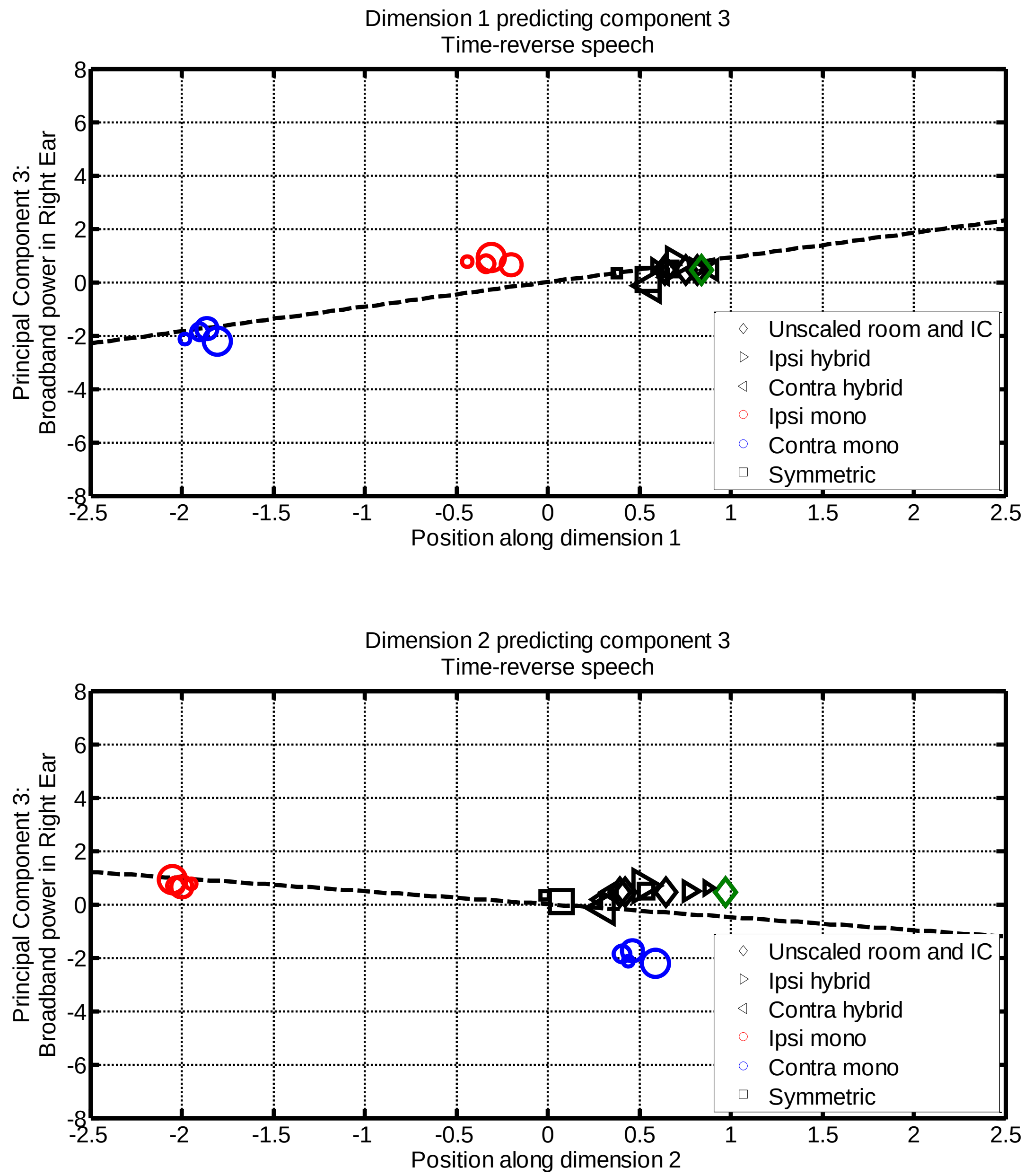


Figure 15. Position along dimensions 1 and 2 as predictors of broadband right ear power. Symbols use the same parameters as Figure 10. Dotted line is the least squares linear regression line. $R^2 = 0.9492$, $p < 0.001$.

Comparing time forward and time-reversed results

In a previous section, YHI and YNH listeners were pooled to test if the groups were statistically significantly different. While it would be desirable to do a similar test for the time forward and time-reverse data, INDSCAL is not capable of performing such a test. Since each stimulus must be compared with each other stimulus to make a complete dissimilarity matrix, listeners would have had to judge time forward and time-reverse stimuli against one another. In the present experiment, time forward and time-reverse stimuli were blocked separately and tested in different sessions. This means we cannot use the same technique to test for a difference between time forward and time-reverse stimuli.

Overall, YHI and YNH listeners are sensitive to listening condition and reverberation level. The order of these dimensions is important, however. For INDSCAL, dimension 1 explains the most variance in the rating data, dimension 2 the second most, and so on (Kruskal & Wish, 1978; Schiffman et al., 1981). This means that the dimensions can be interpreted as the cues listeners will use to determine their similarity ratings. Listeners will preferentially use dimension 1 first, dimension 2 second, and so on (Kruskal & Wish, 1978; Schiffman et al., 1981). We can conclude that listeners perform slightly differently for the time forward and time-reversed speech.

In the time forward speech condition, listeners will first separate the stimuli by listening condition (dimension 1), then will use information about physical amount of reverberation and IACC (dimension 2), then will make another decision about grouping (dimension 3). For the time-reverse speech condition, listeners separate the stimuli by listening condition first (dimensions 1 and 2), then use physical amount of reverberation and IACC last (dimension 3). There is no analysis to the author's knowledge to see if this change is meaningful or not, but it is of qualitative interest—listeners slightly change their strategy when speech is time-reversed relative to when it is time forward.

Comparing of YNH and YHI reverberation level weights

In a previous section, ANAVA was used to test whether YNH and YHI listeners used different cues in their similarity judgments. This test served to examine whether these groups differed in all three dimensions at once. After analyzing the stimulus space, there is evidence that dimension 2 of the time forward space and dimension 3 of the time-reverse space are both related to reverberation level. Do YNH and YHI listeners differ in how much weight they place on reverberation level? Independent samples *t*-tests were conducted to answer this question.

For the time forward space, we tested the difference between YNH and YHI weights on dimension 2. There was no significant difference between mean YNH and YHI weights ($t(24) = -1.9066$, $p = 0.0686$). As described in the subject space analysis above, listeners ZGT and ZGZ were omitted due to being outliers along all three dimensions. With these listeners excluded, the differences between groups was smaller and still not significantly different from 0 ($t(22) = -1.3360$, $p = 0.1952$). Differences in weight along dimension 3 of the time-reverse space were not significantly different between the YNH and YHI listeners ($t(21) = -0.9292$, $p = 0.3633$). This indicates that the nearly-significant differences observed in the subject space analyses above were not due to reverberation level weighting differences between the groups.

Discussion

We tested young normal hearing (YNH) and young hearing-impaired (YHI) perception of reverberant sounds covering a wide range of listening conditions (monaural, binaural), reverberation amounts (from -12 dB to +6 dB), and IACC values (a range between 0 and 1). Our results showed that across this range of stimuli, YNH and YHI listeners use the same cues to group reverberant stimuli. Bearing in mind that some of the peripheral effects of hearing loss were controlled with the NAL-R gain rule, suprathreshold reverberation perception may not be so different between YNH and YHI listeners. The similarity in suprathreshold reverberation is due in part to the conductive nature of some of the YHI losses in our sample. Listeners with conductive losses are more receptive to gain adjustments than the sensorineural population, so the NAL-R gain rule would restore audibility more easily. This could be part of the reason why there was no significant difference found between the YNH and YHI groups.

Based on the analysis relating reverberation level to dimension 2 for the time forward stimuli and dimension 3 for the time-reversed stimuli, we can conclude that YNH and YHI listeners are sensitive to reverberation level. This sensitivity is affected by listening condition. These results are qualitatively consistent with those from the motivating experiment (Chapter 2). In Chapter 2, we also observed changes to reverberation sensitivity depending on listening condition. A notable difference between the results explored in the motivating experiment (Chapter 2) and the results of the present experiment (Chapter 4) is that listeners here do not show the complete insensitivity to changes in reverberation level under hybrid listening conditions we saw in Chapter 2. In the present experiment we do observe less sensitivity to changes in reverberation level under hybrid listening conditions (See Figure 12). It is possible that an open-ended task like MDS gave listeners the opportunity to use other cues that they were specifically requested to ignore in the unidimensional motivating experiment (Chapter 2).

Our results also showed that, regardless of hearing status, listeners preferentially group stimuli by listening condition, separating the monaural and binaural stimuli. Results from the literature have shown that both NH and HI listeners are sensitive to differences between monaural and binaural stimuli (Whitmer, Seeber, & Akeroyd, 2012), and this is likely to be a salient cue. It makes sense that listeners would use a strong cue as a first pass to parse the stimulus space. Listeners placed a larger weight on the grouping dimensions for the time-reverse stimuli than they did for the time forward stimuli. A plausible reason for this increased reliance on listening condition is that the time-reverse stimuli are unfamiliar. Listeners are not used to extracting reverberation information from time-reverse stimuli which could cause the extracted reverberation information to be unreliable. When participants are given a choice between a strong, consistent cue (i.e., listening condition) and a weak, inconsistent one (i.e., extracted reverberation information from an unfamiliar stimulus), participants are more likely to place a higher weight on the strong, reliable cue (Ernst & Banks, 2002).

CHAPTER V

MODELING PERCEIVED AMOUNT OF REVERBERATION

Modeling

Part of the goal of this project was to better understand the mechanisms underlying reverberation perception in NH and HI listeners. To work toward this goal, a reverberation perception model based on van Dorp Schuitman, de Vries, and Lindau (2013) was implemented to explain the results from Chapters 2 and 4. The model as implemented here can be broken down into two subsections defined by van Dorp Schuitman et al.: a peripheral processor representing the ear canal, cochlea, and much of the brainstem processing; and a central processor that makes decisions about reverberation perception. The model as described in van Dorp Schuitman et al. (2013) includes a binaural processing step that was not included here. This step was excluded because it was unclear how it was implemented in the original paper. It also describes how to calculate perceptual aspects of reverberation not relevant to the experiments here. All modeling was done in MATLAB using custom software.

The peripheral processor

The first step of the model accounts for the processing done by the outer, middle, and inner ear. Much of this process is based on models described by Breebaart, van de Par, and Kohlrausch (2001) and Dau, Puschel, and Kohlrausch (1996). Per the recommendations of the authors, the stage modeling the outer ear was skipped because non-individualized HRTFs were used to spatialize the stimuli (Breebaart et al., 2001). To simulate cochlear filtering, the signal was passed through a fourth-order gammatone filterbank with 41 ERB filters centered on frequencies from 27 to 20,577 Hz. The action of the inner hair cells was modeled by passing each band through a half-wave rectifier, then filtered by a fifth-order low-pass filter with a cutoff frequency at 770 Hz. This simulated phase locking in the auditory nerve. The signal was half-wave rectified again before being logarithmically compressed to account for adaptation of the auditory nerve (Dau et al., 1996). This second half-wave rectification prevented imaginary numbers due to the logarithmic compression. This entire procedure was performed for the left and right ears independently.

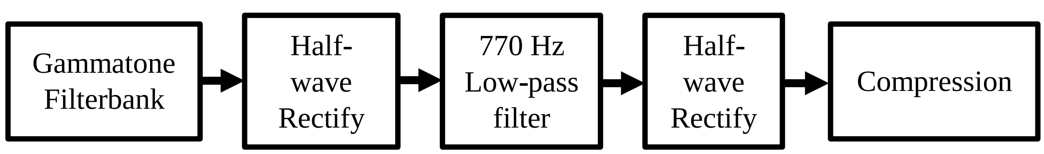


Figure 16. Peripheral processor of the auditory model. This process was completed for both the left and right ears independently. The output of the compressor for the left and right ears served as the input to the central processor

The central processor

The central processor was implemented based on the method used in van Dorp Schuitman et al. (2013). The inputs to the central processor were the outputs of the left and right peripheral processors. A reverberation extraction algorithm was implemented to separate the reverberant sound from the direct sound. The first step of the extraction algorithm calculated average level within each frequency band. This average level was used as a threshold to separate direct from reverberant sound. The second step segregated the signal within frequency band into reverberant and non-reverberant energy. Any time portions of the signal that fell below the average level within a band for 10 ms were considered reverberant energy. Other time portions were discarded.

Once reverberation was extracted, perceived reverberation within a frequency channel was calculated by squaring the reverberant time portions of the signal in both ears for that frequency channel, summing them, then taking the square root of the result. An overall prediction of perceived reverberation was calculated by averaging across frequency channel and time. The resulting estimate of perceived reverberation is an arbitrary model unit (MU).

Modeling and past results

We used data from the motivating experiments (Chapter 2) to validate the model. Five forms of the model were tested: 1) A modified version of van Dorp Schuitman's model as described above, 2) Model without the reverberation extraction algorithm and reverberation only, 3) Model without the reverberation extraction algorithm and the full signal, 4) A band-limited form of the model, and 5) a model with no peripheral processor.

These forms of the model were tested to determine which steps in the original model were most important for predicting the data in the motivating experiments (Chapter 2). We tested the predictive power of each of the forms of the model using general linear models (GLM) with model output (MU) as an independent variable, observed perceived reverberation from Chapter 2 data as a dependent variable, and the five listening conditions (ipsilateral hybrid, contralateral hybrid, ipsilateral monaural, contralateral monaural, and symmetric) as factors.

