LOUDNESS PERCEPTION AT AND NEAR ELEVATED THRESHOLD: IS SOFT STILL SOFT?

by

Reem Mulla

BA, King Saud University, 2006

MA, University of Pittsburgh, 2010

AuD, University of Pittsburgh, 2011

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2018

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This dissertation was presented

by

Reem Mulla

It was defended on

October 2, 2018

and approved by

Christopher Brown, PhD, Associate Professor, Department of Communication Science and Disorders

Sheila Pratt, PhD, Professor, Department of Communication Science and Disorders

Stuart Steinhauer, PhD, Adjunct Research Associate Professor, Department of Psychiatry University of Pittsburgh School of Medicine

Dissertation Advisor: Catherine Palmer, PhD, Associate Professor, Department of Communication Science and Disorders

LOUDNESS PERCEPTION AT AND NEAR ELEVATED THRESHOLD: IS SOFT STILL SOFT?

Reem Mulla, PhD

University of Pittsburgh, 2018

Background: Two differing models of loudness recruitment (abnormally fast growth of loudness above elevated thresholds) for individuals with sensory hearing loss have been described (Buus & Florentine, 2002; Moore & Glasberg, 2004). The two models provide conflicting data related to perceived loudness at elevated thresholds and loudness growth near threshold in listeners with sensory hearing loss compared to normally hearing listeners. The present study was conducted to gain insight into this discrepancy.

Methods: 29 listeners with normal hearing and 29 listeners with hearing loss participated in a simple yes/no detection task for 4000 Hz tones presented at and near their hearing threshold (at -4, 0, 4, 10, and 16 SL) while their pupil dilation response was recorded. Participants also completed a subjective rating task to judge the loudness of the same tones and other at higher levels up to 28 dB SL.

Results: A significant difference between groups was seen in the pupil dilation response at threshold (0 SL) and 10 SL conditions. At threshold, pupil dilation in normal hearing listeners initiated earlier and was sustained longer compared to listeners with hearing loss consistent with increased difficulty of sound detection at threshold. Similar response behavior was observed at -4 SL. At 10 SL, pupil dilation in listeners with hearing loss was sustained longer compared to normal hearing listeners. Pupil dilation to tones at other levels (4 and 16 SL conditions) was not different between groups. Both groups subjectively rated the loudness of tones at all levels similarly with similar loudness growth patterns.

Conclusion: Results suggest that normal hearing listeners experienced more difficulty in the sound detection task at threshold, as well as more uncertainty in decision making. This observation may be consistent with a louder perception for tones at threshold in listeners with hearing loss, which supports the softness imperception loudness model put forth by Florentine et al. In general, caution should be exercised when interpreting pupillary responses to directly indicate perceived loudness or psychoacoustic sensation as task induced cognitive processing may more heavily contribute to the response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This dissertation is the fruit of hard work that lasted for several years. I couldn't wait to reach that point when I thank all the people who were behind this work and this achievement. I am there now, and would like to reflect on those people who have supported and helped me throughout this period.

I would like to express the deepest appreciation to my committee chair Dr. Catherine Palmer who has the attitude and the substance of a genius. She continually conveyed a spirit of adventure regarding research and scholarship, and an excitement in teaching and mentoring. I was amazed by your personality since day one of my AuD program. I knew I would learn lifelong lessons from you, so I decided to keep you close by and requested that you be my advisor. It was the right decision. After being part of your remarkable lab, I fell in love with research and science and stayed longer to pursue my PhD. What you have taught me throughout the years cannot be found in books. You modeled a great thinker, a sharp observer, a real leader, an excellent clinician, and above all that, a woman and mother with a balanced work-life. I am grateful for your inspiration and endless support.

I am also grateful for the considered thoughts of my committee members throughout the dissertation process. Thank you, Dr. Christopher Brown, for always welcoming me when I kept showing up in your lab "unannounced" to bounce ideas off you, to borrow countless pieces of equipment (mics, cable, adaptors), or just to have a cup of your great coffee. Your expertise was overwhelming at times, but it sharpened my investigation and facilitated the technical setup of my experiment stimuli. Thank you Dr. Sheila Pratt for your spare, but highly insightful comments and guidance which helped me improve the quality of my study. Thank you, Dr. Stuart Steinhauer, for

guiding me in my pupillometry learning journey. I am forever honored that I have worked with you.

Thank you, Dr. Elaine Mormer, for always reminding me that “the lowest times in my PhD life will be equally the highest times after my PhD”. That really kept me going. Thank you for always opening your door to me from day one before I started my AuD, until the last day of my PhD.

Thank you AuD students Chelsie Williams, Kayla Copperthite, and Jenna Loffredo for your help keeping this study going when I struggled to keep up.

I would like to express my special thanks to Ronen Hershman. Without his amazing CHAP software, I would still be at work manually processing my data. Thank you for all the time you generously spent with me discussing my research and reviewing my data files.

To my father, my role model, Mohamed Jamil, thank you! I believe that everything I try to do in life, including this PhD, reflects who you are and what you have taught me. Thank you for your endless support and encouragement throughout my PhD journey. My mother, Rawiah, I truly believe that your prayers were what made me succeed in this endeavor. Thank you for every moment you thought and worried about me.

My husband, Tamim. I don’t think I will ever be able to thank you enough or ever a way to return what you have given me. Thank you for all those times you listened patiently as I talked to you about my research and bounced ideas off you and even asked for your opinion. You made me truly believe that you were interested in what I do when I know that you just wanted me to GET IT DONE! I hope I can offer you the same support to achieve your dreams. Thank you for believing in my passion and encouraging me to reach my career ambitions.

I thank my sisters Rana and Ru'a, and my brothers, Sami, Ahmad, and Feras. Despite our living all these years apart, in different continents, all of you have been a constant source of encouragement and unconditional support.

Thanks to my lab mates whose presence made the lab feel like another home. Our continuous discussions and occasional "wining" about PhD life helped relieve the pressures from time to time.

Thank you to my friends in Saudi Arabia and in the United States who were part of this long journey. Every single one of you modeled an important key to success. Thank you all for standing by me even though I was thousands of miles away.

Finally, I thank my biggest accomplishment in life: my children Bijad and Rayana. I will not say that I could not have done this without you because I know I could have, and maybe a lot faster! But I would not have wanted to do it any other way. You gave everything I did throughout this journey a deeper meaning and a greater purpose. Thank you for all the cuddles, hugs and smiles that were literally my "fuel" to go back the next day and continue in this journey. You made reaching the finish line a happier moment. I want you to know that in the spring of 2018, I enjoyed earning my PhD, but not nearly as much as I enjoyed going back to be with both of you.

1.0 **INTRODUCTION**

1.1 STATEMENT OF PROBLEM

Hearing losses of primarily cochlear origin are known to not only impair the detection of sound at the peripheral level but also to alter the perception of sound in the psychological domain. Loudness perception of sounds around us is an auditory process that involves all levels of the auditory pathway and any disruption in cochlear function affects this perception.

The physiology of the hearing mechanism has two facets: the mechanical processing of acoustical signals and the neurological processing of the acoustical signal information. Research has described changes in the hearing mechanism that result in hearing loss; however, since hearing is not a passive function, understanding the different ways in which the physical properties of sounds are neurally encoded in an impaired auditory system and presented as mental phenomena is still an ongoing, active interest in psychoacoustic studies.

In recent decades models have been developed to describe how listeners with cochlear hearing loss perceive loudness of sounds within their reduced dynamic range (Moore & Glasberg, 1997; Moore & Glasberg, 2004). Normalizing loudness perception across all input levels or overall loudness has been one of the main goals of a hearing aid fitting (Keidser, Dillon, Carter, & O'Brien, 2012; Rajkumar, Muttan, Jaya, & Vignesh, 2013; Scollie et al., 2005). Other goals of hearing aid fittings include achieving audibility, enhancing speech intelligibility, and improving safety. Loudness models serve as a framework for hearing aid designers and rehabilitationists. The goal is to use amplification to compensate for the changes in the perception of loudness and the loudness growth profile (Rajkumar et al., 2013) with the hopes of returning normal loudness perception.

The perception of amplified soft sounds by new hearing aid users has been consistently problematic, as reported in the hearing aid literature (Blamey & Martin, 2009; Johnson & Cox, 2013; Mueller & Powers, 2001; Shi, Doherty, & Zwislocki, 2007). Soft sounds are the low-level sounds that are amplified to be at a level that is just above threshold for a listener with hearing loss. Despite the continuous enhancements to hearing aid designs, and continuous refining of hearing aid fitting protocols to follow evidence-based practice, returning normal loudness perception of soft sounds to hearing aid users has not been achieved; soft sounds are consistently perceived louder than target loudness despite the chance for adaptation over time (Blamey & Martin, 2009; Johnson & Cox, 2013; Mueller & Powers, 2001; Shi et al., 2007).

Currently, hearing aid designers use the loudness model by Moore and Glasberg (2004) as the gold standard to predict loudness and the loudness growth profile of input auditory signals for listeners with hearing loss. This loudness model has failed to account for the loudness perception associated with near threshold level presentations for many hearing aid users. A more recent body of literature proposes that listeners with hearing loss of cochlear origin manifest different loudness and loudness growth patterns for low-level sounds (soft sounds) (Buus & Florentine, 2002; Florentine, Buus, & Rosenberg, 2005) than what has been proposed by the currently accepted gold standard.

Given the possible shortcomings in Moore & Glasberg's model and growing debate about the nature of loudness perception of low-level sounds in listeners with cochlear hearing loss, the purpose of this background (section 2.0) is three-fold: 1) to review the theoretical framework of loudness perception in listeners with hearing loss of cochlear origin; 2) to further examine the potential gap in the literature by examining the theoretical rationales behind the two differing views; and 3) to identify research questions that might enhance our understanding of how listeners with cochlear hearing loss perceive loudness compared to normal hearing listeners at or near threshold levels.

1.2 SIGNIFICANCE OF WORK: LOUDNESS AT THRESHOLD - WHAT WE KNOW AND WHY IT IS IMPORTANT

The category of sounds that are just above a normal hearing listener's threshold of audibility and up to about 50 dB SPL are the sounds that are perceived as soft by normal hearing listeners. Soft sounds are all the sounds and voices around us when it is quiet, such as the humming of a ventilation system, ocean waves, birds twittering, people chatting across the street, etc. Soft sounds include white noise or background noises that we typically ignore or wish to ignore, but also can be speech signals that are at a distance from the listener to which we might wish to listen.

Soft sounds are present within moderate and loud conversations. For example, speech is an amplitude modulated signal with segments, such as sibilants, that are lower in amplitude than the average level of the speech signal at conversational level. Soft sounds are not only important for the positive listening experience they carry, but because their presence in conversational speech preserves the modulation that is a cue for speech perception and understanding.

The category of soft sounds, with its psychological associations and contribution to speech cues, is the first one to be missed in the dynamic range of a listener with hearing loss. In listeners with mild to moderate hearing loss, the sounds that are just above their elevated threshold are now a different group of sounds with different acoustical properties than those that are perceived at moderate/average loudness for normal hearing listeners. An extensive amount of research has focused on how loudness perception is altered with hearing loss, but the majority of these studies have been conducted at suprathreshold levels - the levels of everyday conversations. Little research has focused on the change in perception of loudness at and near the elevated thresholds.

There are several possible reasons for this. First, researchers are generally more interested in knowing how loudness grows within the dynamic range once the signal is audible because it is important for the design of hearing aids and for psychoacoustic studies. In psychoacoustic loudness models, loudness at elevated thresholds is always assumed to be equal to loudness at normal thresholds. This assumption makes sense and lends itself to the valid use of psychoacoustic models at suprathreshold levels. Second, the assumption of equal loudness at threshold has good face validity, and has led to reliable, consistent model outcomes at suprathresholds. Lastly, it is generally assumed that less communication and interesting auditory signals are located at threshold level.