Modified reverberation extraction model

Using the modified reverberation model as described above, we found that the GLM explained a significant proportion of the variance ($F(9,58) = 30.9369$, $R^2 = 0.8276$, $p < 0.001$). There was a significant interaction between listening condition and predicted perceived reverberation indicating that at least one of the slopes was different from the others ($F(4,58) = 11.4980$, $p < 0.001$). Bonferroni post-hoc tests were conducted to see which of the slopes differed from the others. Since 10 tests were run, the adjusted α level was set to 0.005. The following slopes were significantly different: ipsilateral hybrid and contralateral monaural ($t(27) = -5.2652$, $p < 0.001$), ipsilateral monaural and contralateral monaural ($t(14) = -3.4282$, $p = 0.0041$), ipsilateral hybrid and symmetric ($t(40) = -5.7910$, $p < 0.001$), ipsilateral monaural and symmetric ($t(27) = -4.1970$, $p < 0.001$). The data are plotted in Figure 17.

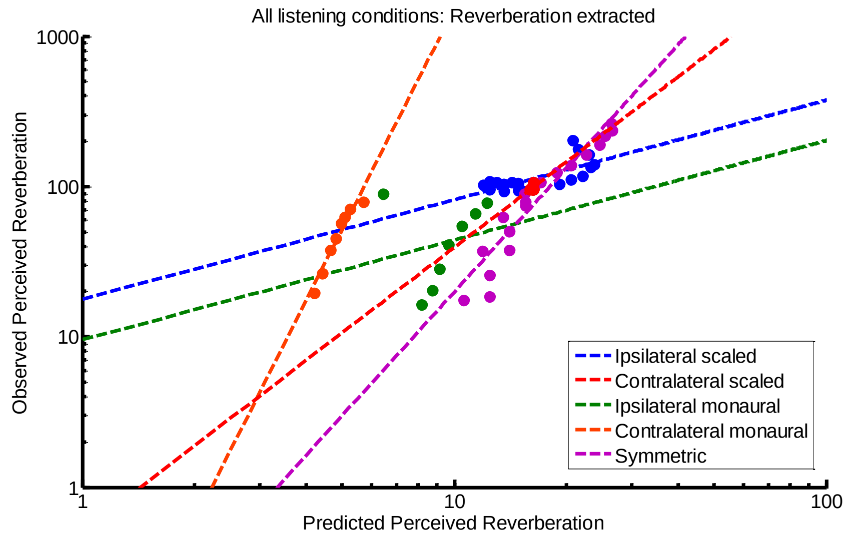


Figure 17. Using the modified reverberation extraction model to predict results of Chapter 2. This version of the model extracted reverberant energy from the signal using a level-based approach (van Dorp Schuitman et al., 2013). The fit explains a significant proportion of the variance (R^2 = 0.8276, p < 0.001), however there are some inconsistently fitted points.

While the GLM explained a significant proportion of the variance, it did not predict the results of the ipsilateral monaural condition very well. It is possible the extraction method was not appropriate for this set of stimuli. To test whether the extraction method was working properly, we ran the signals through the model with only reverberant energy. This effectively extracted the reverberation for the model. The reverb only signal was achieved by removing the direct sound of the impulse response (the first 2.5 ms based on Wightman & Kistler, 1989, similar method to how the stimuli were generated in Experiment 1/Chapter 3) before convolution with the same speech signals used in the motivating experiments (Chapter 2). Since the reverberation was already extracted, the reverberation extraction step of the model was bypassed before running the reverberation only stimuli.

Broadband reverberation only model

The same general linear model was used as above except predicted perceived reverberation was calculated using only the reverberant energy. The overall fit explained a significant proportion of the variance ($F(9,58) = 37.4885$, $R^2 = 0.8533$, $p < 0.001$). There was a significant interaction between listening condition and predicted perceived reverberation indicating that at least one of the slopes was different from the others ($F(4,58) = 7.0457$, $p < 0.001$). Bonferroni post-hoc tests were conducted to see which of the slopes differed from the others. The same adjustments to alpha were made as for the last model. The following slopes were significantly different: ipsilateral hybrid and ipsilateral monaural ($t(27) = -5.0635$, $p < 0.001$), and ipsilateral hybrid and symmetric ($t(40) = -3.9316$, $p < 0.001$). The data are plotted in Figure 18.

These results indicate that summation of reverberant energy is sufficient to predict the data from the motivating experiments (Chapter 2) well. The ipsilateral monaural data are much more linearly related to one another, indicating that this approach is stronger than the reverberation extraction method.

Full signal model

While total reverberant power predicted the results well, it is possible that total power of the signal would as well. To test whether overall signal power (i.e., direct + reverberant energy) was predictive of perceived reverberation, the full signals were put through the model without the reverberation extraction step.

The same general linear model was used as above except predicted perceived reverberation was calculated using total summed power. The overall fit explained a significant proportion of the variance ($F(9,58) = 20.4225$, $R^2 = 0.7601$, $p < 0.001$). There was a significant interaction between listening condition and predicted perceived reverberation indicating that at least one of the slopes was different from the others ($F(4,58) = 3.5344$, $p = 0.0120$). Bonferroni post-hoc tests were conducted to see which of the slopes differed from the others. The same adjustments to alpha were made as for the last model. The ipsilateral monaural and symmetric listening conditions were significantly different from one another ($t(27) = -4.2947$, $p < 0.001$). The data are plotted in Figure 19.

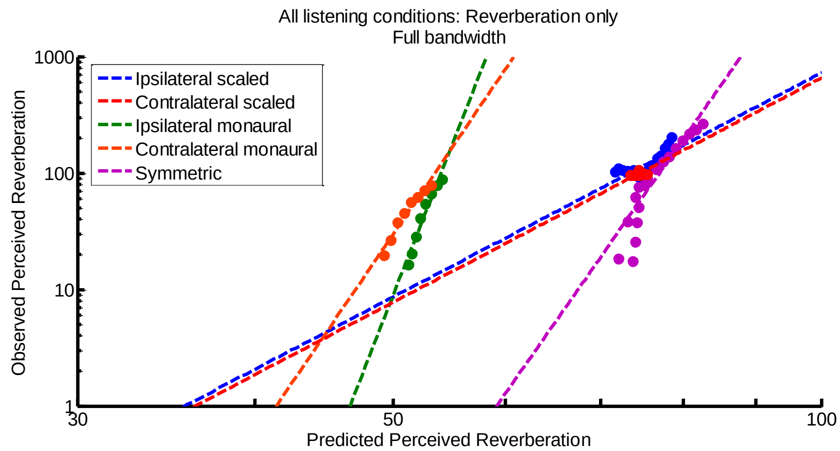


Figure 18. Using reverberation only as input to the model to predict results of the motivating experiments (Chapter 2). The model was fed signals that only contained reverberant energy, then the extraction step was skipped. The fit explains a significant proportion of the variance (R^2 = 0.8533, p < 0.001), and is much more consistent than the fit in Figure 17.

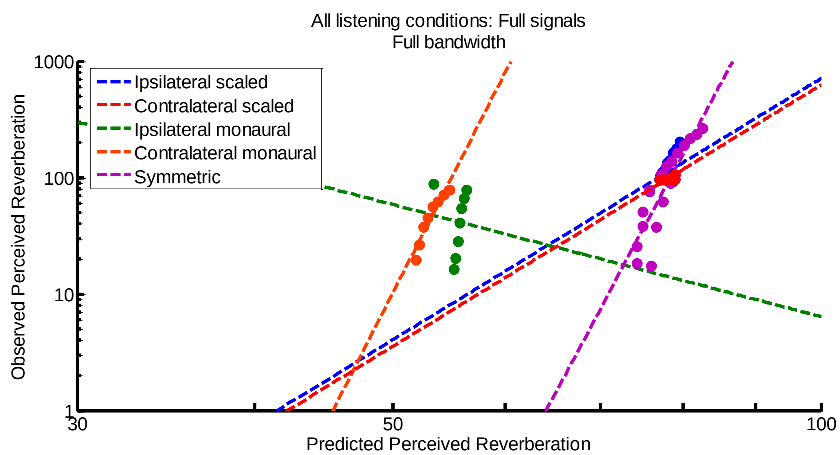


Figure 19. Using full signals as inputs to the model to predict results of Chapter 2. The model was fed full signals (i.e., those that contained direct and reverberant energy), then the extraction step was skipped. This was to test whether a simple summation of power at the two ears explained the results. While the fit explains a significant proportion of the

variance (R^2 = 0.7601, p < 0.001), it is similarly inconsistent to the fit in Figure 17. It is also qualitatively worse than the fit in Figure 18.

These results look like those seen for the modified reverberation extraction model. While this model explained a significant proportion of the variance, the ipsilateral monaural data are not well predicted by the model. The amount of variance explained in this model is also qualitatively lower than the other two (Extraction: R^2 = 0.8276; Reverb only: R^2 = 0.8533; Full signal: R^2 = 0.7601).

The previous three models demonstrate that the summation of reverberant energy across the two ears is predictive of the results in the motivating experiments (Chapter 2). Is there a narrow band of frequencies sufficient to predict these results? Low-frequency residual hearing is common in HI listeners (including a few of those sampled here), so low-frequency reverberant information is likely to be used by this population for judgments of reverberation. To see if low frequencies are sufficient to model the results of the motivating experiments (Chapter 2), only the outputs of auditory filters with center frequencies at or below 1025 Hz were used to predict perceived amount of reverberation.

Narrowband reverberation only model

Using only the outputs of the auditory filters centered on frequencies at 1025 Hz and below, the overall fit explained a significant proportion of the variance ($F(9,58)$ = 55.705, R^2 = 0.8963, p < 0.001). There was a significant interaction between listening condition and predicted perceived reverberation indicating that at least one of the slopes was different from the others ($F(4,58)$ = 6.74, p < 0.001). Bonferroni post-hoc tests were conducted to see which of the slopes differed from the others. The same adjustments to alpha were made as for the last model. The following groups had significantly different slopes: ipsilateral hybrid and ipsilateral monaural ($t(27) = -5.7291$, $p < 0.001$), contralateral hybrid and ipsilateral monaural ($t(16) = -3.3265$, $p = 0.004$), ipsilateral monaural and contralateral monaural ($t(14) = -4.1649$, $p = 0.001$), and ipsilateral monaural and symmetric ($t(27) = -7.0708$, $p < 0.001$). The data are plotted in Figure 20.

Model without peripheral processing

Finally, a model with no peripheral processing was run to test the effect of reverberation summation alone. This model explained a significant proportion of the variance ($F(9,58) = 19.1271$, $R^2 = 0.7480$, $p < 0.001$). There was a significant interaction between listening condition and predicted perceived reverberation indicating that at least one of the slopes was different from the others ($F(4,58) = 14.90$, $p < 0.001$). Bonferroni post-hoc tests were conducted to see which of the slopes differed from the others. The same adjustments to alpha were made as for the last model. The following groups had significantly different slopes: ipsilateral hybrid and symmetric ($t(40) = -4.3992$, $p < 0.001$) and contralateral hybrid and symmetric ($t(29) = -4.1546$, $p < 0.001$). The data are plotted in Figure 21.

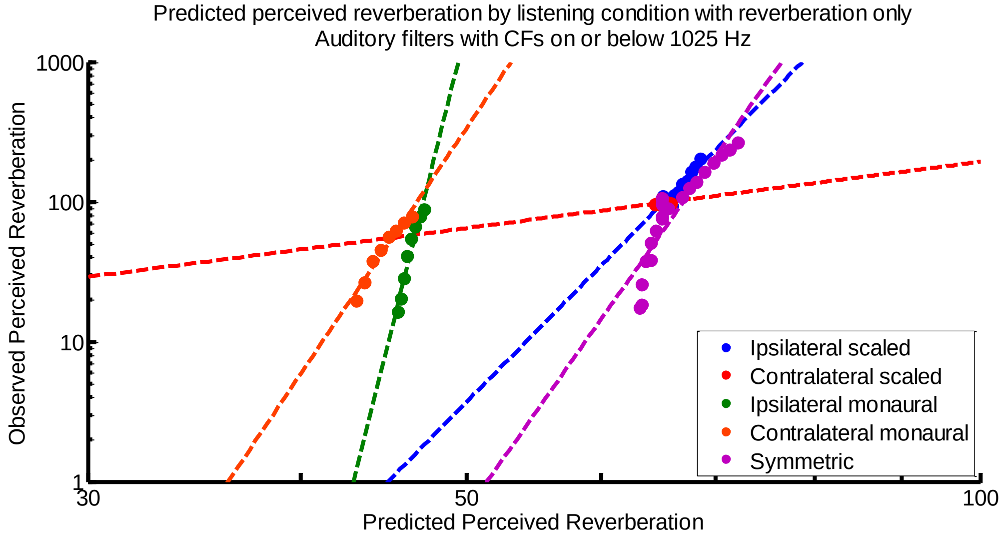


Figure 20. Using reverberation only from auditory filters centered on 1025 Hz or lower to predict results of the motivating experiments (Chapter 2). This fit explains a significant proportion of the variance ($R^2 = 0.8963$, $p < 0.001$).