However, there are reasons why a better understanding of perceived loudness at threshold when a listener's threshold is elevated is important and can enhance our understanding of hearing loss in general and can influence future directions in audiology and psychoacoustic research. If loudness at threshold for some listeners with cochlear hearing loss is greater than for listeners with normal thresholds, current goals and techniques of hearing aid fittings might need to be modified. First, the expectation that new hearing aid users will fully adapt to newly audible soft sounds and perceive these sounds as soft after some period of full time hearing aid use may be misguided. Second, the use of performance of normally hearing individuals to generate the targets for loudness of soft sounds for individual with hearing loss who are using hearing aids may not be justified. Psychoacousticians rely heavily on the concept of recruitment in designing studies and interpreting results when listeners with sensorineural hearing loss are involved. Better understanding of loudness perception at and near elevated threshold in individuals with cochlear hearing loss may lead to future changes in the definition and concept of recruitment. Thus, understanding loudness perception of barely audible signals at elevated threshold has potential significance for understanding sensorineural hearing loss and for hearing aid design.

2.0 BACKGROUND

2.1 THE PERCEPTION OF LOUDNESS WITH COCHLEAR HEARING LOSS

2.1.1 What is Loudness?

Definition

Loudness is a psychological term. It can be defined as the psychological attribute of sound intensity, or the perceptual attribute of a sound that changes when intensity is varied (Scharf, 1978). Although these definitions are accurate, they might not be comprehensive. Definitions that limit the change in perception of loudness only to the change in intensity of the acoustic stimulus discard all other factors that might contribute to the perception of loudness.

Loudness is primarily dependent on the intensity of the sound, but other acoustical variables of a sound contribute to its perceived loudness, such as frequency, duration, bandwidth, and mode of listening (binaural vs. monaural, in quiet vs. background noise). Therefore, a more comprehensive definition of loudness might be "the magnitude of an auditory sensation" by Fletcher and Munson (1933). Fletcher's auditory sensation implies endless contributions to this sensation at the stimulus level and includes the role played by the listener. Listeners' factors can be divided into physiological (integrity of the auditory system) and psychological (cognitive, cross-cultural, fatigue, etc.) factors. In the dictionary, loudness is defined as "the attribute of a sound that determines the magnitude of the auditory sensation produced and that primarily depends on the amplitude of the sound wave involved" (Merriam-Webster, 2004, p. 736). The American

National Standards Institute (ANSI) defines loudness as that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud (ANSI, 1994).

Sound: Pressure, intensity, and level

Sounds are carried by a pressure variation in some material. Sounds occurring in the air around us are the ones most commonly listened to. The ear is capable of detecting very tiny pressure variations that reach less than one billionth of atmospheric pressure. The intensity of a sound is the amount of energy transmitted per second through a square meter of air and is proportional to the square of the pressure.

Our auditory system operates over a wide range of intensities since the pressure of a sound near pain threshold is a million times greater than a sound near absolute threshold. The logarithmic decibel (dB) scale was developed to express this wide range of sound level intensities that our auditory system perceives in a manageable way (Plack, 2004). The logarithmic scale expresses the ratio of two intensities, reference intensity and another intensity that is expressed relative to it (Moore, 2012). A constant increase in dB scale (from 10 to 20 to 30 in dB) corresponds to a constant multiplication of the sound intensity (multiplied by 10, by 100, by 1000…). The reference of this dB scale is 0 dB SPL (Sound Pressure Level) and this reference is the lowest sound level we can hear with a pressure of 0.00002 Newtons per meter.

The rest of this chapter is divided into four sections. The first section reviews the impact that acoustical parameters of different stimuli have on the perception of loudness. The second section focuses on the physiology underlying intensity coding in the auditory system. The third section reviews the theories and models related to loudness prediction. The final section reviews the correlates of loudness that have been used as measures of loudness.

2.1.2 Loudness and Physical Properties of Sound

Many auditory perceptual attributes depend primarily on one physical aspect of the sound and secondarily to other physical aspects. Pitch, timbre, and sound location, for example, depend on relatively complex spectral and temporal aspects of sounds. It can be argued, however, that loudness is closely related to one physical aspect of sound: the intensity level. Loudness changes primarily as a function of level, but it also changes as a function of frequency, bandwidth, duration, spectral complexity, and the presence of other sounds. Outside of well-controlled laboratories, there is never a one-to-one correspondence between loudness and any one physical property of a sound, including level. In this section, attempts to quantify loudness as a function of physical properties of sounds are reviewed.

2.1.2.1 Loudness units: Phons and Sones

Phons and sones are two different units that are used to express loudness. They are both based on 1000 Hz reference tones. Since loudness is primarily correlated with sound intensity, loudness is often described as a function of intensity.

The sone is an arbitrary unit of perceived loudness proposed by Stevens (Stevens, 1936, 1955, 1956). It is a relative unit because it is referenced to a 1000 Hz tone at 40 dB SPL presented binaurally in a free field at 0 azimuth degrees. The loudness of 1 sone is assigned to that sound (Stevens, 1936). A sound that is twice as loud as a 1000 Hz tone at 40 dB SPL would be assigned a loudness of 2 sones. A sound that is half as loud as 1000 Hz at 40 dB SPL would be 0.5 sones loud, and so on. Therefore, the sone is always a positive number (Epstein & Marozeau, 2010).

For sound levels above 40 dB SPL, doubling the perceived loudness (in sones) corresponds approximately to a 10 dB SPL increase in sound level (Hellman, 1976; Moore, 2012; Schreiner &

Malone, 2015). For sound levels below 40 dB SPL, where loudness changes more rapidly with increasing level, this relationship does not hold.

Loudness in sones is a sensation value, but loudness level in phones comprises both sensation and physical values (Zwicker & Fastl, 2013). Indeed, the phon is a unit for loudness *level*. Loudness level of a given sound in phon is the level of a 1000 Hz tone in dB SPL presented binaurally in a free field at 0 azimuth, to which it sounds equally loud to that given sound (Fletcher & Munson, 1933). Table 1 shows the relationship between loudness in sones and loudness level in phons. Every doubling in perceived loudness in sones requires a 10 dB SPL level increase.

Table 1: The relationship between phons and sones. The relationship between loudness level in phons of a 1 KHz tone and loudness in sones based on the prediction of a loudness model (Moore, Glasberg, & Baer, 1997).

sone	1	2	4	8	16	32	64	128	256	512	1024
phon	40	50	60	70	80	90	100	110	120	130	140

2.1.2.2 Sound level and loudness: Loudness functions

Loudness function is a plot of loudness estimates on a logarithmic scale as a function of the audible range of sound levels (intensities) in dB SPL. Intensity is the primary physical parameter that influences loudness, and its effect on loudness is intuitive. Psychacousticians have been trying to come up with a mathematical formula that predicts subjective loudness of a sound based on its known physical intensity. These efforts were started in the nineteenth century by physicists and philosophers who aimed to develop such formulas using magnitude estimation and magnitude production methods (Hellman & Zwislocki, 1963).

The pioneer at that time was Weber who was followed by Fechner. Fechner combined Weber's empirical observation of just noticeable differences (JNDs), the minimum amount of intensity increase needed to induce a perceptual loudness difference, with the assumption that JNDs should show an equal interval scale, i.e., each sensation increase would be identical in magnitude (Stern & Morgan, 2012). Earlier work by Stevens (1955, 1957) modified the loudness growth function and introduced it as a power function of sound level.

More investigations on loudness growth function over time resulted in the inflected exponential function (INEX) (Buus & Florentine, 2001; Florentine & Epstein, 2006). The INEX function is considered one of the most accurate functions that describe the growth of loudness as intensity level increases for normal hearing listeners. The INEX's modifications to the power function were more successful in predicting loudness data across wider intensity levels (Florentine & Epstein, 2006). More details on the differences between the loudness functions are discussed in section 2.4.

The form of the loudness function does not change much for persons with normal hearing at frequencies between 500 Hz and 8000 Hz (Scharf, 1978). Loudness function is normally given for the 1000 Hz tone. However, it can be plotted for other frequencies using the equal loudness contours.

Loudness functions for a 1 KHz tone presented in free field with frontal incidence can be divided into three phases, with each phase having a different growth slope. The relationship between sound level and loudness is always monotonic (i.e., an increase in level will always be perceived as an increase in loudness). However, this relationship is not always linear. For levels between 30-40 and 70 dB SPL, loudness grows in a compressive fashion, where loudness approximately doubles for every 10 dB increase in sound level (Chen, Hu, Glasberg, & Moore, 2011; Hellman, 1976; Hellman & Zwislocki, 1961). This relationship is usually described as a power function: Loudness $=kI^{0.3}$ where k is a consonant depending on the subject and the units used and *I* is the stimulus intensity (Hellman, 1976). This compressive relationship is a result of outer hair cell (OHC) function that amplifies the basilar membrane (BM) displacement in response to low-level sounds and not to high-level sounds. Below and above this range, the slope becomes progressively steeper and approaches unity (Chen et al., 2011; Hellman, 1976). More details on the impact of OHC function on BM displacement and loudness are covered in section 2.3.

Because sounds around us are complex, complex models have been developed to predict loudness from the spectrum of complex sounds. More details about these models are reviewed in the loudness models section of this paper.

The loudness-intensity function is independent of sound duration (Epstein & Florentine, 2005b) and the presentation mode, monaural vs. binaural (Marozeau, Epstein, Florentine, & Daley, 2006; Whilby, Florentine, Wagner, & Marozeau, 2006b). Monaural and binaural loudness functions are parallel when plotted on a log scale (Whilby, Florentine, Wagner, & Marozeau, 2006a). However, the loudness function does depend to some extent on the spectral content of the sound (Scharf, 1959).

The growth of loudness mirrors the response of the auditory periphery and has always been related to BM displacement. Above 30 dB SPL, there is a compressive relationship between the amplitude of BM displacement and stimulus intensity, where the growth of the BM displacement is non-linear with increasing intensity. For low-level sounds near threshold, the growth of BM displacement with increasing intensity is steeper and so is the growth of loudness. At these low levels, outer hair cell compression is not yet active. At very high intensities, the BM displacement returns to grow more steeply with increasing intensity, when the passive component of BM vibration overtakes the saturating active component (Pickles, 2012). Studies that derived the loudness growth slope from BM displacement amplitude at the characteristic frequency (CF) have sometimes underestimated the loudness growth slope (Yates, 1990). These findings may suggest that the BM displacement and the loudness function are related, but the relationship is not straightforward (Zhang & Zwislocki, 1992).

Loudness at threshold: The starting point of the loudness function

Marking the starting point of the loudness growth function on the Y-axis of a graph received notable attention in the loudness literature for a period of time. For many years, the loudness of sounds at a normal threshold was assumed to be zero. For a 1 KHz tone, the loudness of a 3 dB SPL tone, which is the threshold of hearing, reaches zero. This zero corresponds to a minus infinity value on a logarithmic scale (Fastl & Zwicker, 2007). Buus, Müsch, and Florentine (1998) pioneered a change in this concept by concluding that loudness at normal thresholds exceeds zero, at least in most listeners, meaning that once a sound is heard, there must be a percept associated with it. Their study aimed to determine the form of loudness function at low-level sounds and whether the loudness at threshold is greater than zero. Six subjects matched the loudness of tone complexes whose components had equal SLs and were separated by one to as many as six critical bands, and a 1600 Hz tone in an adaptive paradigm. Subjects matched tone complexes of subthreshold components to a pure tone that was a few dB above threshold. Based on a simple loudness summation model, the authors concluded that loudness of tones at or even below threshold is greater than zero. Models that predict loudness for normal hearing listeners incorporated this modification to more accurately predict loudness (Moore et al., 1997). Models that predict loudness for listeners with hearing loss assume that loudness at elevated threshold equals the loudness at threshold for normal hearing listeners. Section 3 of this paper will review the literature related to this concept and discuss the opposing views on loudness at threshold.

2.1.2.3 Frequency and loudness: equal loudness contours

Loudness level was introduced in the twenties by Barkhausen (Fastl & Zwicker, 2007). The loudness level of a sound is the sound pressure level of a 1000 Hz tone presented horizontally at a 0 azimuth that is as loud as the sound (Zwicker & Fastl, 2013). As mentioned earlier, the unit of measure for loudness level is phon. When a 1000 Hz tone is fixed in level and test tones of various frequencies are adjusted to match the loudness of the fixed tone, an equal-loudness contour is generated for that fixed level in phon (Fletcher & Munson, 1933). Equal loudness contours illustrate that loudness also depends on frequency, in addition to intensity, and that it does so differently at different intensities. An equal loudness contour curve does not provide an absolute measure of loudness. A contour only provides loudness level, a relative measure to a 1000 Hz tone's loudness. Absolute hearing thresholds across frequencies also form an equal loudness contour that indicates when the lower limit of loudness sensation is reached. The absolute threshold loudness contour corresponds to 3-4 dB at 1K Hz.

The shapes of equal loudness contours vary between studies depending on methodological details, such as the test-tone level range employed (Gabriel, Kollmeier, & Mellert, 1997). The graph in Figure 1 shows examples of the equal loudness contours that are predictions based on different loudness models and contour curves that were obtained empirically from listeners with normal hearing.

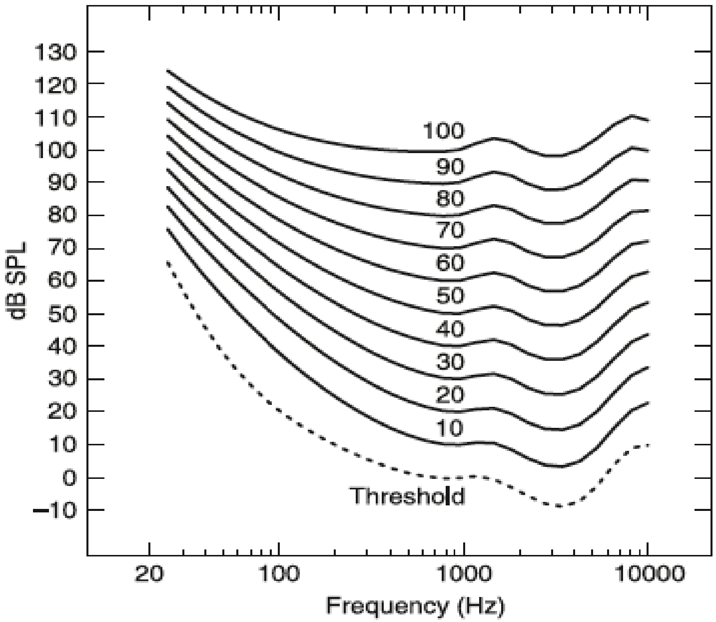


Figure 1: The equal loudness contours specified in the International Organization for Standardization (ISO). Graph obtained from Florentine (2011) with permission.

When we examine the equal loudness contour curves in Figure 1, tones of the same intensity within the frequency range between 600 and 2000 Hz, are perceived to have almost equal loudness. Higher frequency sounds presented at the same intensity, especially between 3 to 4 KHz where the ear is most sensitive due to ear canal resonance, are perceived to be louder. Lower frequency sounds (especially below 300 Hz) are perceived to be less loud since the ear is least sensitive to this frequency range. The contour lines in this very low-frequency area are closer together, implying that loudness grows faster with increasing level.

2.1.2.4 Duration, bandwidth and loudness

Duration

The human ear averages sound energy over approximately 200 ms (approximately 400 ms for broadband signals). A normal listener's ability to make loudness judgments when a sound's duration is less than 200 ms is compromised because below this threshold, loudness grows as a power function of duration (Stevens & Hall, 1966). Above this temporal threshold, and assuming the listener continues to attend to the stimuli, the duration of the sound should not impact its perceived loudness.

The loudness matching paradigm is used to measure the effect of duration on loudness where subjects match the loudness of a sound with a fixed short duration to a standard sound with a longer fixed duration. Due to the effect of level and frequency on this critical time, a handful of studies have reported different values; some as small as 15 ms to 500 ms, and others up to one second (Ekman, Berglund, & Berglund, 1966; Small, Brandt, & Cox, 1962).

Bandwidth

In the case of complex tones, loudness is affected by the spectrum. Critical bandwidth is the bandwidth where the perceptual change occurs (Scharf, 1970). The more a complex tone or noise is spread out across many critical bands, the louder it is perceived even if the overall intensity is fixed (Zwicker, Flottorp, & Stevens, 1957). This phenomenon is called spectral summation. Varying the bandwidth of a noise or the separation between complex tone components within a single critical band will result in a total loudness of the sum of the intensities and will remain constant as long as the bandwidth is less than the critical bandwidth.

In contrast to normal hearing listeners, for listeners with hearing loss, the loudness of sounds of a constant intensity does not markedly increase with increasing bandwidth due to their wider critical bands (Scharf & Hellman, 1966).

2.1.3 The Underlying Physiology of Loudness Perception

The field of psychoacoustics was established in the early 1800s. Beginning in the 1900s, the field became closely connected to the physiology of hearing. Psychoacousticians sought to determine the mechanisms that might contribute to the formation of the perceptual attributes of sounds. Their studies were limited to animal models due to the invasive nature of the procedures needed to obtain neural responses at different stations along the auditory pathway.

Modern psychoacousticians have continued to investigate the physiological mechanisms underlying loudness perception. It has been hypothesized that loudness sensation depends on the sum of activities carried along the auditory nerve (AN) (Evans, 1981; Fletcher & Munson, 1933). The AN, which conveys impulses from the cochlea to the central auditory system, is thought of as the bottleneck; the auditory centers in the brain are limited by the information conveyed through the AN. At the same time, there has been no consensus among researchers on the adequacy of this and other hypotheses to inform us about all the actual physiological mechanisms that may determine the sensation of loudness.

Only recently has research been directed toward investigation of the contribution of central processes to the perception of loudness (Röhl & Uppenkamp, 2012). In the next section, loudness coding at the peripheral level will be reviewed for both the normally functioning cochlea and the pathological cochlea. Loudness coding at the central level also will be reviewed in a following section.

2.1.3.1 Loudness coding at the peripheral level: normally functioning cochlea

Coding of intensity changes of a sound starts in the cochlea, where mechanical vibrations get transduced and encoded into neural impulses in the auditory nerve. The sound vibrations traveling through the ear canal, the middle ear and finally to the cochlear fluid set the basilar membrane into a vertical displacement. The basilar membrane displacement excites the hair cells in the Organ of Corti, which in turn elicits the auditory nerve fiber's action potential. The greater the intensity, the greater the amplitude of the basilar membrane displacement. This difference in basilar membrane size of displacement for different sound levels is translated into different amounts of perceived loudness based on three theories: firing rate theory (Fletcher & Munson, 1933), spread of excitation theory (Rose, Hind, Anderson, & Brugge, 1971; Ruggero, 1992), and temporal cues (phase locking) theory (Javel & Mott, 1988; Rose, Brugge, Anderson, & Hind, 1967). The first two theories appear to contribute more to explaining the coding of loudness at the peripheral level than the third theory (phase locking).

Firing rate hypothesis

It was initially believed that loudness is simply proportional to the sum of all the firing rates of auditory nerve neurons: the average number of times a nerve fiber activates per second for a given stimulus (Fletcher & Munson, 1933). In other words, according to this hypothesis, the more spikes per second in the nerve fiber, the louder the sound is. This relationship can be seen in a rate-level function. Auditory nerve fibers show spontaneous firing in the absence of any sound stimulation. For the detection of a sound, the sound must elicit a change in the firing rate of neurons with central frequencies close to the stimulus frequency. The threshold of a neuron is the lowest level of a sound at which such change occurs. Although the rate-level function is probably the best theory of loudness physiology, there has not been consensus on the adequacy of firing rate theory to accommodate what has been called the dynamic range problem.

The dynamic range of individual auditory nerve fibers, defined as the stimulus levels over which firing rate increases from threshold to saturation, does not cover the dynamic range of all intensity levels the human auditory system can perceive. The firing rate saturation of most auditory nerve fibers is limited to 35-60 dB above the threshold at which the firing rate is independent of the stimulus intensity (Sachs & Abbas, 1974). Auditory nerve fibers cannot fire faster than about 200 spikes per second. Above about the 30 dB SPL range, the vibration of the basilar membrane starts to saturate, that is, it starts to grow more slowly as a function of intensity (Pickles, 2012). Around this level, the active mechanism of the outer hair cells (OHC) that magnifies the amplitude of the basilar membrane starts to saturate. This is where the excitation of the basilar membrane starts to widen as well. Intensity discrimination of higher-level sounds cannot be explained by the firing rate theory alone. Another mechanism must complement the firing rate theory to accommodate the loudness perception of intensity change and discrimination of higher-level sounds.

The findings of some studies questioned the firing rate theory as being the only physiological correlate of loudness. Discrepancies were found between the growth of auditory nerve firing rate with increasing tone intensity and the growth of loudness, where growth of firing rate was shallower (Jeng, 1992; Relkin & Doucet, 1997). Relkin and Doucet (1997) found that the firing rate equal-count contours obtained from a chinchilla did not match the equal-loudness contours of human listeners. The growth of the auditory nerve firing rate in response to pure tones was slower than the growth of loudness of pure tones. Similar results were found by Jeng (1992).

Spread of excitation theory

As mentioned previously, with high-level sounds, the firing rate hypothesis is not adequate to code the high intensity perception. The spread of the excitation pattern is what is thought to contribute to intensity coding at such levels. Because the auditory nerve has 30,000 nerve fibers, many fibers collaborate to code intensity. They share the workload by having different dynamic ranges and different spontaneous firing rates and thresholds. With high level sounds, the basilar membrane does not only show greater vertical displacement, but also a wider region gets involved in this displacement. The excitation pattern can be defined as the representation of the distribution of activity by a sound along the basilar membrane (Fletcher, 1940 cited in Chen et al., 2011; Zwicker, 1956). This is simply the shape of the traveling wave along the BM in response to sound. The width of the excitation pattern is determined by the width of the auditory filter for that sound's characteristic frequency (CF). One can think of the sound entering the cochlea as being decomposed and fed to different filters, each centered at a different CF. The width of each filter varies with frequency. The higher the CF, the wider the auditory filter. The excitation patterns are the output of those overlapping auditory filters as a function of characteristic frequencies.

Just like the cochlea is tonotopic, so are the corresponding auditory nerve fibers. Nerve fibers exhibit different firing thresholds and different saturation levels as a function of frequency. They also exhibit different spontaneous firing rates (i.e., spontaneous spiking in the absence of a stimulus) (Liberman, 1982). Neurons with high spontaneous firing rates (18-250 spikes per sec) are associated with low firing thresholds (as low as 0 dB SPL) and small dynamic range (15-30 dB); and vice versa, neurons with low spontaneous firing rates (<0.5 spikes per sec) are associated with high firing thresholds (as high as 80 dB SPL) and large dynamic range (as much as 60 dB) (Liberman, 1978; Rose et al., 1971). Each individual neuron has its own dynamic range that may be only between 15-60 dB. However, this combination of neurons is what allows the cochlea to respond to all levels within the dynamic range of our auditory system (up to 120 dB), and therefore, solves the dynamic range problem.

The higher the sound level, the wider the excitation pattern becomes. Wider excitation patterns imply that a wider area of the basilar membrane is stimulated and a larger number of nerve fibers that have characteristic frequencies adjacent to the tone frequency start to fire. This increased number of firing neurons is coded as greater loudness. The combination of neurons with different thresholds and different dynamic ranges assures the ability to code intensities across the entire dynamic range of hearing.

The spread of excitation supplements the intensity coding for two reasons. First, it increases the amount of firings coming from neighboring neurons with characteristic frequencies just above and below the stimulus. These additional firing rates beside the characteristic frequency neuron's firing rate will be summed at the auditory nerve level. Second, spread of excitation also increases the number of neurons involved in this increased firing rate. Both cues play a role in coding loudness.

Intensity encoding by the change in average firing rate of AN fibers as a function of level, and the spread of excitation across the population of fibers with increasing level are sometimes referred to as rate-place schemes, as opposed to the temporal scheme, namely the phase locking theory.