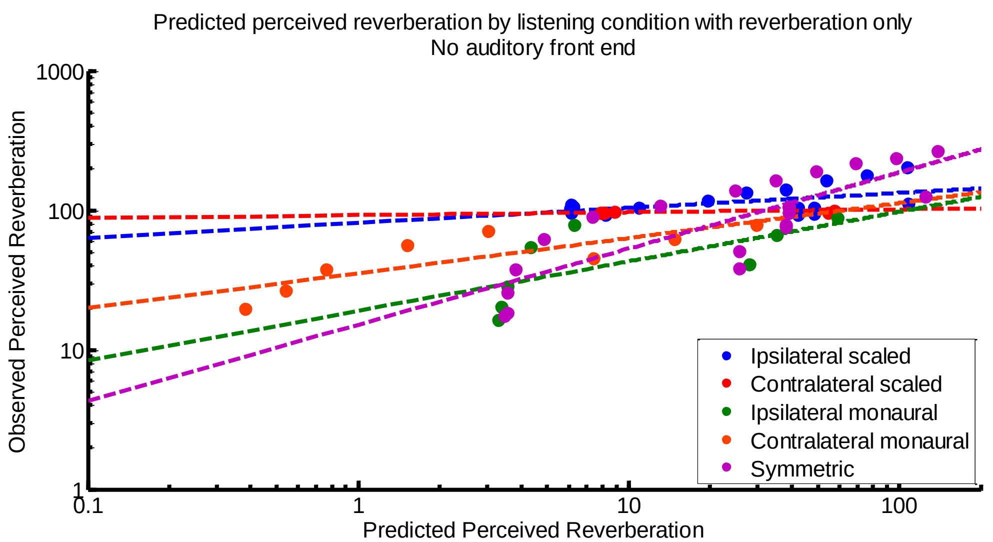


Figure 21. Using reverberation only and no auditory periphery to predict the results of the motivating experiments (Chapter 2). This fit explains a significant proportion of the variance ($R^2 = 0.7480$, $p < 0.001$).

We can now compare the five modeling approaches to one another (see Table 7 for a summary). The reverberation extraction and full signal models explained significant proportions of the variance (R^2 = 0.8276 and R^2 = 0.7601, respectively), but produced results that are not well-explained by the general linear model used here. The variability in the ipsilateral monaural listening condition in these two models is large. For this dataset and the stimuli examined here, neither the reverberation extraction nor full bandwidth summation modeling approaches are appropriate.

Table 7.

Summary of models and fits

Model	$\boldsymbol{R^2}$	**Qualitative Fit Observations**
Modified reverberation extraction model (van Dorp Schuitman et al., 2013)	0.8276	Does not fit ipsilateral monaural well
Broadband reverberation only	0.8533	Fits most listening conditions well, except some symmetric stimuli
Full signal	0.7601	Does not fit ipsilateral monaural well
Low- frequency reverberation only	0.8963	Fits most listening conditions well
Broadband reverberation only. No auditory periphery	0.7480	Inconsistent fit, nonlinear

The broadband reverberation only model fits the observed data well. In this model, the reverberation is extracted for the model, so the important part is the summation across frequency bands. That the model predicts the data accurately indicates that a summation of reverberant power at the two ears is important for judgments of perceived amount of reverberation. Power summation has been shown to be important for overall loudness judgments (B. C. J. Moore, Glasberg, & Baer, 1997), and for loudness judgments as a function of azimuth (Sivonen, 2007). Here, the summation of reverberant energy alone at the ears is important for judgments of perceived amount of reverberation.

The auditory periphery contributes to controlling the variability of the data somewhat. The sole difference between the broadband reverberation only model and the no auditory periphery model is the exclusion of the auditory periphery in the latter. This means that the nonlinear action of the auditory periphery is contributing some proportion of explained variance. Here, it is about 10%.

Of the five methods of modelling used above, the low-frequency only model explained the most variance and was qualitatively the best fit. This provides evidence that while broadband summation of reverberant energy predicts the data well, low-frequency information is sufficient. This low-frequency information could be used by hearing-impaired listeners to make judgments of reverberation perception. For this reason, and because the low-frequency version of the model with the auditory periphery was the best predictor of the results of the motivating experiments (Chapter 2), this version of the model was used to analyze the results of experiment 2 (Chapter 4)

Modeling results and Experiment 2 (Chapter 4)

We continued to use the low-frequency, reverberation only version of the model to further analyze the results of Experiment 2 (Chapter 4). The time forward and time-reverse speech stimuli were resynthesized to include only the early reflections and late reverberation as in the section above. The output of the model was predicted perceived reverberation amount in arbitrary units. The effect of hearing loss was not taken into account in the model.

Recall the results of Experiment 2. For the time forward speech stimuli, dimension 2 was related to physical amount of reverberation removed from the signal and IACC. Since dimensions 1 and 3 served to group the stimuli into binaural and monaural stimuli, dimension 2 is the most likely dimension to be related to perceived reverberation. Dimension 2 was used to predict the model output. The stimuli were put into two groups: monaural and binaural stimuli. For the time forward stimuli, a GLM was run with dimension 2 as an independent variable, model output at and below 1025 Hz (MU) as a dependent variable and listening condition (monaural and binaural) as a factor. For the time-reverse stimuli, a GLM was run with dimension 3 as an independent variable, model output at and below 1025 Hz (MU) as a dependent variable and listening condition (monaural and binaural) as a factor.

The time forward data were analyzed first. The GLM explained a significant proportion of the variance ($F(3,17) = 1259.4$, $R^2 = 0.9955$, $p < 0.001$). There was not a significant interaction between listening condition and predicted perceived reverberation indicating that the slopes were not different from one another ($F(1,17) = 0.1$, $p = 0.7564$). The model was rerun without the interaction term. The new model also explained a significant proportion of the variance ($F(2,18) = 1988.5$, $R^2 = 0.9955$, $p < 0.001$). Because it was more parsimonious, this model was selected as being more appropriate. A post-hoc test showed that the intercepts were significantly different ($t(19) = 22.82$, $p < 0.001$). This means that at a given position in the stimulus space for the time forward stimuli, listeners are predicted to perceive more reverberation in the binaural stimuli than in the monaural stimuli.

A similar analysis was run for the time-reversed speech. The GLM explained a significant proportion of the variance ($F(3,17) = 1540.6$, $R^2 = 0.9963$, $p < 0.001$). There was not a significant interaction between listening condition and predicted perceived reverberation indicating that the slopes were not different from one another ($F(1,17) = 1.14$, $p = 0.30$). The model was rerun without the interaction term. The new model also explained a significant proportion of the variance ($F(2,18) = 2292$, $R^2 = 0.9961$, $p < 0.001$). A post-hoc test showed that the intercepts were significantly different ($t(19) = 26.256$, $p < 0.001$). This means that at a given position in the stimulus space for the time-reversed stimuli, listeners should be perceiving more reverberation in the binaural stimuli than in the monaural stimuli. See Figure 22 for the time forward speech and Figure 23 for the time-reversed speech.

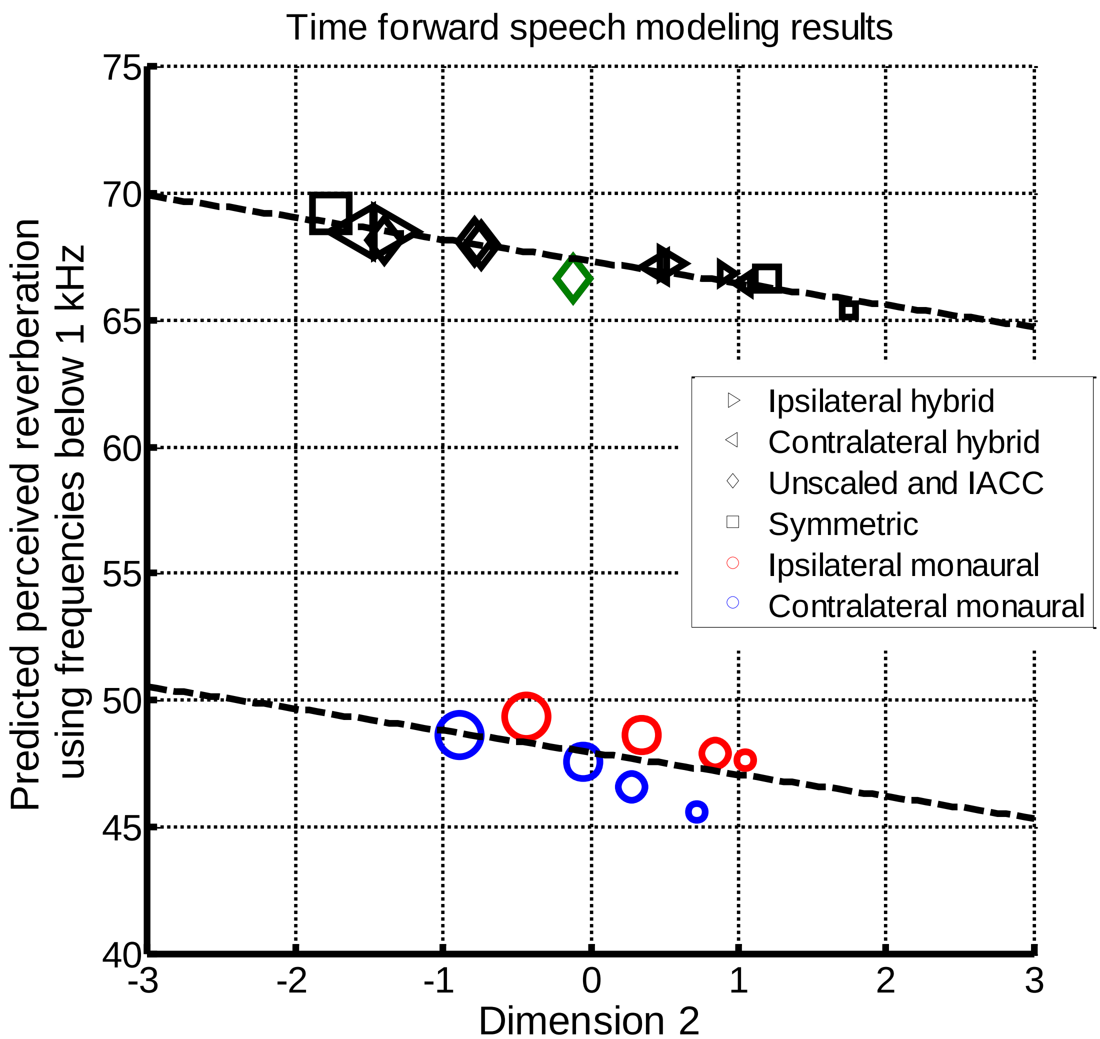


Figure 22. Dimension 2 of the time forward speech stimulus space predicting modeled perceived reverberation using only auditory filters centered at or below 1025 Hz. Size represents reverberation level. From smallest to largest: -12 dB, -6 dB, 0 dB, +6 dB. Color: red, ipsilateral monaural; blue, contralateral monaural; black, binaural. Shape: listening condition. See legend. Lines represent least-squares regression lines (R^2 = 0.9955, $p < 0.001$). Dashed lines represent best fits for the binaural and monaural data.

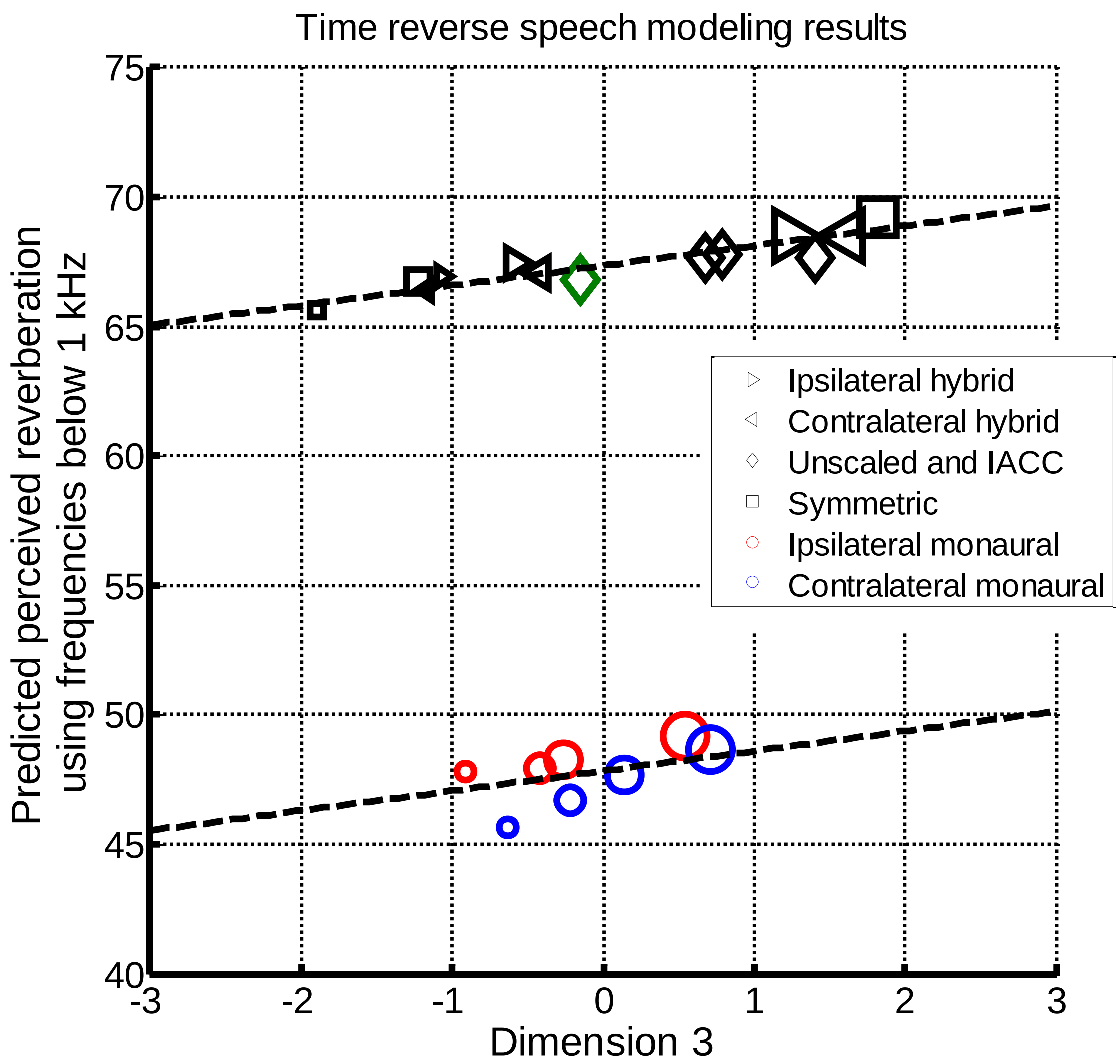


Figure 23. Dimension 3 of the time-reverse speech stimulus space predicting modeled perceived reverberation using only auditory filters centered at or below 1025 Hz. Size represents reverberation level. From smallest to largest: -12 dB, -6 dB, 0 dB, +6 dB. Color: red, ipsilateral monaural; blue, contralateral monaural; black, binaural. Shape: listening condition. See legend. Lines represent least-squares regression lines (R^2 = 0.9963, $p < 0.001$). Dashed lines represent best fits for the binaural and monaural data.