Phase locking theory

Phase locking occurs when the exact timing of the auditory nerve fibers' spikes lock to specific phases of the stimulus waveform and occur roughly at the same phase each time. Cues related to phase locking may play a role in intensity coding and discrimination, particularly for complex stimuli where there are different relative levels of frequency components (such as speech). The firing rate of an auditory nerve fiber is average information (average number of pulses per second). When the firing rate, on average, reaches saturation, the timing pattern information of this firing rate might still vary and convey an additional cue about the intensity of the stimuli at the auditory nerve level (Carney, 1994; Pickles, 2012).

Another theory states that with increasing sound level, more neurons phase-lock to it, which increases the overall synchrony across the population of auditory nerve fibers (Javel & Mott, 1988; Rose et al., 1967). Phase locking improves in precision with increased sound level at low levels. At medium and high levels, phase locking precision stays roughly constant. Above 5 KHz, however, phase locking is lost (Johnson, 1980). Therefore, this spatio-temporal pattern cue cannot be the neural correlate of loudness of higher frequency stimuli.

In summary, the intensity coding literature suggests that loudness, at least of a pure tone, is directly related to the BM response (measured BM velocity) that is translated to the AN. The neural correlate of loudness at the AN is based on the firing rate of each auditory neuron, the spread of excitation across the population of AN fibers, as well as the spatial patterns of the auditory nerve activity when all the neural activities are combined. All the codes of signal intensity are maintained within the AN.

2.1.3.2 Loudness coding at the peripheral level: abnormally functioning cochlea

When hearing loss involves damage to the structure and function of the cochlea, it is referred to as cochlear hearing loss. Cochlear damage can arise in many ways, for example by noise exposure, ototoxic drugs, infection, metabolic disturbances, autoimmune disorders, or genetic factors. When the effect(s) of the damage extends beyond the cochlea to the auditory nerve and higher centers in the auditory pathway, the hearing loss is more generally referred to as sensorineural hearing loss. In this paper, the phrase hearing loss will be used to refer to cochlear hearing loss, the hearing loss caused by cochlear damage. Most sensorineural hearing losses result from damage within the cochlea. It is common for this damage to be largely confined to specific structures within the cochlea. Pure cochlear hearing loss can easily be produced in animal models, whereas in humans with hearing loss it is rarely a pure cochlear hearing loss (Moore, 2007).

Cochlear hearing loss involves damage to the inner hair cells (IHCs) and/or outer hair cells (OHCs), damage or destruction of the stereocilia, or total death of the hair cells. Cochlear physiology evidence supports the notion that the normal function of the cochlea depends largely on the active mechanism that is strongly linked to OHC function.

The basilar membrane (BM) in a healthy cochlea is very sharply tuned to frequencies. It moves in a very specific area that best corresponds to the stimulating sound's frequency. Each point along the BM shows the greatest displacement to a certain frequency, the characteristic frequency (CF), and responds progressively less as the frequency moves away from the CF. Two mechanisms in cochlear function are believed to be responsible for this tuning of the BM: the passive mechanism and the active mechanism. The passive mechanism depends on the mechanical aspects of the BM and surrounding structures, and it operates in a roughly linear manner (Békésy & Wever, 1960). The progressive variation in the width and stiffness of the BM across its length causes the BM to respond best to different frequencies at different points starting from high frequencies at the base to low frequencies at the apex (Békésy & Wever, 1960; Moore, 2003). The active mechanism, on the other hand, depends on the function of the OHCs and it operates in a nonlinear manner. The active mechanism is responsible for three cochlear functions: 1) sensitivity to low-level sounds; 2) sharp tuning of the BM; and 3) compressive nonlinearity of the BM's input/output (I/O) function.

The OHCs increase the amount of vibration on the BM in response to low-level sounds. The loss or dysfunction of OHCs elevates the absolute threshold due to the loss or reduction of this boost to low-level sounds. The dysfunction of IHCs also elevates the threshold due to the reduction in the efficiency of neural transduction to the AN fibers.

In terms of sharp tuning, the sharp tip of the BM traveling wave envelope comes from normally functioning OHCs. Figure 2 shows the contribution of the active mechanism in the two cochlear functions mentioned above (sensitivity to low-level sounds and sharp tuning). The more damage there is to the OHC, the more the sensitivity to low-level sounds and the tuning of the BM are expected to be impaired as the cochlea will function on the remaining passive mechanism only. The effects of OHC loss on auditory neuron tuning curves described in the literature are thought to reflect changes in BM tuning curves (Narayan, Temchin, Recio, & Ruggero, 1998). Therefore, similar effects are observed from auditory neurons and BM tuning curves for the same cochlear region.

Auditory filters were first described by Fletcher (1940) as a bank of overlapping bandpass filters in the auditory system. It is a concept that is largely related to critical bandwidth by Zwicker et al. (1957). Although our auditory system functions in an orchestrated manner across its pathway, the auditory filters are an example of when the auditory system does not actually work together to produce the perception of the sound. With the reduction of the BM tuning, the auditory filters become broader and all the peaks and dips of the spectral envelope in the auditory signal are no longer reflected as sharp peaks and dips in the BM, a phenomenon referred to as reduced frequency selectivity (Glasberg & Moore, 1986). Figure 3 shows the shape of normal and impaired auditory filters obtained for listeners with unilateral hearing loss. Auditory filters obtained from impaired ears imply that the target sound (at CF) can now be easily masked by adjacent frequency sounds especially lower frequency sounds where the impaired auditory filters show a broader side. This phenomenon contributes to what is called upward spread of masking. Excitation patterns, which are derived from auditory filters of a given sound, are thought of as the output of the auditory filters as a function of their center frequency (Moore & Glasberg, 1983) (See figure 4). Excitation patterns that are derived from wider auditory filters are wider than normal, as well. Auditory filter bandwidth increases with increasing frequency (e.g., (Glasberg & Moore, 1990; Moore, 2003).

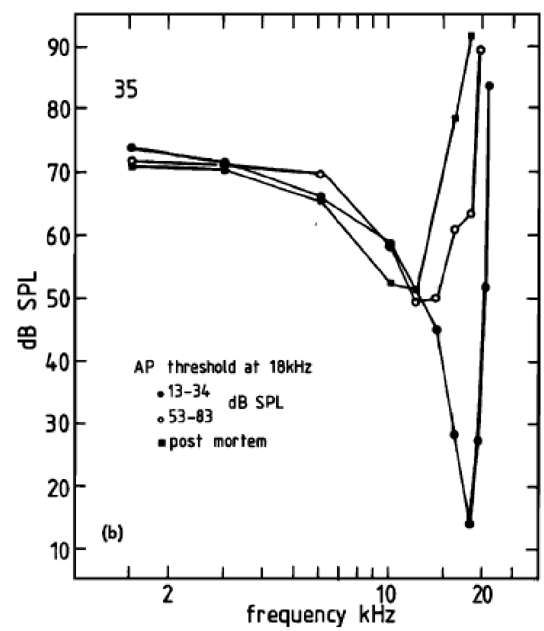


Figure 2: Basilar membrane velocity curves (tuning curves) of 18 KHz characteristic frequency fiber when threshold was normal (filled circles) during the experiment and after the animal's death (filled squares). The basilar membrane tuning curve became less sharp after loss of active mechanism due to outer hair cell loss. Graph from (Sellick, Patuzzi, & Johnstone, 1982) with permission.

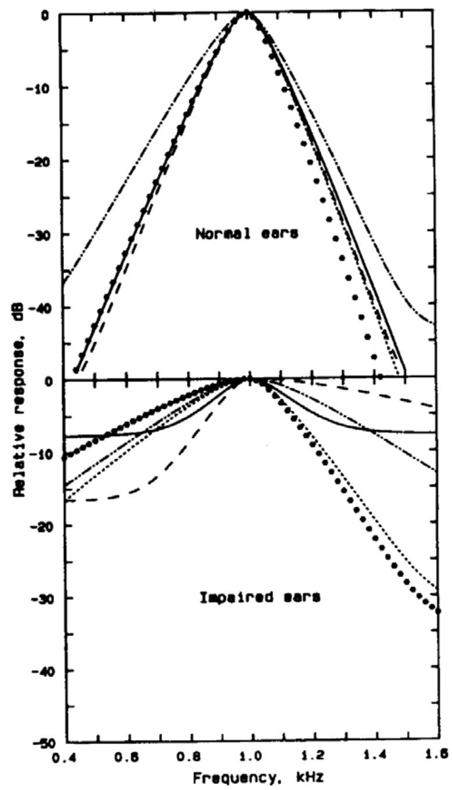


Figure 3: Shapes of auditory filters at 1 KHz characteristic frequency in normal ears (top) and impaired ears (bottom) for six listeners with unilateral cochlear hearing loss. Graph from (Glasberg & Moore, 1986) with permission.

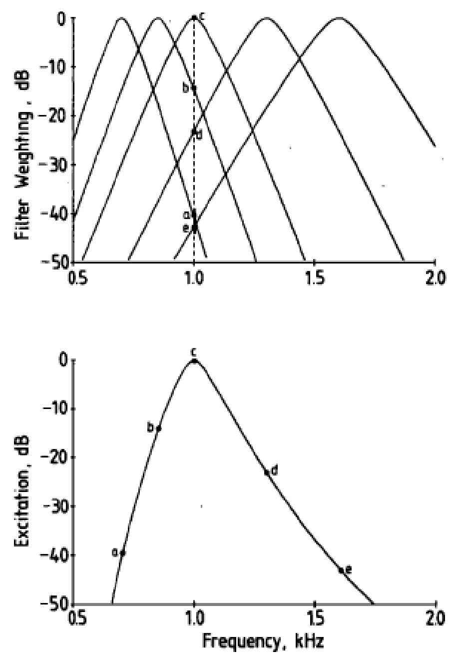


Figure 4: Deriving excitation pattern (bottom graph) for 1 KHz tone (dashed line) from auditory filters (top graph) by calculating the output of each filter as a function of filter center frequency. Graph from (Moore & Glasberg, 1983) with permission.

The third function of the active mechanism is the compressive nonlinearity of the basilar membrane's input/output (I/O) function. Outer hair cells have a motility function that provides level-dependent gain to the BM displacement. The first in vivo measurement of basilar membrane displacement was by Rhode (1971). Low-level sounds (below 30 dB SPL) receive the greatest gain and then, with greater sound levels, the BM grows in a compressive manner (30-90 dB SPL). In other words, the OHCs' gain progressively *decreases* with increasing sound level. The growth of BM response then returns to linear at higher levels (Moore, 2003; Ruggero & Rich, 1991; Ruggero, Rich, Recio, Narayan, & Robles, 1997) (Fig. 5). Without functioning OHCs, the BM response loses its compressive nonlinearity (Fridberger, Zheng, & Nuttall, 2002; Ruggero & Rich, 1991). The responses of AN neurons mirror the BM's responses; therefore, OHC dysfunction modifies AN responses through its impact on BM response.

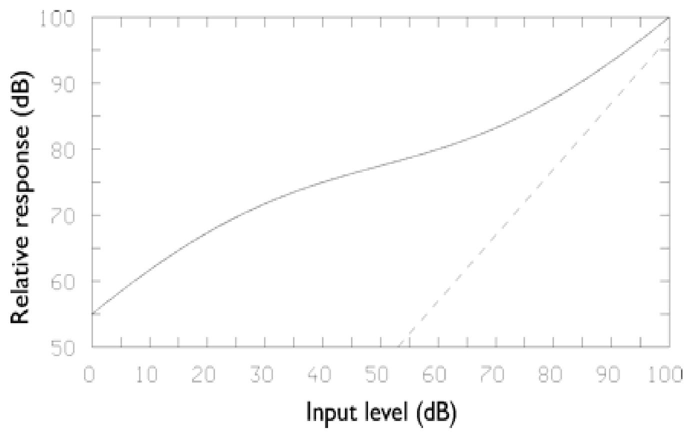


Figure 5: Illustration of the input-output function of the basilar membrane of a tone at characteristic frequency (solid line) and for a tone well below characteristic frequency (dashed line) at which the active mechanism is not involved. Graph from (Moore, 2003) with permission.