Conclusions

The auditory model examined here has a straightforward front end representing the auditory periphery and a reverberation processor that sums reverberant power at the two ears. This model predicts the results of Chapter 2 (the motivating experiment) well, and accounts for 85% of the variance in those data. The summation of reverberant power at the two ears contributes in a large way to the perceived amount of reverberation, accounting for 75% of the variance on its own. This summation is an important factor in how much reverberation listeners perceive. In addition to the summation, the auditory front end contributes the remaining 10% of the variance observed here.

The low-frequency reverberation only version of the model almost perfectly predicts the results of Chapter 4 (Experiment 2) when listening condition is accounted for ($R^2 > 0.99$). This indicates that low frequencies are sufficient for YNH and YHI listeners to make judgments about the similarity of reverberant stimuli.

CHAPTER VI

GENERAL DISCUSSION

This project examined reverberation perception using different techniques across different listeners at separate times. That the conclusions from each of the experiments in Chapters 2, 3, and 4 are in general agreement provides strong across-experiment evidence for the results.

The experiment in Chapter 2 demonstrated that listeners judge perceived amount of reverberation differently based on listening condition. In symmetric and monaural listening conditions, listeners report hearing changes in the amount of perceived reverberation as a function of physical reverberation removed from the signal. In hybrid listening conditions, listeners do not report hearing any change in the amount of reverberation as a function of the amount of physical reverberation removed from the signal. This is true even when there is no reverberation present in one ear. Though the hybrid listening conditions are unlikely to occur in a natural listening environment, these findings provide insights into how the auditory system must process reverberant sound.

The experiment in Chapter 3 demonstrated that listeners can effectively judge the perceived amount of reverberation in the absence of the reverberant tail of the signal. This finding indicates that the reverberant tail is not necessary for making judgments of perceived reverberation—that is, the ongoing reverberation is sufficient. While the ecological relevance of this experiment could be questioned (real-world listening rarely results in absent reverberant tails), it provides insight into the nature of reverberant sound processing strategies in the auditory system.

The experiment in Chapter 4 extends the experiment from Chapter 2 to both YNH and YHI listeners, using different (MDS) measurement techniques. Results demonstrated that YNH and YHI listeners did not differ in their judgments of reverberant sound similarity. Results were also found to be qualitatively similar to those reported in Chapter 2, which provides cross validation of the perceptual measurement techniques used in both experiments.

The modeling work in Chapter 5 well predicts the data reported in both Chapters 2 (scaling) and 4 (MDS). It also provides insight into the nature of reverberant sound processing strategies in the auditory system. As was demonstrated in Chapter 5, the summation of reverberant energy at the two ears is a primary factor in the amount of reverberation listeners perceive. Recall that the full signal model that did not explicitly separate reverberant from direct sound energy provided unsatisfactory results. The success of the reverberation-only model and the unsatisfactory fit of the full signal model are consistent with the hypothesis that the auditory system can successfully extract the reverberation from a sound in a room. The results of Chapter 3 further suggest that this can be done with only the ongoing reverberant signal—the tail is not necessary for reverberation extraction.

While the precise mechanism for reverberation extraction is unknown, the model in Chapter 5 sheds some light on what the mechanism may *not* be. The reverberation extraction mechanism described by van Dorp Schuitman et al. (2013) uses a level-based approach to separate the reverberant and direct energy. For the stimuli we tested in

Chapter 5, the level-based approach did not work well. While the fit explained a significant proportion of the variance, the fit was not appropriate for all listening conditions. Therefore, it seems unlikely that the level rule proposed by van Dorp Schuitman et al is used generally by the human auditory system. Further evidence against the level rule was found when the extraction mechanism was bypassed and reverberant energy was fed into the model on its own. This resulted in very accurate model predictions of the data presented in Chapters 2 and 4.

It should be noted that this level-based approach is not likely to fail in *all* environments. It would theoretically work better in environments where the ratio of direct-to-reverberant energy is relatively high. In these environments, the level differences between the direct and reverberant energy would be large and easy to detect. Level-based approaches may also work for stimuli that are highly impulsive (e.g., clicks, hand claps, etc.). These impulse stimuli in anechoic space are characterized by a high-amplitude onset followed by silence. In a reverberant environment, this silent interval will be "filled in" by reverberant energy caused by the impulse. Ongoing stimuli (i.e., speech and some types of music) are characterized by continuous—or near-continuous—sound even in anechoic environments. When subjected to reverberation, there is no way for a model to know whether the low amplitude points of the waveform are reverberant energy or simply a low-amplitude time-portion of the waveform. As a result, reverberant speech and certain types of music may not be suitable signals for level-based attempts to segregate direct and reverberant energy.

Although a level-based approach to segregating direct and reverberant energy doesn't seem to work for the stimuli tested in this dissertation, there is converging evidence that the auditory system can segregate direct and reverberant energy in a room to some degree. This claim is supposed in a study conducted by Traer and McDermott (2016). There were two conditions of interest in Traer and McDermott's study. Both conditions required listeners to perform an "oddball" 3-AFC task. In the source discrimination condition, listeners had to choose the source (amplitude-modulated noise with statistical properties similar to natural sounds, see McDermott, Wrobleski, and Oxenham (2011) for synthesis details) that was different from the other two. All three sources were convolved with impulse responses that had different direct-to-reverberant energy ratios (DR). Since DR is a cue to distance, this manipulation effectively placed the sources at different distances in a given room (Zahorik, 2002a, 2002b). To perform this task, listeners had to control for the impulse responses to detect the source that was different. According to Traer and McDermott, listeners did this by segregating the source from the impulse response. In a second condition, listeners had to choose the oddball impulse response among three different source signals. Two impulse responses were identical while the third had a different reverberation time. The length of the sources was jittered so listeners could not use absolute time judgments as a cue. Again, Traer and McDermott argued that listeners were separating the source from the impulse response to do this task.

In addition to Traer and McDermott's evidence for the segregation of direct and reverberant energy, there is indirect evidence in the concert hall acoustics literature that listeners do this–the field of concert hall acoustics agrees that listeners are capable of segregating direct and reverberant energy. This is implicit in the findings that perceptual attributes are tied to physical measurements based on relative energy in early and late portions of the room impulse response. The early and late time portions of a sound in a room are associated with auditory source width (*ASW*) and listener envelopment (*LEV*), respectively. *ASW* and *LEV* are distinct, separate perceptual phenomena (Beranek, 2004; Bradley & Soulodre, 1995; Hidaka, Beranek, & Okano, 1995; Kuttruff, 2000; Rumsey, 2002) related to measures that use nonoverlapping time portions of an impulse response. *ASW* is associated with the first 80 ms and *LEV* is associated with everything after the first 80 ms. If the auditory system is calculating *ASW* and *LEV* using solely the physical acoustics of a sound, then it must separate the source and the room to do this. While these perceptual measures were not examined here, the fact that *ASW* and *LEV* are defined as separate perceptual entities and do not overlap temporally in the physical impulse response suggests that listeners must be able to segregate the direct from the reverberant sound and then perform calculations on the reverberant sound.

Studies have found populations of neurons in the inferior colliculus (IC) to be sensitive to a wide variety of binaural properties including interaural coherence and timing cues (Day & Delgutte, 2016; Fitzpatrick, Kuwada, & Batra, 2002; Jiang, McAlpine, & Palmer, 1997; Kuwada & Yin, 1983; Kuwada, Yin, Syka, Buunen, & Wickesberg, 1984; Brian C. J. Moore, 2013; Palmer, Jiang, & McAlpine, 1999, 2000; Yin & Kuwada, 1983a, 1983b). Acoustically, the direct time portion and early reflections are very highly correlated whereas reverberation is diffuse and decorrelated. These binaural acoustic properties could be used by the auditory system to separate early from late energy. Evidence has also shown that some percentage of neurons in unanesthetized rabbit IC are resistant to reverberation—that is, their response properties are robust to the spectrotemporal smearing caused by reverberation (Kuwada, Bishop, & Kim, 2014).

Human listeners also perform better on psychophysical tasks in reverberation than would be predicted based on the acoustics alone (Zahorik et al., 2012; Zahorik et al., 2011), a fact that could be mediated by these reverberation resistant neural populations if present in human listeners. This population of reverberation resistant neurons, in conjunction with neurons sensitive to interaural coherence/timing cues, could in theory serve to separate reverberant and nonreverberant sound.

In addition to binaural cues like interaural coherence and timing, monaural cues like amplitude modulation could underlie a mechanism segregating direct from reverberant energy. Though there are more neurons in unanesthetized rabbit IC sensitive to binaural information than monaural information (Kuwada et al., 2014), it has been shown that there is a population of neurons in the same brainstem region sensitive to distance via monaural amplitude modulation (AM) in reverberant environments (Kim, Zahorik, Carney, Bishop, & Kuwada, 2015). Kim et al. showed that these neurons are insensitive to changes in distance when the noise is unmodulated or when the noise is modulated and in an anechoic environment, suggesting these neurons are selectively sensitive to AM noises in reverberant environments (2015). This neural population could underlie a mechanism that uses monaural AM in reverberant environments to segregate direct from reverberant energy.

If reverberant and nonreverberant energy is separated in this way, reverberation-resistant neurons could eventually feedforward to networks in the brain that care about semantic properties of speech, emotional aspects of music, and other centers responsible for judgments of the direct portion of a sound. Neural populations sensitive to reverberation could feedforward to networks in the brain responsible for making judgments of the room—*ASW*, *LEV*, etc. These populations of reverberation-resistant and reverberation-sensitive neurons would thus serve to separate the direct and reverberant sound from one another to be processed in parallel streams.

So far, the direct and reverberant energy segregation mechanism has been examined from a purely bottom-up point of view. There are almost certainly top-down effects that have not been measured here. A simple thought experiment demonstrates this. In everyday listening, people are rarely aware of the reverberant properties of the room in which they are standing; however, if they are asked to attend to the reverberant properties, they will likely become aware of coloration and reverberation in just about any room. The reader may wish to try this now. Anechoic chambers demonstrate this sub-attentional awareness of reverberation as well—it can be disorienting to stand in a room with absolutely no reverberation. The effects of attention, working memory, and other cognitive factors have yet to be explored with respect to reverberation perception, but they likely play a role. It is theoretically reasonable that listeners with better attentional control can better attend to reverberant properties of a sound in a room and can therefore better extract it from the ongoing signal. It follows therefore that listeners with better attentional control may perform better in speech understanding tasks in reverberation. These cognitive aspects of reverberation perception are promising future directions.

Regardless of the precise extraction mechanism and its properties, it is clear from the data and modeling described in this dissertation that once the reverberation is extracted, the summation of reverberant energy at the two ears in the low frequencies is instrumental in the perception of reverberation. This simple summation model strongly supports the results of the motivating experiments (Chapter 2) and predicts the results of

Experiment 2 (Chapter 4) with high accuracy. While the auditory periphery also plays an important role in the model, the summation of reverberant energy accounts for the majority of the variance in the results of the motivating experiments (Chapter 2).

Adding to this, the results of Experiment 2 (Chapter 4) suggest that both young normal hearing (YNH) and young hearing-impaired (YHI) listeners with mixed conductive/sensorineural losses have intact reverberation segregation abilities. These samples of listeners were closely matched on the cues they used to judge the similarity of a wide range of reverberant stimuli. Both groups of listeners used listening condition and the amount of reverberation in the sound to judge how similar or different the stimuli were. To the author's knowledge, this is the first time this has been observed. It would be beneficial for future studies to cross validate these results with different samples of YHI listeners.

Assuming our sample of YHI listeners was representative of the population of YHI listeners, our results suggest aging may have a bigger impact on this segregation mechanism than hearing loss alone. Results from Whitmer et al. (2012) suggest that older listeners have difficulty discriminating interaural coherence compared to younger listeners. We also know that there is an effect of aging on hearing when understanding speech in reverberation (Nabelek & Robinson, 1982). This age effect on speech intelligibility in reverberation could be due to a deteriorating segregation mechanism. While this sounds plausible, age and auditory thresholds are often highly correlated. This serves to confound results. Future studies seeking to examine the effects of age and hearing status should draw samples from every combination of age and hearing status

(YNH, YHI, ONH, OHI). This would allow for a better understanding of the individual contributions of age and hearing status to performance in psychophysical tasks.

The results observed in our YNH and YHI listeners may appear to be at odds with the literature indicating that speech understanding in reverberant environments is generally worse for HI listeners than for NH listeners (Gelfand & Hochberg, 1976; Harris, 1965; Nabelek & Pickett, 1974a). It is important to note that speech understanding in reverberation and reverberation perception are not the same thing. While it is known that speech understanding varies with different T_{60} (Gelfand & Hochberg, 1976; Harris, 1965; Moncur & Dirks, 1967; Nabelek & Pickett, 1974a, 1974b; Nabelek & Robinson, 1982), it remains to be seen how other perceptual qualities of reverberation may affect speech understanding. One could imagine a speech understanding task in reverberation in which a group of listeners scores 50% correct. It has not been shown that changing the perceptual properties of the reverberation (e.g., *ASW*, *LEV*, etc.) directly affects performance on the speech understanding task. In this way, certain perceptual qualities of reverberation may be orthogonal to speech understanding.

General Conclusions

Across several different experimental techniques and samples of listeners, this dissertation has provided evidence that:

- The amount of reverberation that normal hearing listeners perceive is well-predicted by a model that represents the signal processing of the auditory periphery and then sums the reverberant sound power from the two ears.
- Normal hearing listeners are capable of segregating reverberant energy from ongoing reverberant time portions of speech in a room.
- Young normal hearing listeners and young hearing-impaired listeners with mixed losses use similar cues when making judgments about perceived reverberation.

REFERENCES

Allen, J. B., & Berkley, D. A. (1979). Image method for efficiently simulating small-room acoustics. *Journal of the Acoustical Society of America, 65*(4), 943-950.

Anderson, P. W., & Zahorik, P. (2014). Auditory/visual distance estimation: accuracy and variability. *Front Psychol, 5*, 1097. doi:10.3389/fpsyg.2014.01097

Barron, M., & Marshall, A. H. (1981). Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure. *Journal of Sound and Vibration, 77*(2), 211-232.

Bell, D. W., Kruel, E. J., & Nixon, J. C. (1972). Reliability of the modified rhyme test for hearing. *Journal of Speech and Hearing Research, 15*, 287-295.

Beranek, L. L. (2004). *Concert halls and opera houses : music, acoustics, and architecture* (2nd ed.). New York, NY: Springer.

Blauert, J. (1983). *Spatial hearing : the psychophysics of human sound localization*. Cambridge, Mass.: MIT Press.

Bolia, R. S., Nelson, W. T., Ericson, M. A., & Simpson, B. D. (2000). A speech corpus for multitalker communications research. *Journal of the Acoustical Society of America, 107*(2), 1065-1066. doi:Doi 10.1121/1.428288

Bradley, J. S., & Soulodre, G. A. (1995). Objective measures of listener envelopment. *Journal of the Acoustical Society of America, 98*(5), 2590 - 2597. doi:10.1121/1.413225

Breebaart, J., van de Par, S., & Kohlrausch, A. (2001). Binaural processing model based on contralateral inhibition. I. Model structure. *J Acoust Soc Am, 110*(2), 1074-1088.

Brungart, D. S., Cohen, J. I., Zion, D., & Romigh, G. (2017). The localization of non-individualized virtual sounds by hearing impaired listeners. *J Acoust Soc Am, 141*(4), 2870. doi:10.1121/1.4979462

Byrne, D., & Dillon, H. (1986). The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear, 7*(4), 257-265.

Carroll, J. D., & Chang, J. (1970). Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*(3), 283-319.

. Circular data analysis. (2018). In *NCCS Statistical Software*.

Dau, T., Puschel, D., & Kohlrausch, A. (1996). A quantitative model of the "effective" signal processing in the auditory system. I. Model structure. *J Acoust Soc Am, 99*(6), 3615-3622.

Day, M. L., & Delgutte, B. (2016). Neural population encoding and decoding of sound source location across sound level in the rabbit inferior colliculus. *J Neurophysiol, 115*(1), 193-207. doi:10.1152/jn.00643.2015

Dolan, T. R., & Robinson, D. E. (1967). Explanation of masking-level differences that result from interaural intensive disparities of noise. *J Acoust Soc Am, 42*(5), 977-981.

Egan, J. P. (1948). Articulation testing methods. *Laryngoscope, 58*(9), 955-991. doi:10.1288/00005537-194809000-00002

Ellis, G. M., Zahorik, P., & Hartmann, W. M. (2016). Using multidimensional scaling techniques to quantify binaural squelch. *Proc Meet Acoust, 23*(1), 1-10. doi:10.1121/2.0000164

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415*(6870), 429-433. doi:10.1038/415429a

Fitzpatrick, D. C., Kuwada, S., & Batra, R. (2002). Transformations in processing interaural time differences between the superior olivary complex and inferior colliculus: beyond the Jeffress model. *Hear Res, 168*(1-2), 79-89.

Gelfand, S. A., & Hochberg, I. (1976). Binaural and monaural speech discrimination under reverberation. *Audiology, 15*(1), 72-84.

Gescheider, G. A. (1985). *Psychophysics: Method, theory, and application*. New Jersey: Lawrence Erlbaum Associates, Inc.

Gilkey, R., & Anderson, T. R. (Eds.). (1997). *Binaural and spatial hearing in real and virtual environments*. New York, NY: Lawrence Erlbaum Associates, Inc.

Gilkey, R. H., & Anderson, T. R. (1997). *Binaural and spatial hearing in real and virtual environments*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Grey, J. M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustical Society of America, 61*(5), 1270-1277. doi:Doi 10.1121/1.381428

Harris, J. D. (1965). Monaural and Binaural Speech Intelligibility and the Stereophonic Effect Based Upon Temporal Cues. *Laryngoscope, 75*, 428-446. doi:10.1288/00005537-196503000-00003

Hassager, H. G., Wiinberg, A., & Dau, T. (2017). Effects of hearing-aid dynamic range compression on spatial perception in a reverberant environment. *J Acoust Soc Am, 141*(4), 2556. doi:10.1121/1.4979783

Hidaka, T., Beranek, L. L., & Okano, T. (1995). Interaural Cross-Correlation, Lateral Fraction, and Low-Frequency and High-Frequency Sound Levels as Measures of Acoustical Quality in Concert-Halls. *Journal of the Acoustical Society of America, 98*(2), 988-1007. doi:Doi 10.1121/1.414451

IEEE recommended practice for speech quality measurements. (1969). *IEEE Transactions on Audio and Electroacoustics, 17*(3), 225-246.

ISO-3382. (1997). Measurement of the reverbertion time of rooms with reference to other acoustical parameters. In *Acoustics*.

Jerger, J. (1973). Audiological findings in aging. *Adv Otorhinolaryngol, 20*, 115-124.

Jiang, D., McAlpine, D., & Palmer, A. R. (1997). Responses of neurons in the inferior colliculus to binaural masking level difference stimuli measured by rate-versus-level functions. *J Neurophysiol, 77*(6), 3085-3106.

Kim, D. O., Zahorik, P., Carney, L. H., Bishop, B. B., & Kuwada, S. (2015). Auditory distance coding in rabbit midbrain neurons and human perception: monaural amplitude modulation depth as a cue. *J Neurosci, 35*(13), 5360-5372. doi:10.1523/JNEUROSCI.3798-14.2015

Koenig, A. H. (1950). Subjective effects in binaural hearing. *Journal of the Acoustical Society of America, 22*(1), 61-62.

Koenig, A. H., Allen, J. B., Berkley, D. A., & Curtis, T. H. (1977). Determination of masking-level differences in a reverberant environment. *J Acoust Soc Am, 61*(5), 1374-1376.

Kruel, E. J., Nixon, J. C., Kryter, K. D., Bell, D. W., Land, J. S., & Schubert, E. D. (1968). A proposed clinical test of speech discrimination. *Journal of Speech and Hearing Research, 11*, 536-552.

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, Calif.: Sage Publications.

Kuttruff, H. (2000). *Room acoustics* (4th ed.). London, England ; New York, NY: Spon Press.

Kuwada, S., Bishop, B. B., & Kim, D. O. (2014). Azimuth and envelope coding in the inferior colliculus of the unanesthetized rabbit: Effect of reverberation and distance. *J Neurophysiol, 112*, 1340-1355.

Kuwada, S., & Yin, T. C. (1983). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. I. Effects of long interaural delays, intensity, and repetition rate on interaural delay function. *J Neurophysiol, 50*(4), 981-999.

Kuwada, S., Yin, T. C., Syka, J., Buunen, T. J., & Wickesberg, R. E. (1984). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. IV. Comparison of monaural and binaural response properties. *J Neurophysiol, 51*(6), 1306-1325.

Lokki, T., Patynen, J., Kuusinen, A., & Tervo, S. (2012). Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. *J Acoust Soc Am, 132*(5), 3148-3161. doi:10.1121/1.4756826

Lokki, T., Patynen, J., Kuusinen, A., Vertanen, H., & Tervo, S. (2011). Concert hall acoustics assessment with individually elicited attributes. *J Acoust Soc Am, 130*(2), 835-849. doi:10.1121/1.3607422

McDermott, J. H., Wrobleski, D., & Oxenham, A. J. (2011). Recovering sound sources from embedded repetition. *Proc Natl Acad Sci U S A, 108*(3), 1188-1193. doi:10.1073/pnas.1004765108

Moncur, J. P., & Dirks, D. (1967). Binaural and monaural speech intelligibility in reverberation. *J Speech Hear Res, 10*(2), 186-195.

Moore, B. C. J. (2013). *An introduction to the psychology of hearing* (Sixth edition. ed.). Leiden: Brill.

Moore, B. C. J., Glasberg, B. R., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *J. Audio Eng. Soc., 45*, 224-239.

Nabelek, A. K., & Pickett, J. M. (1974a). Monaural and binaural speech perception through hearing aids under noise and reverberation with normal and hearing-impaired listeners. *J Speech Hear Res, 17*(4), 724-739.

Nabelek, A. K., & Pickett, J. M. (1974b). Reception of consonants in a classroom as affected by monaural and binaural listening, noise, reverberation, and hearing aids. *J Acoust Soc Am, 56*(2), 628-639.

Nabelek, A. K., & Robinson, P. K. (1982). Monaural and binaural speech perception in reverberation for listeners of various ages. *J Acoust Soc Am, 71*(5), 1242-1248.

Nielsen, J. B., & Dau, T. (2010). Revisiting perceptual compensation for effects of reverberation in speech identification. *Journal of the Acoustical Society of America, 128*(5).

Okano, T. (2002). Judgments of noticeable differences in sound fields of concert halls caused by intensity variations in early reflections. *J Acoust Soc Am, 111*(1 Pt 1), 217-229.

Palmer, A. R., Jiang, D., & McAlpine, D. (1999). Desynchronizing responses to correlated noise: A mechanism for binaural masking level differences at the inferior colliculus. *J Neurophysiol, 81*(2), 722-734.

Palmer, A. R., Jiang, D., & McAlpine, D. (2000). Neural responses in the inferior colliculus to binaural masking level differences created by inverting the noise in one ear. *J Neurophysiol, 84*(2), 844-852.

Reinhart, P. N., Souza, P. E., Srinivasan, N. K., & Gallun, F. J. (2016). Effects of Reverberation and Compression on Consonant Identification in Individuals with Hearing Impairment. *Ear Hear, 37*(2), 144-152. doi:10.1097/AUD.0000000000000229

Robinson, D. E., & Jeffress, L. A. (1963). Effect of varying the interaural noise correlation on the detectability of tonal signals. *Journal of the Acoustical Society of America, 35*(12), 1947-1952.

Rumsey, F. (2002). Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm. *Journal of the Audio Engineering Society, 50,* 651-666.

Schiffman, S. S., Reynolds, M. L., & Young, F. W. (1981). *Introduction to multidimensional scaling : theory, methods, and applications*. New York: Academic Press.

Shore, A., Hartmann, W. M., Rakerd, B., Ellis, G. M., & Zahorik, P. (2016). *Squelch of room effects in everyday conversation*. Paper presented at the 171st Meeting of the Acoustical Society of America, Salt Lake City, UT.

Sivonen, V. (2007). Directional loudness and binaural summation for wideband and reverberant sounds. *Journal of the Acoustical Society of America, 121*(5), 2852 - 2861.

Sivonen, V., Alanko, M., Gamper, H., Raummukainen, O., & Pulkki, V. (2011). Reverberation perception with one and two ears. *Forum Acusticum*.

Stevens, S. S. (1957). On the psychophysical law. *Psychol Rev, 64*(3), 153-181.

Traer, J., & McDermott, J. H. (2016). Statistics of natural reverberation enable perceptual separation of sound and space. *Proc Natl Acad Sci U S A, 113*(48), E7856-E7865. doi:10.1073/pnas.1612524113

van Dorp Schuitman, J., de Vries, D., & Lindau, A. (2013). Deriving content-specific measures of room acoustic perception using a binaural, nonlinear auditory model. *J Acoust Soc Am, 133*(3), 1572-1585. doi:10.1121/1.4789357

Watkins, A. J. (2005). Perceptual compensation for effects of reverberation in speech identification. *Journal of the Acoustical Society of America, 118*(1).

Watkins, A. J., & Raimond, A. P. (2013). Perceptual compensation when isolated test words are heard in room reverberation. In B. C. J. Moore, R. Patterson, W. I., R. P. Carlyon, & G. H. (Eds.), *Basic Aspects of Hearing: Advances in Experimental Medicine and Biology* (Vol. 787). New York, NY: Springer.