The flowchart in figure 6 below summarizes the three cochlear functions that are affected by the loss of the active mechanism: reduced sensitivity to low-level sounds, loss of BM compression, and loss of BM frequency selectivity. It also shows the hair cell type and damage that is responsible.

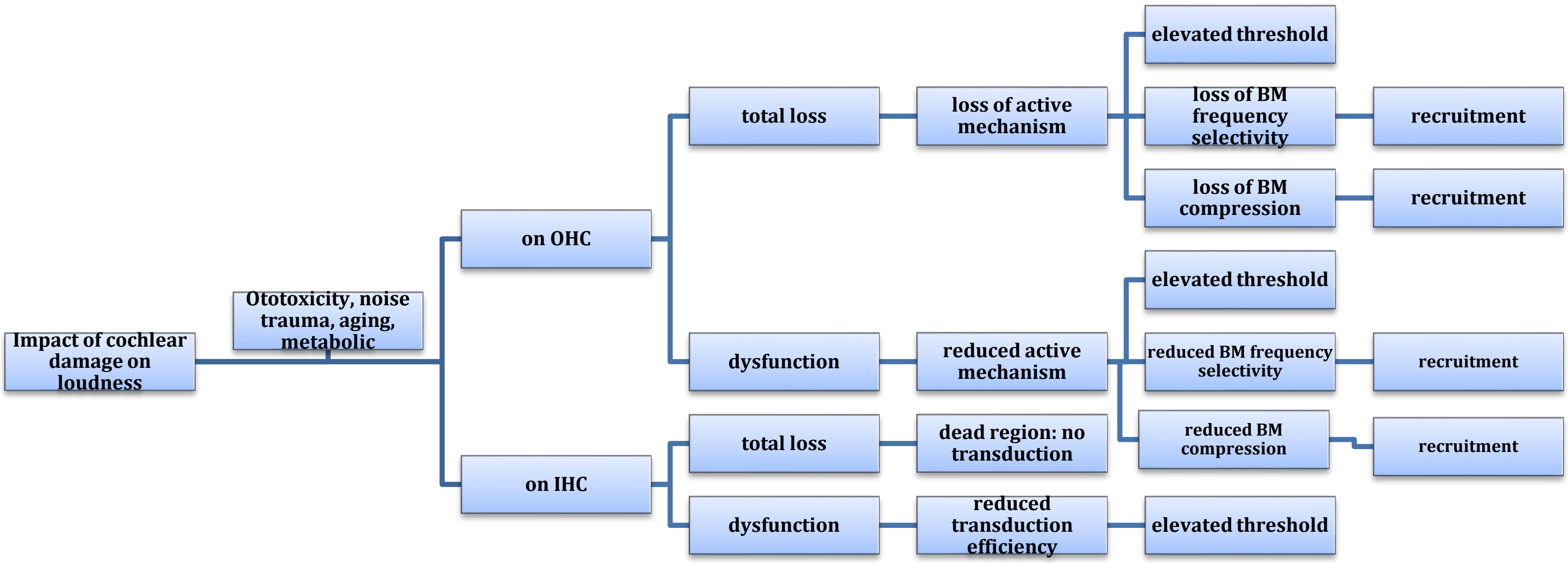


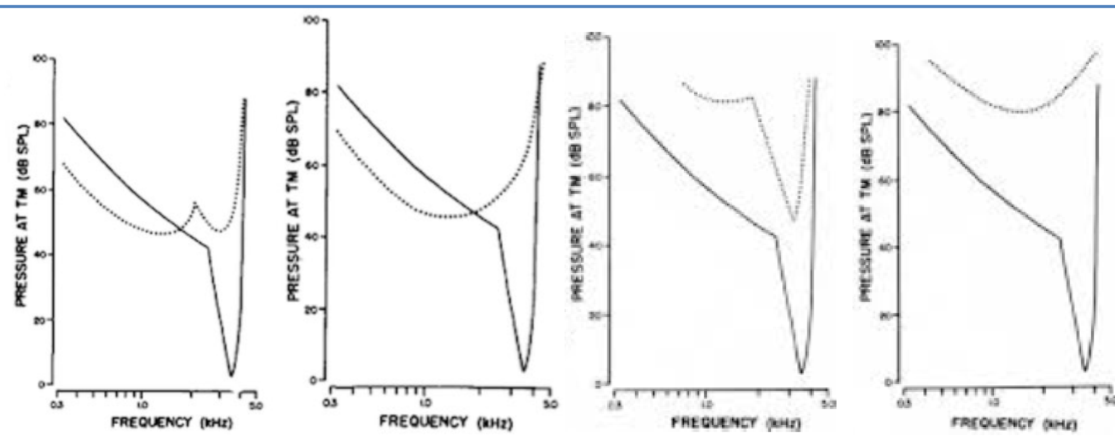
Figure 6: Impact of cochlear damage on loudness.

In most common cases of cochlear hearing loss, the damage is primarily to the OHC. Inner hair cells are less likely to be affected. Outer hair cells are more vulnerable to damage than IHCs. Outer hair cells are more susceptible to metabolic, chemical, and pathological factors (e.g., age-related degeneration, Stria vascularis malfunctioning, exposure to noise trauma and damage due to ototoxic drugs; Patuzzi et al., 1989; Ruggero and Rich, 1991; Ruggero et al., 1992, 1993).

Inner hair cell loss or damage can account for part of a cochlear hearing loss. Inner hair cells act as the transducer in the auditory system. They detect the BM's vibration and convert it into action potentials in the AN. The majority of the signals that are transferred from the ear to the brain are carried by the neurons that are attached to the IHCs. Therefore, IHC damage (without complete loss of function) reduces the efficiency of this transduction process and a higher-level signal will be needed to trigger transduction. Damaged OHCs also lead to elevated absolute threshold. Total loss of IHC in certain regions in the cochlea produces what are called dead regions where no transduction of sound can occur (Moore, 2000). There is typically no response to tones at the frequencies within the dead region. The change in loudness perception with the presence of dead regions in the cochlea is beyond the scope of this review.

Liberman and Dodds (1984a) established a correlation between structural hair cell damage and the resulting changes in AN tuning curve properties. Figure 7 shows the different types and severities of hair cell damage that can occur as a result of various changes in cochlear mechanisms.

Location of cochlear damage	Typical cause	Tuning curve description	Explanation of tuning curve change	Figure
OHC intact IHC intact		Normal: sharply tuned tip Broadly tuned tail		Dark line in graphs
OHC partial IHC intact	Moderate doses of ototoxic drugs	Elevated tip Hypersensitive tail (lower threshold in tail region)	Reduced active mechanism	1
OHC total IHC intact	Large dose of ototoxic drugs	Bowl-shaped curve Lacks the sharp tip Broad curve	Destroyed active mechanism Depends on passive mechanism	2
OHC stereocilia partial IHC stereocilia total		Normal sharp tip Tip and tail both shifted upwards by about 40 dB	Active mechanism intact Reduced sensitivity of transduction mechanism	3
OHC stereocilia total (first row) IHC stereocilia total	Acoustic trauma	Lacks the sharp tip Whole curve is shifted upward (loss of sensitivity)	Destroyed active mechanism	4



1 **2** **3** **4**

Figure 7: The impact of different types of hair cell loss on auditory neuron tuning curves. Figure adapted from Liberman, Dodds, and Learson (1986). Graphs from Liberman et al. (1986) with permission.

The above-discussed changes in cochlear functions are what are believed to change the perception of loudness in listeners with hearing loss. The impact of each of these effects on loudness perception will be described in the following section.

Effects of cochlear damage on loudness perception

Changes in loudness perception with hearing loss are largely related to OHC loss and, consequently, reduction or loss of the active mechanism in the cochlea. For cochlear hearing losses that are moderate in severity (i.e., threshold is elevated by less than 50 dB), the main cause of the hearing loss is OHC damage. With increasing severity of hearing loss (i.e., when absolute threshold is elevated more than 50 dB) complete loss of OHCs is expected and IHC damage is probably more involved (Moore, 2000). Individual differences in the amount of OHC and IHC damage between listeners might account for variations seen in loudness perception and loudness growth among listeners with similar absolute hearing thresholds (Moore, 2007).

The impact of cochlear damage on the perception of loudness has always been discussed in terms of the loss of sensitivity to low-level sounds and the concept of recruitment. More details on this altered perception and its underlying physiology are provided in the following section.

Increased absolute threshold

As discussed above, the elevation in perceptual threshold corresponds to the elevated thresholds of auditory neurons due to the dysfunction or loss of OHCs, which reduces or eliminates the active mechanism, or the dysfunction or loss of IHCs.

Recruitment

The concept of recruitment describes how loudness grows with increasing input level above the elevated threshold within the reduced dynamic range. It is defined as an abnormal growth of loudness above elevated threshold (Moore, 2007). A more detailed definition is as follows: Once a sound is increased in level above the elevated absolute threshold due to cochlear hearing loss, the rate of growth of loudness with increasing signal level is greater than normal, but grows to near normal at high intensities (around 100 dB SPL). The recruitment phenomenon was first noted in 1928 by Fowler when he found that loudness was equivalent at 5 SL in the impaired ear and 10-25 SL in the normal ear of the same listener with unilateral hearing loss. When loudness at high level in an impaired ear reached normal loudness perception, Fowler described this phenomenon as a boosting effect (1936). In an attempt to explain this phenomenon within a physiological model, Fowler (1937) claimed that even with the loss of some hair cells, a strong enough stimulus would still elicit the same number of impulses from the ear to the brain because each nerve fiber connects to a number of hair cells. Fowler's physiologic definition did not include what we currently know about cochlear functioning, such as the separate innervations of inner and outer hair cells and the active function of outer hair cells. Current understanding of the cochlear compressive mechanism due to OHC motility and basilar membrane mechanics contributed to a more accurate description of recruitment as being due to the loss of cochlear compressive nonlinearity resulting in a steepening of the BM input-output (I/O) function.

Many physiological models attempt to explain the recruitment phenomenon. It is apparent in the literature that the underlying physiology of recruitment is not fully understood. The majority of loudness recruitment models concur that recruitment reflects the loss or reduction of the active mechanism of the cochlea, i.e., the loss or dysfunction of OHCs. Researchers report that loudness growth with increasing signal level in human listeners corresponds well to the BM displacement in animal experiments, both with normal and impaired hearing (Fridberger et al., 2002; Moore, 1995; Ruggero & Rich, 1991; Schlauch, DiGiovanni, & Ries, 1998). The mechanism by which this change in BM response is conveyed as a neural correlate at higher levels of the auditory system remains unclear. Research groups seem divided in terms of what is believed to be the location and the physiology of the neural correlates of recruitment.

Neural correlates of recruitment at the peripheral level

Since it is commonly thought that the loudness of a sound reflects the total activity of the auditory nerve (AN), it is logical to assume that loudness recruitment also is coded at that level. Models that explain neural correlates of recruitment at the AN level have hypothesized three different physiological changes to explain the faster growth of loudness with increasing level in listeners with OHC loss. The three hypothetical changes are schematically illustrated in Figure 8. The first change hypothesizes a steeper rate-level function (firing rate change as a function of sound level) of AN fibers post cochlear trauma, as would be expected from the loss of the BM compressive nonlinearity (Harrison, 1981; Yates, 1990). The second change hypothesizes a faster spread of excitation as would be expected from the increased width of auditory filters (reduced frequency selectivity) (Evans, 1975; Kiang, Moxon, & Levine, 1970). Faster spread of excitation implies that a smaller increase in sound level is required to recruit more neurons from the neighboring auditory filters, hence the name recruitment. The third change hypothesizes a compression of the spread of AN fiber thresholds as a result of cochlear trauma (Moore, Glasberg, Hess, & Birchall, 1985; Zeng & Turner, 1991). In a healthy cochlea, auditory neuron fibers tuned to similar frequencies present with a variety of thresholds and dynamic ranges which are important for sound-level coding. The decrease in this spread of thresholds would cause a steeper growth of firing rate with increasing level.