Whitmer, W. M., Seeber, B. U., & Akeroyd, M. A. (2012). Apparent auditory source width insensitivity in older hearing-impaired individuals. *J Acoust Soc Am, 132*(1), 369-379. doi:10.1121/1.4728200

Wightman, F. L., & Kistler, D. J. (1989a). Headphone simulation of free-field listening. I: Stimulus synthesis. *J Acoust Soc Am, 85*(2), 858-867.

Wightman, F. L., & Kistler, D. J. (1989b). Headphone simulation of free-field listening. II: Psychophysical validation. *J Acoust Soc Am, 85*(2), 868-878.

Wojtczak, M., & Viemeister, N. F. (2005). Forward masking of amplitude modulation: basic characteristics. *J Acoust Soc Am, 118*(5), 3198-3210.

Yin, T. C., & Kuwada, S. (1983a). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. II. Effects of changing rate and direction of interaural phase. *J Neurophysiol, 50*(4), 1000-1019.

Yin, T. C., & Kuwada, S. (1983b). Binaural interaction in low-frequency neurons in inferior colliculus of the cat. III. Effects of changing frequency. *J Neurophysiol, 50*(4), 1020-1042.

Zahorik, P. (2002a). Assessing auditory distance perception using virtual acoustics. *J Acoust Soc Am, 111*(4), 1832-1846.

Zahorik, P. (2002b). Direct-to-reverberant energy ratio sensitivity. *J Acoust Soc Am, 112*(5 Pt 1), 2110-2117.

Zahorik, P. (2009). Perceptually relevant parameters for virtual listening simulation of small room acoustics. *J Acoust Soc Am, 126*(2), 776-791. doi:10.1121/1.3167842

Zahorik, P., & Ellis, G. M. (2016). *An example of dissociation between speech intelligibility and perceived reverberation*. Paper presented at the 171st Meeting of the Acousical Society of America, Salt Lake City, UT.

Zahorik, P., Kim, D. O., Kuwada, S., Anderson, P. W., Brandewie, E., Collecchia, R., & Srinivasan, N. (2012). Amplitude modulation detection by human listeners in reverberant sound fields: Carrier bandwidth effects and binaural versus monaural comparison. *Proc Meet Acoust, 15*. doi:10.1121/1.4733848

Zahorik, P., Kim, D. O., Kuwada, S., Anderson, P. W., Brandewie, E., & Srinivasan, N. (2011). Amplitude modulation detection by human listeners in sound fields. *Proc Meet Acoust*, *12,* 50005-50010.

Zurek, P. M. (1979). Measurements of binaural echo suppression. *J Acoust Soc Am*, *66*(6), 1751-1757.

APPENDIX A

GLOSSARY

Auditory source width (*ASW*): the subjective width of a sound source.

Binaural: of or related to both ears; can refer to a stimulus present at both ears, a cue present at both ears, or a measurement made that requires information from both ears.

Binaural room impulse response (BRIR): the impulse response measured from two microphones in a room (ISO-3382, 1997).

Binaural squelch: increased perception of reverberation in monaural or diotic listening relative to binaural and dichotic listening (Koenig, 1950).

Clarity index (C_t): the ratio of energy before time point *t* to energy after time point *t*, where *t* is in milliseconds. Expressed in dB.

Dichotic: a listening condition in which the signals at the two ears are different

Diotic: a listening condition in which the signals at the two ears are identical

Direct-to-reverberant energy ratio (DR): the ratio of direct energy to reverberant energy expressed in dB.

Direct sound: The sound that propagates directly from a source to a receiver in a room.

Early reflections: The sound that propagates from a source then reflects off one or two surfaces before reaching a receiver in a room. The precise number of reflections depends on the size of the room (Kuttruff, 2000).

Head-related transfer function (HRTF): the filtering properties of the head, shoulders, and ears of an individual listener. Individualized HRTFs are matched to the listener from which they were measured. Non-individualized HRTFs were not.

Interaural cross-correlation (IACC): The peak cross-correlation between the left and right ear BRIR within a time-window of 2.5 ms.

Late reverberation: The diffuse time portion of a sound in a room.

Listener envelopment (*LEV*): the subjective sense that one is completely surrounded by reverberation.

Monaural: A listening condition in which only one ear receives a signal.

Receiver: An apparatus that is capable of receiving and working with incoming pressure waves. A microphone and a human listener are examples of sound sources.

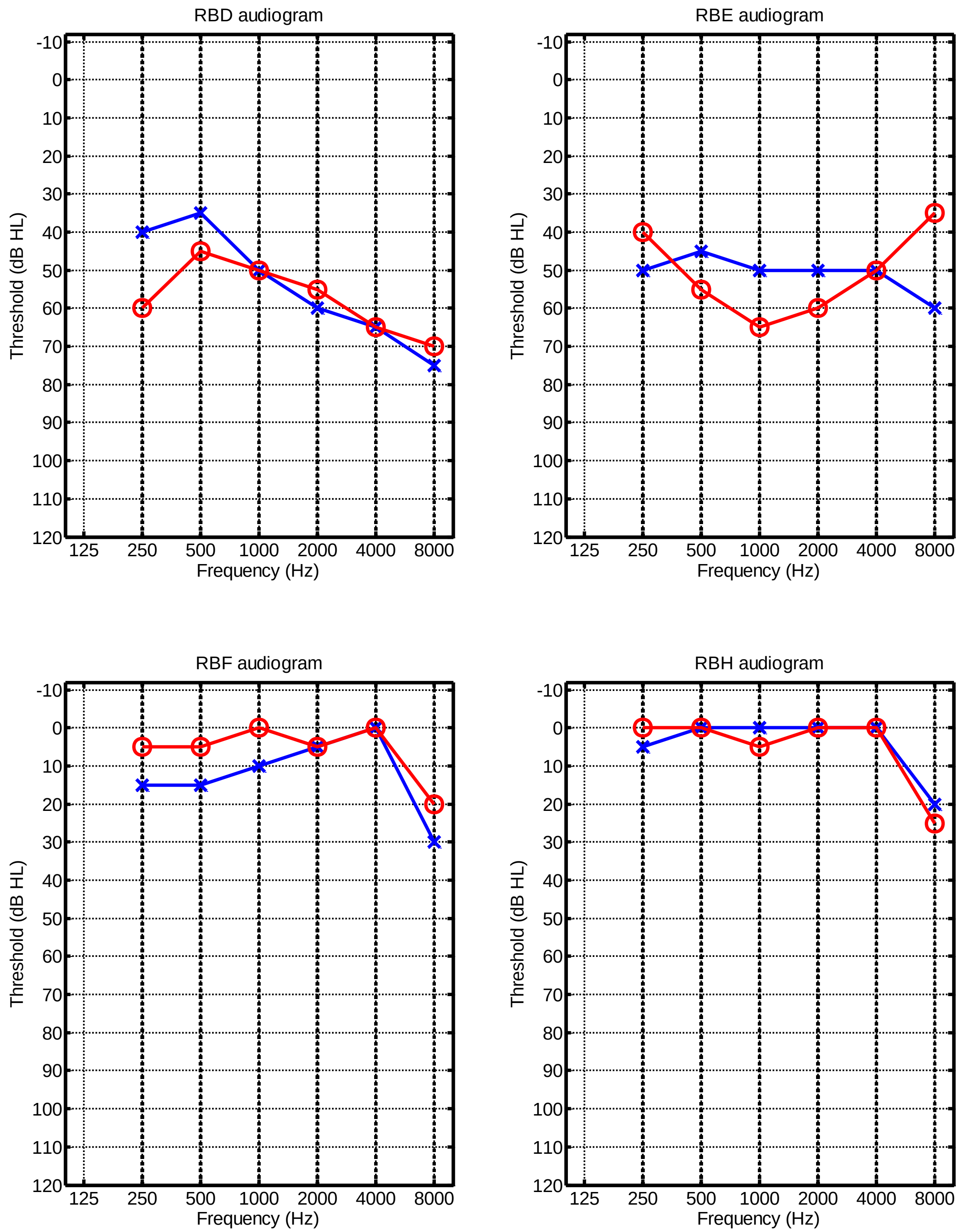
Reverberation time (T_L): Measured from the impulse response, the amount of time it takes for a sound in a room to decay by L dB once it begins to decay linearly.

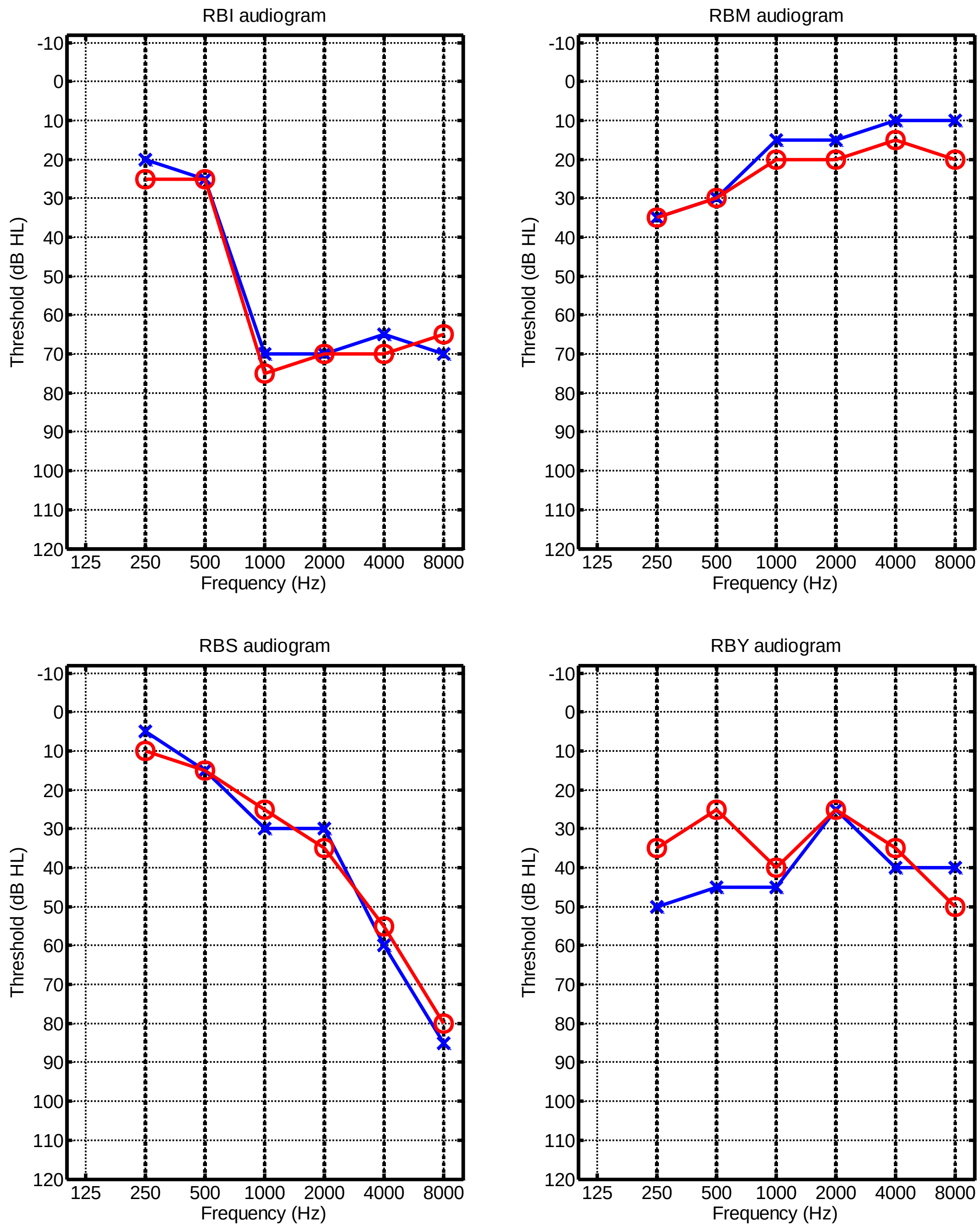
Source: A vibrating body that generates sound waves. A voice, a violin, a loudspeaker, and headphones are all examples of sound sources.