The relationship between acoustic stimuli and the perceptions they elicit has been quantified by human psychoacoustics. In contrast, the relationship between the acoustic stimuli and neural processing has been largely limited to animal studies due to the invasive nature of this type of research.

Studies that have looked at neural correlates of recruitment at the AN do not provide a consistent answer. Table 2 lists some of the studies that have examined the first two change hypotheses, which are more commonly studied, and also shows the inconsistency of their findings.

Studies that have examined the steepening of AN rate-level function have shown inconsistent results. Studies that have not shown a steepening function attribute their findings mostly to two reasons (e.g., Heinz & Young, 2004): first, mixed lesions of both OHCs and IHCs are expected with acoustic trauma (Liberman & Kiang, 1984) and as a result a steepening of rate-level function might not be seen due to the contribution of IHCs damage that shifts the absolute threshold without the loss of BM compression, which in turn might result in shallower than normal rate-level function. Thus, the contribution of IHC lesions complicates the situation. Second, it is not expected that all AN fibers show a steeper response with hearing loss. Only fibers with high threshold (low spontaneous rate) should show a steeper function with cochlear damage given that they were the ones responsible to code the compressed range of sound levels. Low threshold fibers (with high spontaneous rate) are responsible for coding low-level sounds and have linear rate-level function to begin with.

Most of the existing research suggests that recruitment is caused by a steeper input/output function on the BM in abnormal cochleas, which is neurally coded as an increase in AN rate-level function slope, and that the greater spread of excitation has only a minor role in the recruitment phenomenon, contrary to what is suggested by Evans (1975); Kiang et al. (1970).

Changes in temporal aspects (i.e., phase locking), of the response across the AN fibers' population may have an important role in sound level coding. This coding mechanism is thought to be heavily dependent on the presence of compressive nonlinearity in the cochlea (Carney, 1994). It has been proposed that the loss of the compressive nonlinearity after cochlear damage alters these temporal patterns and causes recruitment (Carney, 1994; Heinz, Issa, & Young, 2005). While a few studies found that phase locking was adversely impacted by OHC damage (Woolf, Ryan, & Bone, 1981), others did not report any impact (Harrison & Evans, 1979). When phase locking was impacted, the impact was found on auditory neurons with CFs corresponding to the frequencies where the behavioral threshold was elevated by at least 40 dB. The highest frequency at which phase locking occurred was reduced, and below that frequency, the precision of phase locking over the frequency range also was reduced. Different reasons for why phase locking is impacted by cochlear hearing loss have been suggested by different researchers, but the reason remains unclear. The most frequently proposed reason is the poor mechanical coupling between the tallest stereocilia of the OHC and the tectorial membrane (Woolf et al., 1981).

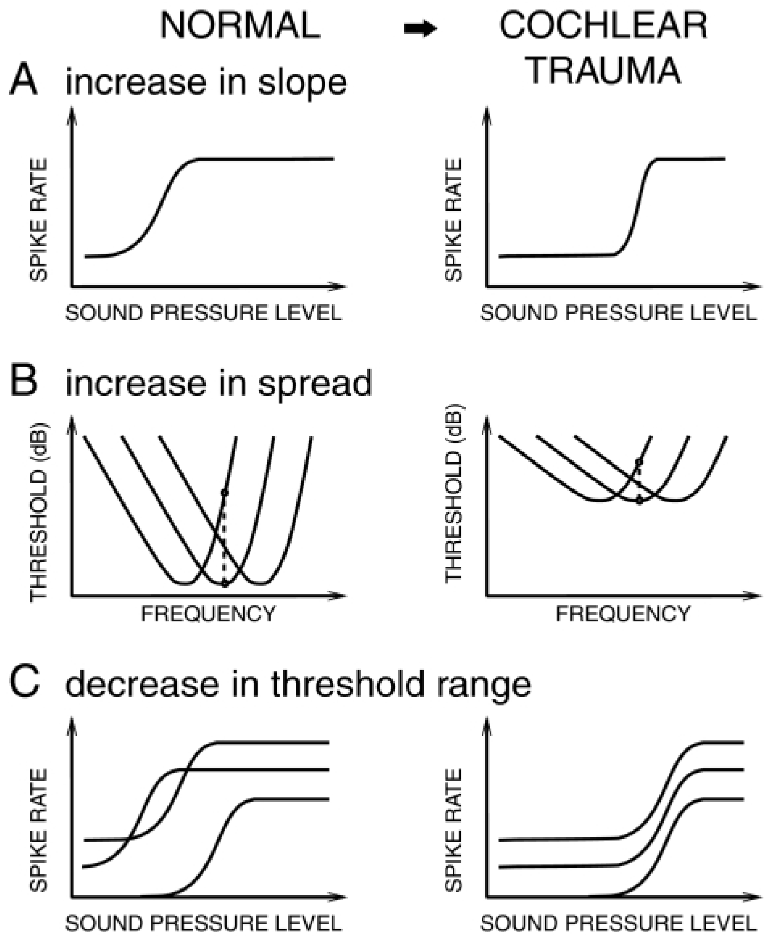


Figure 8: The three hypotheses of the neural correlates of recruitment at the auditory nerve (AN) level. Normal AN fibers response (left) and impaired AN fibers response post cochlear trauma (right). Impaired AN fibers show: A. increase rate-level function slope. B. Increase spread of excitation and as a result a smaller increase in sound level is required to recruit fibers from neighboring filters. C. Decrease in range of AN fibers' thresholds. Graph obtained from (Joris, 2009).

Table 2: Some of the studies that investigated recruitment's neural correlates at the peripheral level.

	Supports the theory?	Study	Purpose/method	Results
If recruitment is caused by loss of BM nonlinearity, it is coded as steeper rate-level function	Yes	Yates (1990) could not access yet	Modified an intensity-rate function model for normal hearing to account for HL by reducing BM nonlinearity.	The modified model showed an increase in rate-level function slope with increasing level
		Harrison (1981)	Compared AN fibers' rate-level function obtained from normal and ototoxic guinea pigs, and from human with normal and pathological cochlea	The slope of rate-level function of fibers near to CF become as steep as those for the tail
	No	Heinz & Young (2004)	Looked at rate-level function of AN fibers in cats with acoustic trauma	Obtained AN rate-level function is not steeper than normal
		Kiang et al (1970)	Obtained AN rate-level function from cats with ototoxicity	Data reported no increase in slope of rate-level functions
If recruitment caused by broader auditory filter, it is coded as increased spread of excitation	Yes	Evans (1975)	Obtained AN fibers' tuning curves from cats with hypoxic cochlea	Proposed a model to explain recruitment based on the assumptions: 1) abnormal cochlea have degraded cochlear tuning. 2) loudness is related to number of active fibers in the AN.
		Kiang et al (1970)	Recorded AN fibers in cats with ototoxic cochlea	Found an abnormal growth of the extent of activity in the AN
	No	Moore, Glasberg, Hess, (1985)	Presented a signal with a notched noise that is designed to mask neighboring neurons of the signal's CF to human listeners with unilateral hearing loss	Observed recruitment was not reduced indicating that recruitment is not a phenomenon of increased spread of excitation

Neural correlates of recruitment at the central level

All the above-mentioned models have a common underlying assumption that loudness recruitment can be fully accounted for by changes in cochlear mechanisms. However, inconsistent findings of the relationship between peripheral neural correlates of recruitment and abnormal rate of growth of loudness, i.e., psychoacoustic measures, have led some researchers to conclude that neural correlates of recruitment may be more pronounced at higher synaptic or cortical levels in the central nervous system (e.g., Boettcher & Salvi, 1993; Cai, Ma, & Young, 2009; Heinz et al., 2005; Langers, van Dijk, Schoenmaker, & Backes, 2007; Morita et al., 2003; Phillips, 1987; Salvi, Wang, & Ding, 2000). This line of research has challenged the underlying peripheral neurophysiology of recruitment and suggested that changes in the input to higher levels of the auditory pathway due to the hearing loss may alter the way these levels handle the signal and recruitment might actually be a manifestation of these alterations (Boettcher & Salvi, 1993). It is also possible that reduced input from the periphery to the higher auditory centers due to hearing loss changes the synaptic gain and/or connection strength as an adjustment to overcome and compensate for this reduction. This concept of hyperexcitability provides a possible mechanism for recruitment.

In the last two decades, studies have shown enhanced neural activity and increased rate of growth of response amplitude in certain regions of the central auditory pathway, e.g., ventral and dorsal cochlear nucleus (VCN, DCN) and inferior colliculus (CI) (Cai et al., 2009; Salvi et al., 2000) following cochlear damage. Advances in brain imaging technology have allowed researchers to identify regions of hyperactivity in listeners with recruitment, as well as other related abnormal perception phenomena that are common symptoms of cochlear damage such as tinnitus and hyperacusis.

Langers et al. (2007) looked at brain cortical activations in humans using functional magnetic resonance imaging (fMRI) in response to increased sound levels in listeners with normal hearing and others with hearing loss and found that auditory cortical activations in the temporal lobes were strongly related to the loudness level of stimuli, as opposed to the intensity level. The authors found a larger increase in activation with increasing level in listeners with hearing loss compared to normal listeners. Similar findings have been reported using magnetoencephalography (MEG) (Morita et al., 2003).

Cai et al. (2009) conducted another study that looked neural correlates of recruitment at the central level. The authors recorded the rate-level function of neurons in the VCN of normal cats and cats with acoustic trauma. Elevated maximum firing rates and steeper rate-level functions in one class of VCN neurons were found in traumatized cats compared to normal ones.

Summary

Figure 9 below summarizes most of the different theories that are studied in the literature of neural correlates of recruitment across the auditory pathway.

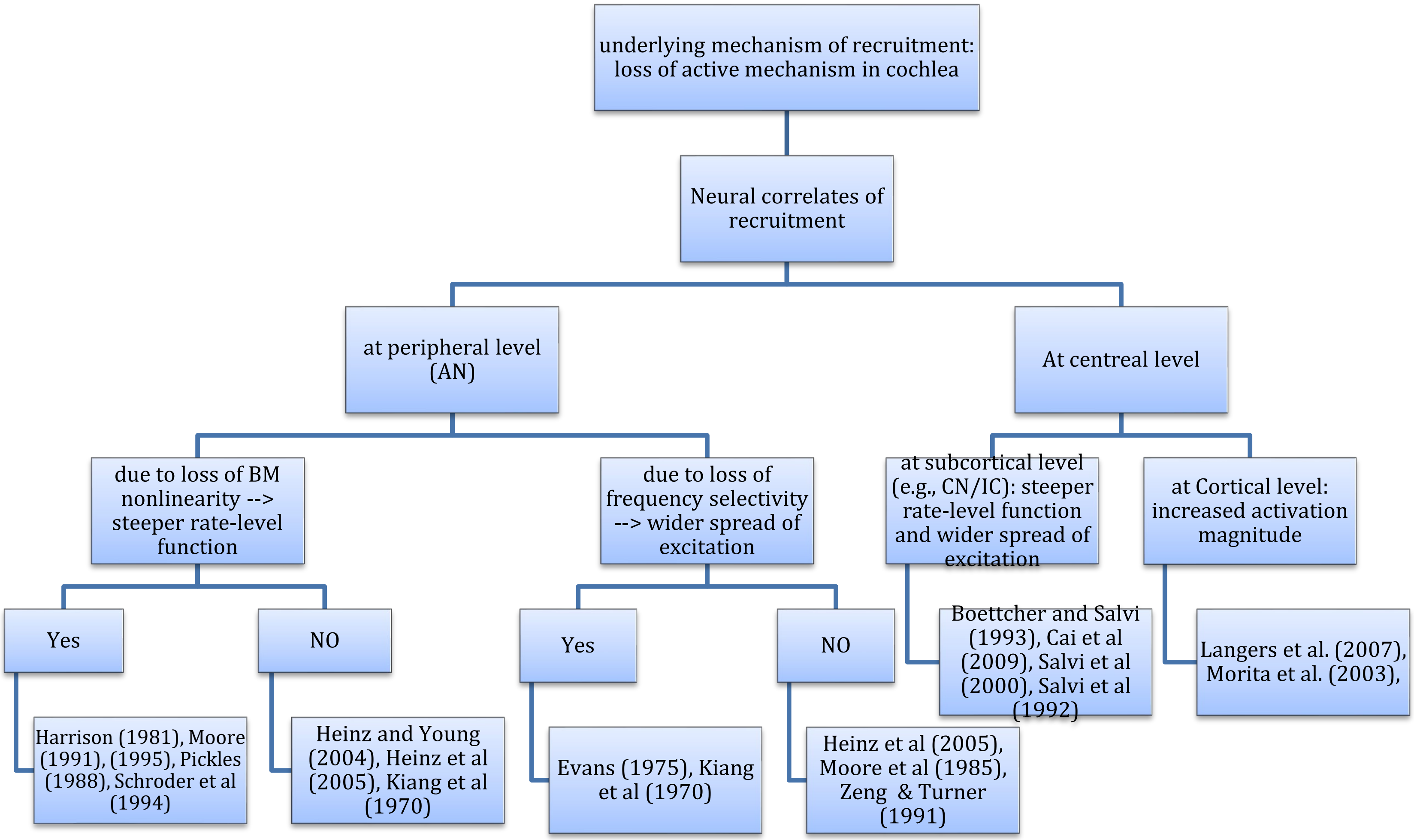


Figure 9: Neural correlates of recruitment.

Relation between the study of neural correlates of recruitment and loudness at threshold

The current views that describe how listeners with hearing loss perceive loudness at and near their elevated thresholds are controversial. One view is in line with the classical description of recruitment that loudness at elevated threshold is equal to loudness at normal thresholds and that loudness grows abnormally fast above the elevated threshold to catch up with normal loudness at high levels (Moore & Glasberg, 1997; Moore & Glasberg, 2004). This concept is referred to, at least in this review, as rapid growth concept (RG). The other view is more recent and describes loudness at elevated threshold as larger than normal and that loudness grows at a normal range 10-15 dB above threshold (Buus & Florentine, 2002; Florentine et al., 2005). The concept is referred to as softness imperceptions (SI). The argument regarding the change in loudness at elevated threshold cannot be disambiguated based on any of the two most studied theories of the neural correlates of recruitment. However, it is possible to consider the third, less studied theory of compressed AN fibers' thresholds distribution in impaired ears (Moore et al., 1985; Zeng & Turner, 1991) (Fig. 8, C). Although not always seen in animal model studies (Heinz et al., 2005), this concept might support the louder perception at elevated threshold. In normal AN, neurons with high spontaneous rates and low threshold are the ones responsible for coding low-level sounds. Loudness grows in a linear manner with increasing level at the levels corresponding to these neurons' linear rate-level function. With cochlear trauma and increased absolute threshold, the response range of these different fibers overlap. High spontaneous-low threshold neurons will have an increased threshold that is similar to other types of neurons' thresholds. If these fiber populations contribute in different ways to the neural coding of loudness, and there is an increase in number of neurons responding to sounds at threshold, a larger loudness might be perceived at threshold.

Larger than normal loudness at elevated thresholds might be explained also in the case of selective loss or dysfunction of high spontaneous-low threshold fibers, which are the most vulnerable to damage than other fibers. As suggested by Florentine (2011), elevated threshold will be determined mostly by lower spontaneous, higher threshold fibers that normally code for higher loudness. If the activity of these neurons is still interpreted by the brain as indicating high loudness, loudness at elevated threshold might be higher than normal.

While studies of neural correlates of recruitment examined the changes in rate-level function and in the firing rate, these physiologic data have never been compared to psychoacoustic data, subjective loudness, obtained from the same listeners. While obtaining subjective loudness data from animals is challenging, it confirms that the listeners that have been studied actually have experienced the recruitment phenomenon as described by human listeners, and might also provide quantitative data to the limited pool of data for loudness at and above threshold. These data are lacking in the loudness literature.

All above discussed hypotheses of recruitment have been postulated, but not fully validated, so there is room for proposing modifications to the existing hypotheses.

The comparison of loudness at elevated thresholds with loudness at normal thresholds did not arise until 2002, when Buus and Florentine added more detail to the notion of recruitment when they reported that loudness near elevated threshold due to cochlear pathology grows at similar rates as loudness near normal threshold (Buus &

Florentine, 2002). They proposed that for the impaired system, the recruitment phenomena might actually start with an abnormally larger loudness at threshold compared to normal loudness (very small loudness) at normal thresholds. There has not been a neurophysiological study to date that specifically addresses these different views of loudness at threshold.

2.1.3.3 Loudness coding at the central level

Every stage in the ascending auditory pathway analyzes and processes auditory information. The auditory signal leaves the cochlea and travels through the auditory nerve to the brainstem. In the brainstem, the signal passes through several nuclei stations starting with the cochlear nucleus (CN), and moving to the superior olivary complex (SOC), the lateral lemniscus (LL), and the inferior colliculus (IC). The auditory information is then passed to the primary auditory cortex via the medial geniculate body (MGB) of the thalamus.

More recent literature has investigated the contribution of central processes to the perception of loudness. This occurred in response to physiological studies of loudness that did not show agreement between the peripheral response and response growth and reported loudness. Pickles (2012, p. 292) stated, “it is highly likely that transformations of the auditory stimulus by the central nervous system play a critical part in determining the loudness of stimuli”.

Neuroimaging studies of the neural correlates of sound intensity in the human central auditory system have demonstrated an increase in neural activation magnitude in auditory cortical (e.g., Hall et al., 2001; Hart, Palmer, & Hall, 2002; Röhl & Uppenkamp,

2012) and subcortical (CN, SOC, IC, MGB) (Sigalovsky & Melcher, 2006) areas as a function of sound level.

For example, using functional magnetic resonance imaging (fMRI), Röhl and Uppenkamp (2012) looked at individual differences in the magnitude of the responses to noise and compared them with individual differences in the judgments of loudness. They found that when comparing different individuals, variation in the magnitude of the fMRI signal in the auditory cortex was related to variation in the subjective magnitude of the loudness. The same correlation was found between average response magnitude data and loudness scaling (Hall et al., 2001). This relationship was not observed at lower levels of the auditory system, where the responses were more closely described by the physical parameters of the stimulus than by its subjective loudness. Primary auditory areas and especially Heschl's gyrus, are more sensitive to sound level changes than secondary auditory areas in the auditory cortex, which suggests larger involvement in sound-level processing (Hart et al., 2002).

Two of the limitations in neuroimaging studies are: 1) They do not look at cortical activation in response to low-level sounds (lower than 60 dB SPL), which have not been the focus of these studies; and 2) They examine the response behavior for only a limited range of sound levels presented (i.e., few tens of dB SPL), which is a small part of our dynamic range that might not provide a full picture of cortical neural coding of loudness or the response curve slope.

Summary

To date there is no single comprehensive theory that incorporates all of the above-mentioned physiological and neurophysiological mechanisms to explain our ability to perceive loudness. Despite the current enhanced understanding of acoustics, the physiology of hearing at the cochlear level and the auditory nerve fiber response, the literature is still not complete with regard to loudness perception. Peripheral neural correlates of loudness have not always successfully explained subjective perceived loudness. In contrast, cortical neural correlates, have so far almost always matched subjective loudness. We do not yet know whether peripheral neural correlates interact to code loudness or whether the final determinant for perceived loudness is at the cortical level. Further research into loudness coding is needed.

2.1.4 Psychoacoustic Models of Loudness: Modeling Perception of Soft Sounds

"The best material model of a cat is another, or preferably the same cat". This is what Norbert Wiener once said as cited in De Cheveigne (2005, p. 3), a chapter on pitch models. However, psychoacousticians generally disagree with this approach because in their view, the real goal of modeling is to simplify that cat. Since loudness perception is determined by the physical characteristics of the sound, as well as the status of the listener, loudness models are limited to predicting loudness based on the variables that can be manipulated in laboratory research, namely the input to the model (physical characteristics of sound) and to a certain extent the sensory status of the listener (integrity of the auditory system). The psychological states of the listener (attention, fatigue, alertness, etc.) are all variables that researchers try to control for and bypass to predict average loudness. These loudness models can appropriately be used for general application as they attempt to calculate the average perceived loudness by a large group of listeners under the least biased conditions (Moore, 2001).

Models of loudness were developed as a consequence of discoveries about loudness perception and its complex interaction with physical properties of sound, and the integrity of the hearing mechanism. Loudness models attempt to incorporate all that is currently known about the physiology of intensity coding, and can be divided into two types: 1) psychoacoustic models that describe and predict the relationships between the stimulus and the perceived subjective loudness; and 2) physiologic models that correlate changes in the stimulus level and the physiologic response to these changes.

The physiologic models of loudness were previously reviewed in section 2.2 of this paper. As discussed earlier, physiologic models of loudness in cases of cochlear hearing loss have explained the recruitment phenomena; however, they have not been extended to calculate subjective loudness functions, i.e., model psychoacoustic data (Launer, 1995). In contrast, the following section will review psychoacoustic models that can predict behavioral data.

Psychoacoustic models of loudness are those that calculate and predict the average loudness that would be perceived by a large group of listeners with normal hearing under conditions where biases are minimized as far as possible. These models are very useful because they act as sound level meters that measure loudness, which is a totally subjective attribute. If we examine the history of modeling the perception of loudness, back in the 1800s, loudness models started as simple formulas that predicted the subjective perceived loudness directly from the physical magnitude of the sound. These models were developed over the years and gradually increased in their complexity and sophistication by adding more variables to the input of the model (e.g., spectral information) and by incorporating the physiological models to more precisely predict loudness of sounds not only from intensity, but from spectral properties (Fastl & Zwicker, 2007; Florentine & Zwicker, 1979; Glasberg & Moore, 2002; Moore & Glasberg, 1996; Moore et al., 1997; Zwicker, 1977).

2.1.4.1 Modeling loudness of pure tones: Loudness functions

Loudness functions for listeners with normal hearing

The earliest effort to predict loudness was started by Fechner in 1860. Based on Weber's law of Just Noticeable Difference (JND), Fechner assumed that each increase in sensation is an identical increase in magnitude (Fig. 10, left panel). His model of loudness was a logarithmic function where loudness has a logarithmic relation with sound intensity (i.e., a linear relation with sound level in dB SPL) (Fechner, 1948). Stevens (1957) in his seminal paper, described the relation between loudness and sound level in dB SPL as a power law function where the magnitude of the subjective loudness was directly proportional to the sound level raised to the power of an exponent (Fig. 10, middle panel) (Stevens, 1957, 1961). In contrast to Fechner's logarithmic function, Stevens' power function was based on the assumption that each JND produced an increase of sensation that was not constant, but proportional to the original sensation. Stevens' data were derived using magnitude estimation and magnitude production procedures. One limitation of Stevens' power law was that it only applied to sound levels above 40 dB SPL. Loudness of lower level sounds changes more rapidly than predicted by the power function.

In 2006, Florentine and colleagues proposed a newer loudness function called the inflected exponential (INEX) (Florentine & Epstein, 2006). The INEX function deviates from the power function in two main variations that cause it to have different slopes across the level range (Fig. 10, right panel). These variations are based on data obtained from different studies (Buus & Florentine, 2001; Buus, Florentine, & Poulsen, 1997; Buus et al., 1998; Hellman & Zwislocki, 1961). The first variation is that the function is steeper at and near threshold than at moderate levels, meaning that for sounds below 30 dB SPL, loudness grows more rapidly. This steeper slope at low-levels has been previously supported (Buus et al., 1998; Hellman & Zwislocki, 1961, 1963). The INEX function suggests that near threshold, the average slope is about unity or slightly larger and as level increases, the slope decreases to 0.36 at 20 dB SL and 0.19 at 40 dB SL.

The second way INEX is different from the power function is that its function is less steep at moderate levels than at low and high levels, which means that loudness grows in a more compressive manner between 30-60 dB SPL, followed by an increased slope at higher levels. Data from studies of spectral and temporal integration of loudness, as well as binaural summation supported this concept (Florentine, Buus, & Poulsen, 1996). The INEX function is claimed by Florentine and Epstein (2006) to be the best description of current behavioral data.