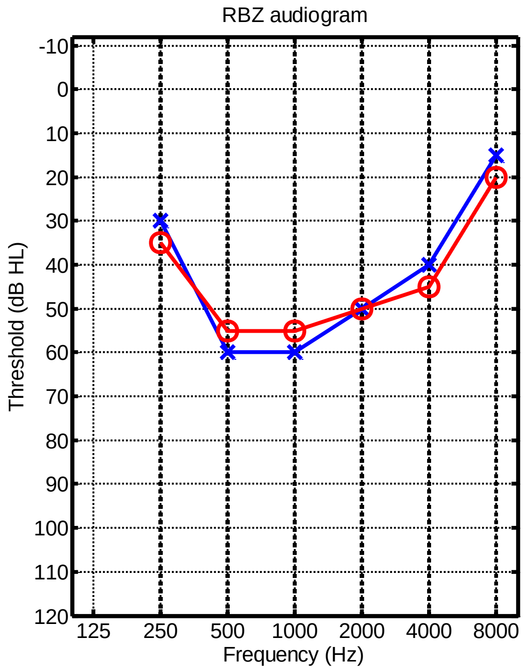
APPENDIX B

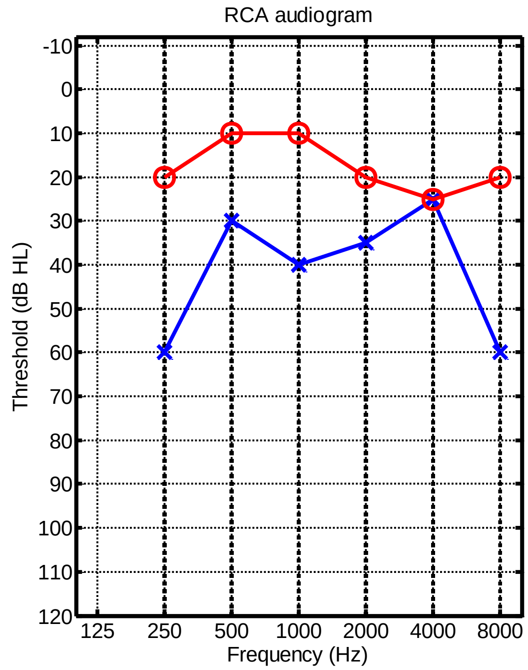
HEARING-IMPAIRED AND NORMAL-HEARING LISTENERS' AUDIOGRAMS

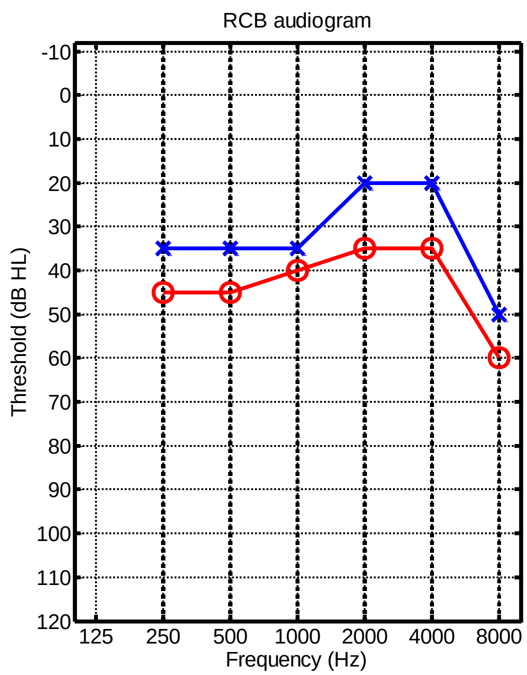
Below are the pure-tone air-conduction audiograms for the 11 hearing-impaired listeners and 15 normal hearing listeners tested in Chapter 4. Hearing-impaired listeners are identified by three-letter codes that start with "R". Normal-hearing listeners are identified by three-letter codes that start with "Z". Pure tone frequency is plotted on the x-axis. Threshold is plotted on the y-axis. There is a large amount of variability between listeners. The losses are all sensorineural with some degree of conductive loss in many of the listeners. The nature of the losses was made jointly with a practicing audiologist.

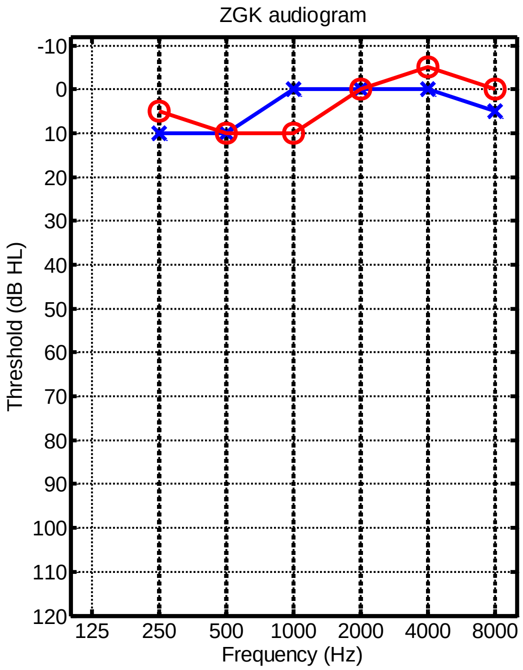

RBD audiogram
RBE audiogram
RBF audiogram
RBH audiogram
Threshold (dB HL)
Frequency (Hz)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000

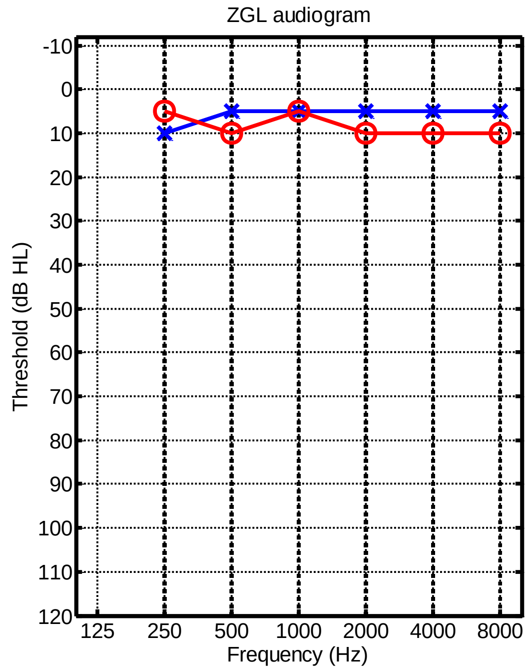

RBI audiogram
RBM audiogram
RBS audiogram
RBY audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

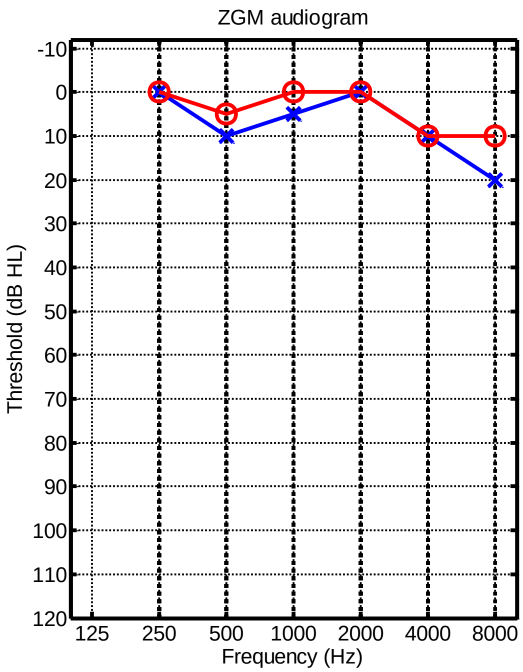

RBZ audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

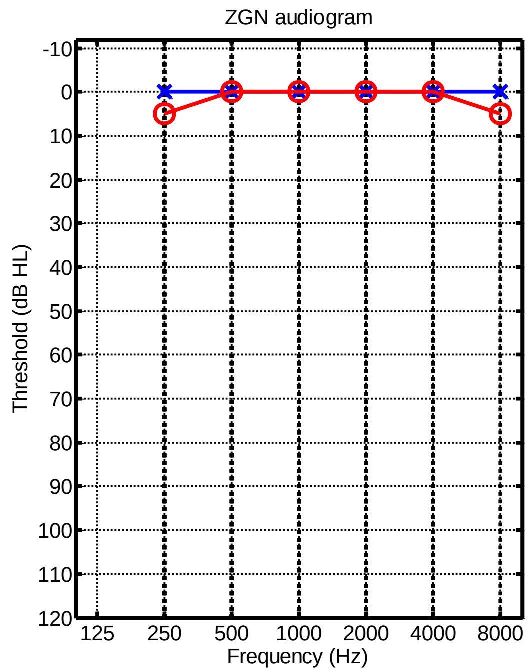

RCA audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

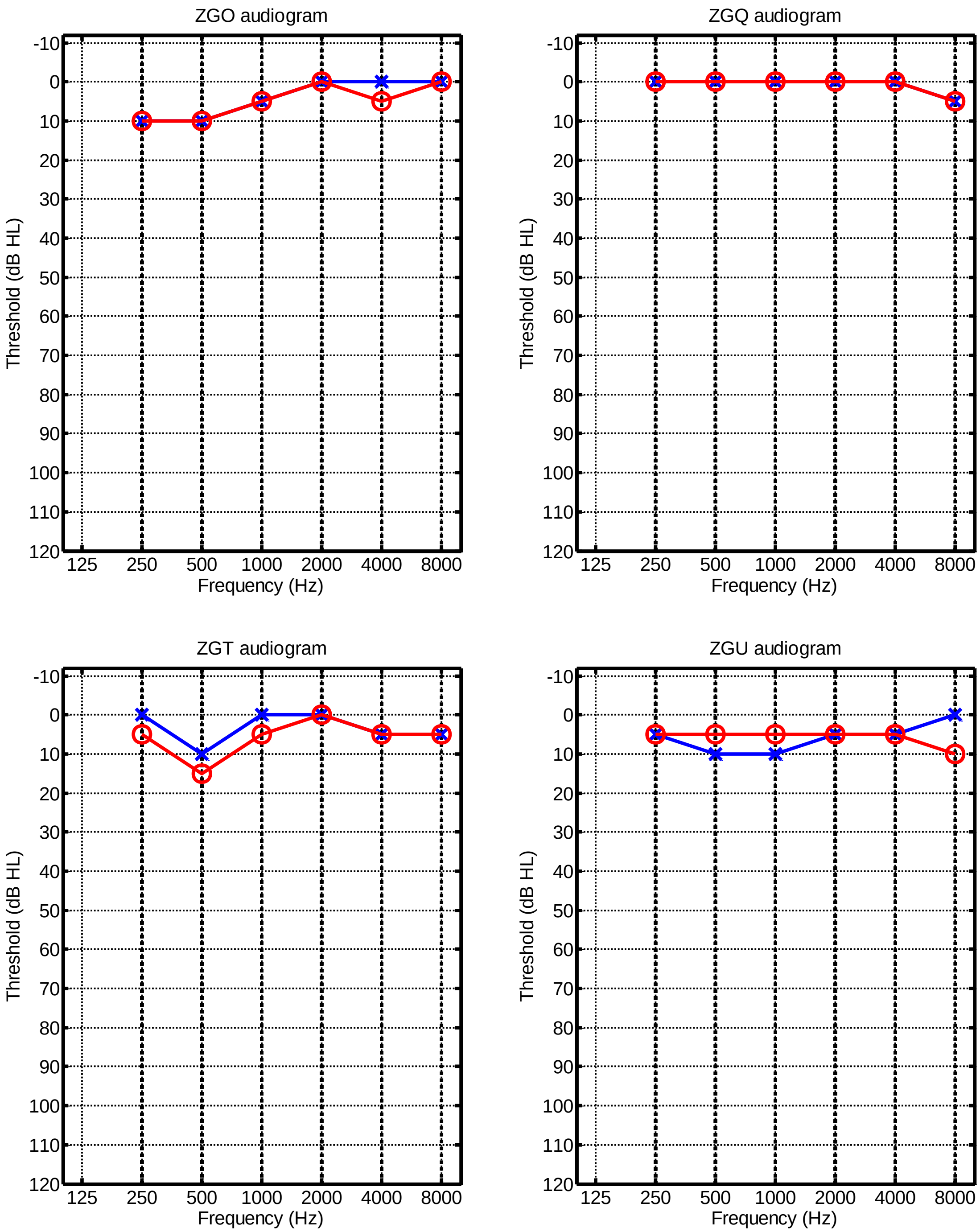

RCB audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

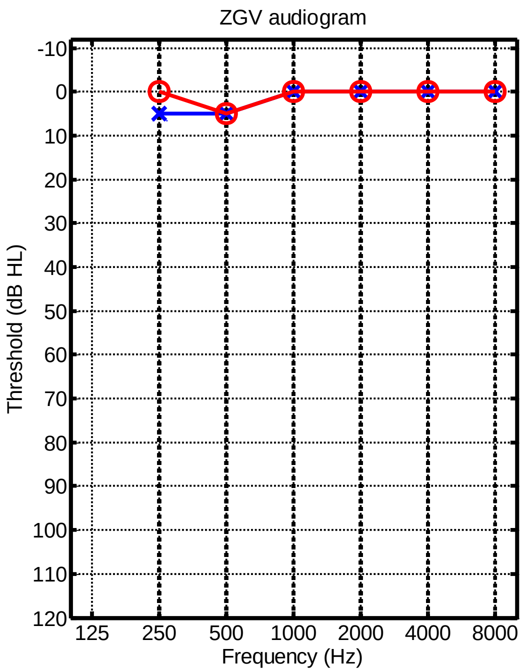

ZGK audiogram
Threshold (dB HL)
Frequency (Hz)

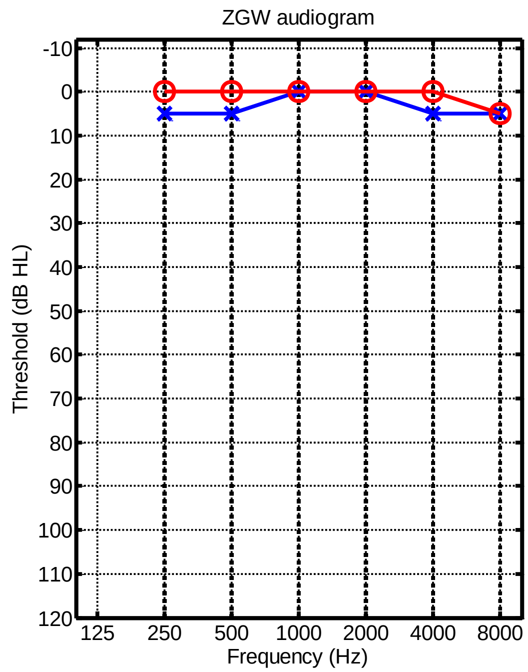

ZGL audiogram
Threshold (dB HL)
Frequency (Hz)

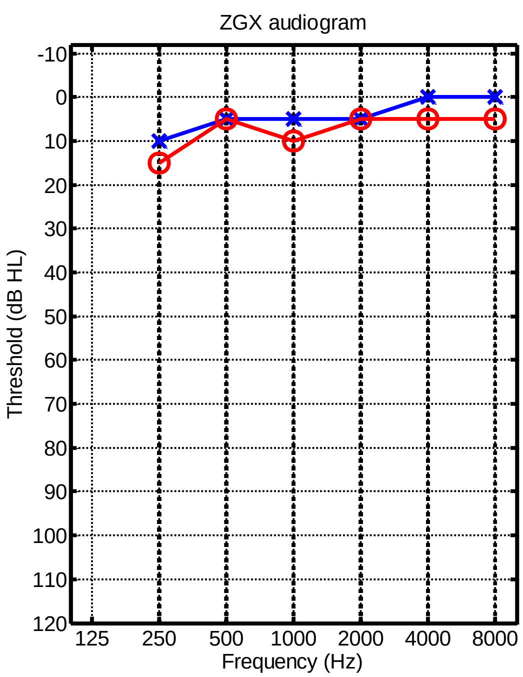

ZGM audiogram
Threshold (dB HL)
Frequency (Hz)

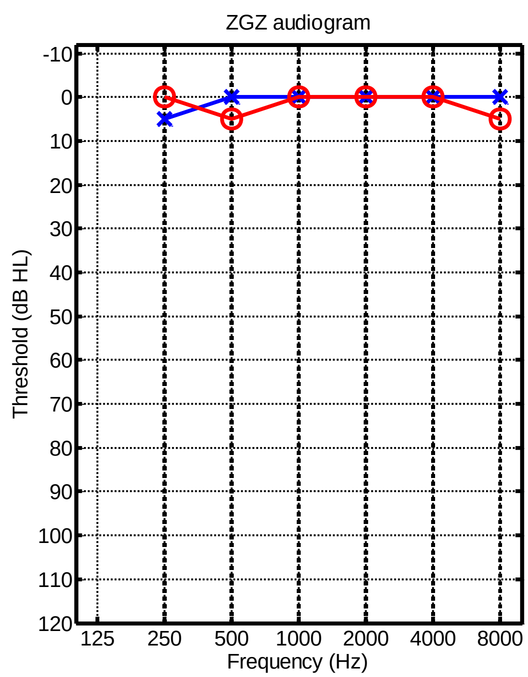

ZGN audiogram
Threshold (dB HL)
Frequency (Hz)

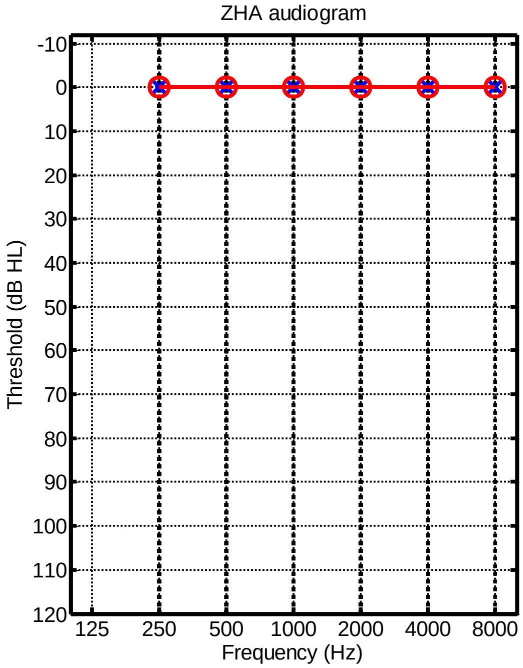

ZGO audiogram
ZGQ audiogram
ZGT audiogram
ZGU audiogram
Threshold (dB HL)
Frequency (Hz)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000

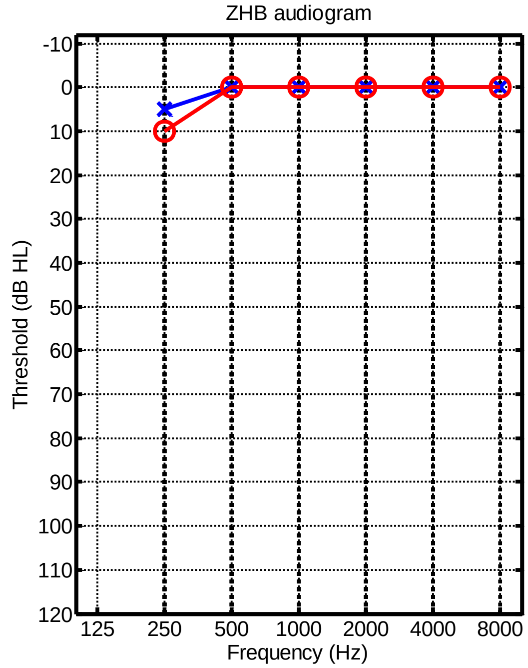

ZGV audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

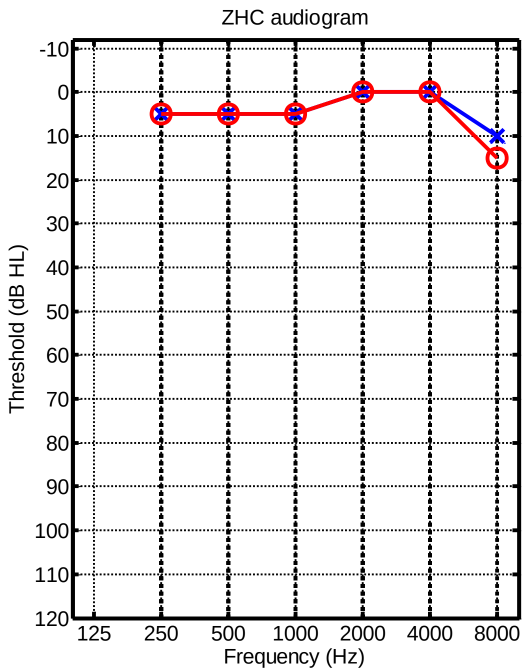

ZGW audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)


ZGX audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)


ZGZ audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)


ZHA audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)


ZHB audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)


ZHC audiogram
Threshold (dB HL)
-10
0
10
20
30
40
50
60
70
80
90
100
110
120
125
250
500
1000
2000
4000
8000
Frequency (Hz)

CURRICULUM VITA

Gregory Matthew Ellis, M.A., M.Sc.

Department of Psychological and Brain Sciences
2301 South Third Street
University of Louisville
Louisville, KY 40292
Phone: 612-703-6992
Email: ellisgm90@gmail.com

EDUCATION

2016 - Present
Ph.D., Experimental Psychology (expected completion, June 2018)
Dissertation: "The effects of monaural and binaural cues on reverberation perception in normal hearing and hearing-impaired listeners"
University of Louisville, Louisville, KY

2014 – 2016
M.Sc., Experimental Psychology
University of Louisville, Louisville. KY

2012 - 2014
M.A., Music Composition
University of Minnesota—Twin Cities, Minneapolis, MN

2008 - 2012
B.A., Music Composition
University of Wisconsin—Eau Claire, Eau Claire, WI

RESEARCH EXPERIENCE

August 2014 - Present
Graduate Research Fellow
Zahorik Auditory Perception Lab, PI: Pavel Zahorik, Ph.D.
Department of Psychological and Brain Sciences
University of Louisville, Louisville, KY

February 2013 - May 2014
Research Assistant
Auditory Perception and Cognition Lab, PI: Andrew Oxenham, Ph.D.
Department of Psychology, Cognitive and Brain Sciences
University of Minnesota-Twin Cities, Minneapolis, MN

Spring 2007 Research Assistant
PI: Gary Don, Ph.D.
Department of Music and Theatre Arts
University of Wisconsin-Eau Claire, Eau Claire, WI

PUBLICATIONS

Peer-reviewed publications

Shore, A., Hartmann, W.M., Rakerd, B., **Ellis, G.M.**, & Zahorik P. (Submitted). "Squelching reverberation and coloration of speech through binaural presentation." Submitted to the Journal of the Acoustical Society of America Express Letters.

Miyazaki, K., Rakowski, A., Makomaska, S., Jiang, C., Tsuzaki, M., Oxenham, A.J., **Ellis, G.M.**, & Lipscomb, S.D. (In Press). "Absolute pitch and relative pitch in music students in the east and the west: Implications for aural-skills education." *Music Perception.*

Stilp, C.E., Anderson, P.W., Assgari, A.A., **Ellis, G.M.**, & Zahorik, P. (2016). "Speech perception adjusts to reliable spectrotemporal properties of the listening environment." *Hearing Research,* 341, 168-178.

Manuscripts in preparation

Ellis, G.M., & Zahorik, P. (in preparation). "Perceived reverberation strength and speech intelligibility in virtual listening conditions."

Conference Proceedings

Ellis, G.M., Zahorik, P., & Hartmann, W.M. (2016). "Using multidimensional scaling techniques to quantify binaural squelch." *Proceedings of Meetings on Acoustics, 23*(1), 1-10.

Talks

Ellis, G.M. & Zahorik, P. (2018, May). "Binaural perceptual weighting of reverberation level in normal hearing listeners." Paper presented at the 175th Meeting of the Acoustical Society of America, Minneapolis, Minnesota.

Abstracts

Ellis, G.M. & Zahorik, P. (2018, February). "Binaural aspects of perceived reverberation strength in normal-hearing and hearing-impaired listeners." Poster presented at the 41st Meeting of the Association for Research in Otolaryngology, San Diego, CA.

Ellis, G.M., & Zahorik, P. (2017, June). "Perceived amount of reverberation consistent with binaural summation model." Poster presented at the 173rd Meeting of the Acoustical Society of America, Boston, Massachusetts.

Miyazaki, K., Rakowski, A., Makomaska, S., Jiang, C., Tsuzaki, M., Oxenham, A. J., **Ellis, G.**, & Lipscomb, S.D. (2016, July). "Absolute pitch and relative pitch in music students: A comparison between East and West." Paper presented at the 31st International Congress of Psychology. Yokohama, Japan.

Zahorik, P., & **Ellis, G.M.** (2016, May). "An example of dissociation between speech intelligibility and perceived reverberation." Paper presented at the 171st Meeting of the Acoustical Society of America, Salt Lake City, Utah.

Shore, A., Hartmann W.M., Rakerd, B., **Ellis, G.M.**, & Zahorik, P. (2016, May). "Squelch of room effects in everyday conversation." Paper presented at the 171st Meeting of the Acousitcal Society of America, Salt Lake City, Utah.

Ellis, G.M., and Zahorik, P. (2016, Feb.). "A situation in which the ipsilateral ear does not contribute to the amount of perceived reverberation." Poster presented at the 39th Meeting of the Association for Research in Otolaryngology, San Diego, CA.

Ellis, G.M., Zahorik, P., & Hartmann, W.M. (2015, May). "Using multidimensional scaling techniques to quantify binaural squelch". Poster presented at the 169th Meeting of the Acoustical Society of America, Pittsburgh, Pennsylvania.

Stilp, C.E., Zahorik, P., Anderson, P.W., Assgari, A.A., **Ellis, G.M.** (2015, May). "Reverberation increases perceptual calibration to reliable spectral peaks". Paper presented at the 169th Meeting of the Acoustical Society of America, Pittsburgh, Pennsylvania.

PROFESSIONAL PRESENTATIONS

Ellis, G.M. (2016). The role of binaural summation in the amount of perceived reverberation. Presented at psychology department meeting.

Ellis, G.M. (2015). On the benefit of having two ears: Perceived reverberation and binaural squelch. Presented at psychology department meeting.

TEACHING EXPERIENCE

August 2016 - Present
Graduate Teaching Assistant
PSYC 301: Quantitative Methods in Psychology
University of Louisville, Louisville, KY

August 2012 - May 2014
Graduate Teaching Assistant
MUSI 1013 Rock I: The Historical Origins and Development of Rock Music to 1970
MUSI 1014 Rock II: Rock Music from 1970 to Present
University of Minnesota--Twin Cities, Minneapolis, MN

GRANTS/AWARDS

May 2016 GSC Travel Funding ($350)
December 2015 ARO Travel Award ($500)
October 2015 GNAS Travel Award ($100)
February 2015 GSC Travel Funding ($250)
September 2014GSU Travel Award ($100)
July 2014 - Present Graduate Student Fellowship, University of Louisville

PROFESSIONAL AFFILIATIONS

Acoustical Society of America
Association for Research in Otolaryngology

SERVICE

July 2015 -
June 2017 Experimental Psychology Graduate Student Representative

REFERENCES

Pavel Zahorik, Ph.D., Associate Professor & Heuser Hearing Research Endowed Chair
Division of Communicative Disorders
Department of Surgery, University of Louisville School of Medicine, Louisville, KY
Email: pavel.zahorik@louisville.edu Phone: 502-852-3843

Christian Stilp, Ph.D., Assistant Professor
Department of Psychological and Brain Sciences, University of Louisville, Louisville KY
Email: christian.stilp@louisville.edu Phone: 502-852-0820

Andrew Oxenham, Ph.D., Professor
Departments of Psychology and Otolaryngology
University of Minnesota–Twin Cities, Minneapolis, MN
Email: oxenham@umn.edu Phone: 612-624-2241