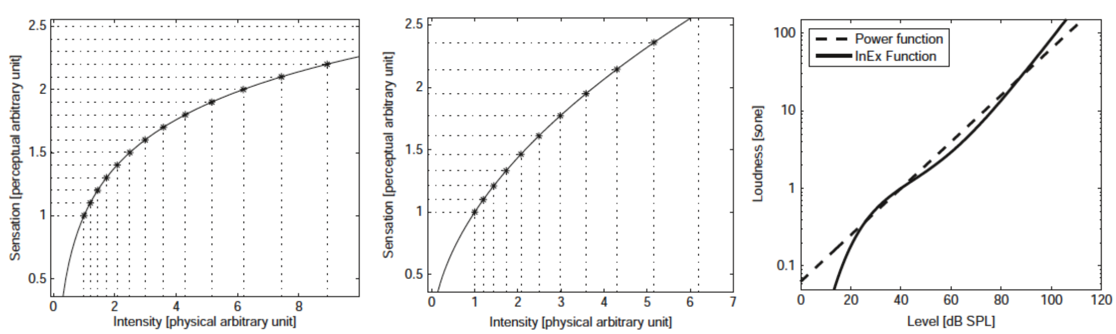


Figure 10: The main three loudness functions for normal hearing listeners proposed over the years: Fechner's logarithmic loudness function (left), Stevens' power law loudness function (middle), and INEX loudness function (right). Graph obtained from Florentine (2011) with permission

Loudness functions for listeners with hearing loss

When we examine loudness functions for listeners with hearing loss, we must bear in mind that a reduced dynamic range is a unique characteristic of sensorineural hearing loss. The intensities perceived by listeners with hearing loss ranges from the level of elevated threshold to about 100 dB SPL, where sounds become uncomfortable. The rate at which loudness grows within this reduced dynamic range is more rapid compared to normal hearing listeners who have a full dynamic range. This rapid growth of loudness above elevated threshold is referred to as recruitment. There are lines of evidence that show that listeners with hearing loss of cochlear origin exhibit a wide range of loudness growth functions that are neither similar to those of normal hearing listeners, nor simply predictable from their threshold (Hellman & Meiselman, 1990; Marozeau & Florentine, 2007; Moore & Glasberg, 2004).

Rapid Growth (RG)

Recruitment has been described for decades as abnormally rapid growth of loudness above elevated thresholds while loudness for high-level sounds is close to normal. Loudness models originally predicted this phenomenon by predicting a markedly greater slope of loudness growth function at levels near threshold for ears with cochlear hearing loss than for normally hearing ears (Moore & Glasberg, 1997).

It was not until recently that the loudness function was adjusted to show that the rate of growth of loudness for levels very close to threshold (3-4 dB above threshold) is similar for normal and impaired ears. This change reflects empirical data from several studies of loudness perception at low levels (Buus & Florentine, 2002; Hellman, 1999;

Hellman & Meiselman, 1990; Moore & Glasberg, 2004). Beyond the 3-4 dB above threshold range, the curve for impaired ears becomes steeper than that for normal ears. Hence, the term Rapid Growth (RG) is used to describe this function.

2.1.4.2 Modeling loudness of complex sounds

With the discoveries of the auditory critical bandwidth and spectral loudness summation, researchers also developed models that predict loudness not only from the sound's intensity but also from other physical parameters of sound such as spectral content. In contrast to simple loudness functions, such models can react with precision to complex environmental sounds. Known processes of the auditory system also were integrated into these models to better predict loudness. Pioneers in this area were Zwicker and more recently Moore and colleagues. Most models are arranged as cascade filters, which means that the output of a filter is an input to the next one.

Many models of loudness perception have been developed over the years. Loudness models that most commonly cited in the literature are included in this section. Details regarding the way these models have been modified to account for hearing loss will be reviewed in this section.

Zwicker's Model

One of the first models of loudness perception for hearing loss was developed by Zwicker in 1960s. One might think of Zwicker's model as the first attempt to link the physiological and psychological models of loudness, combining the concepts of critical bands by Fletcher (1940) and power law by Stevens (1957). Zwicker's model initially calculated loudness of steady-state sounds for normal hearing listeners. Based on the spectral analysis of the sound, the sound is passed into different auditory filters. The output of each filter is then processed in a nonlinear way by applying the power law function. Application of the power law allows us to calculate the specific loudness patterns. The concept of specific loudness is what was new in Zwicker's model (Zwicker & Scharf, 1965). The overall loudness is then calculated as the area under the specific loudness patterns taking into account all of what was known about spectral interactions, such as spectral masking and loudness summation (Zwicker, 1961; Zwicker & Fastl, 2013; Zwicker & Scharf, 1965).

When Zwicker's model was extended to account for time-varying sounds (Zwicker, 1984), it became the normative standard in the ISO 532B (Scheuren, 2014; Zwicker & Fastl, 2013). Zwicker's model, as well as other loudness models for complex sounds were based on the integration of the specific loudness that is derived from excitation patterns. Taken together, the model's three main stages are: 1) The transformation of the signal's long-term power spectrum into a modified discrete Fourier transformation that mimics the cochlear auditory filters (Fig. 11, a, b, and c). This stage integrates the outer and middle ear transform functions (Fig. 11, d) and ends with calculating the level and the spread of the excitation pattern for each critical band (Fig. 2.11, e). 2) Given the spectral masking due to the level differences and the overlap between the excitation patterns, this stage converts the excitation patterns into specific loudness which represent the contribution of each band to the overall loudness (Fig. 11, f and g). This conversion is based on Stevens' power function described in the previous section. The transformation of specific loudness to overall loudness in sons or phones by summing across specific loudness in a mathematical formula (Fig. 11, h).

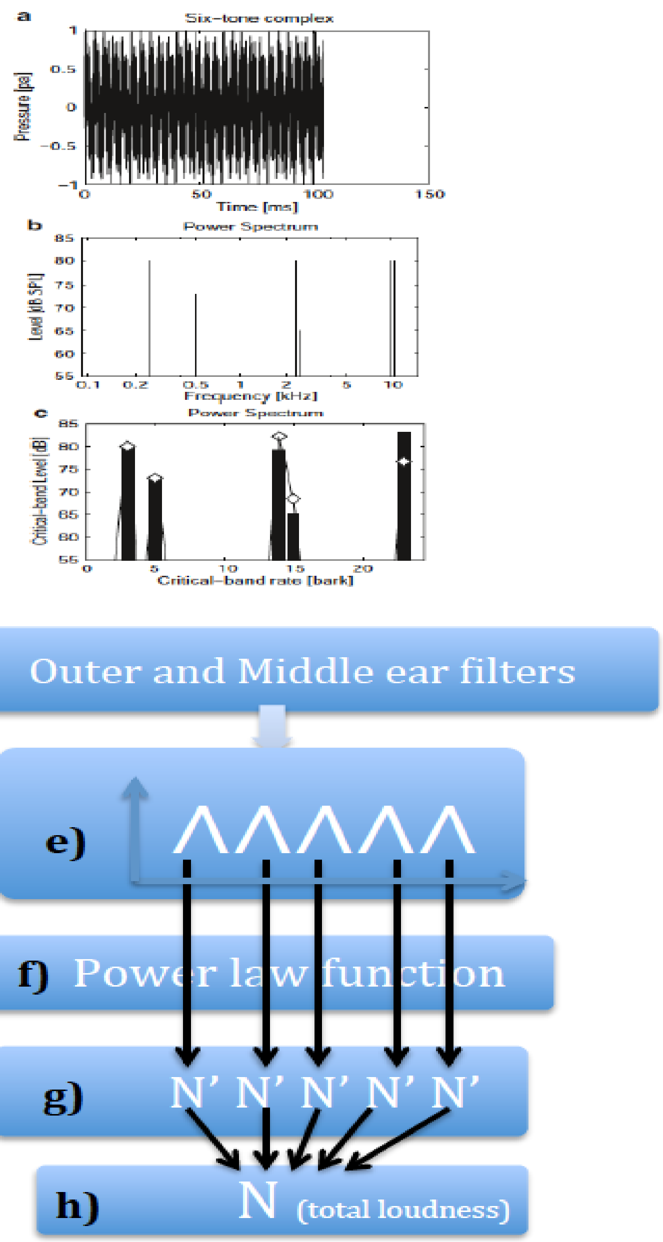


Figure 11: Stages of Zwicker’s loudness model for complex sounds. a,b) The signal is represented in temporal and spectral domains; c) the energy is summed in each critical band corresponding to the complex components; d) a linear filter mimicking the outer and middle ear transfer function; e) excitation patterns are calculated from auditory filters; f) Excitation pattern levels are transferred into specific loudness N’ (g) by the power law function relationship; (h) Overall loudness N is calculated by summing the specific loudness across auditory filters. Graph adapted from Launer (1995) and Florentine (2011).

As mentioned earlier, the concept of specific loudness is essential in loudness models for complex sounds. Instead of using the total loudness as a direct interest, the specific loudness that is produced at different critical bands are used. Specific loudness is developed using Steven's power function. As the power function uses the intensity to predict loudness in sones, the specific loudness transfer function uses excitation level to predict loudness in sones per critical band.

Zwicker's model was revised by multiple research groups to account for listeners with hearing loss by modifying certain parameters of the model (Chalupper & Fastl, 2002; Florentine & Zwicker, 1979; Launer, Hohmann, & Kollmeier, 1995). The main modifications to the model are frequency dependent attenuation of excitation level per critical band to mimic reduced sensitivity (Chalupper & Fastl, 2002; Florentine & Zwicker, 1979; Launer et al., 1995), broadening auditory filters to reflect reduced frequency selectivity (Florentine & Zwicker, 1979), and modifying the power function that calculates the specific loudness to reflect reduced compression by increasing the exponent in the formula (Chalupper & Fastl, 2002; Launer et al., 1995).

Moore et al., Model

The loudness perception model developed by Moore and colleagues in 1996 is based on concepts proposed in previous models by Fletcher and Munson (1933, 1937) and by Zwicker and Scharf (Fletcher, 1937; Zwicker, 1956, 1958; Zwicker & Fastl, 2013; Zwicker & Scharf, 1965). This model shares the basic concept for modeling loudness with Zwicker's model, that is, the loudness of a sound is related to the sum of loudness contributions from different longitudinal places along the BM, expressed mathematically as the sum of loudness contributions per normal auditory filter, or what is usually called specific loudness (Florentine, 2011; Moore & Glasberg, 1996; Moore & Glasberg, 2004; Moore et al., 1997).

The Moore et al model's initial and basic form is composed of five stages. The stages are schematically demonstrated in a block diagram in Figure 12. The first two stages are fixed filters that account for outer and middle ear transfer functions, so the sounds that get filtered out by the outer and the middle ear do not contribute to loudness perception. The remaining three stages are similar in concept to the stages in Zwicker's model (Fig. 11).

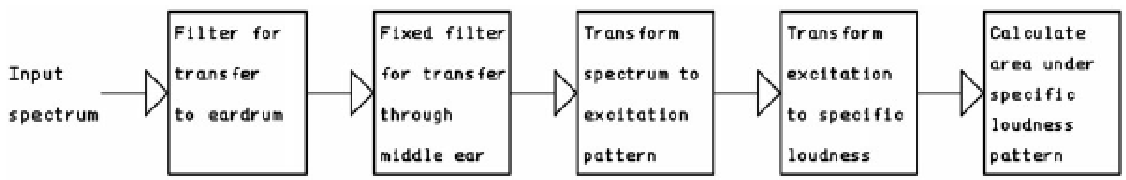


Figure 12: Block diagram of the 1996 loudness model by Moore and Glasberg.

The loudness model for complex sounds by Moore and Glasberg (Moore, 1997; Moore & Glasberg, 2004) shares the basic concepts with the earlier model by Zwicker and colleagues (Zwicker & Scharf, 1965). The main differences between the two models are summarized in Table 3.

Table 3: Summary of the main differences between the two models of loudness: Zwicker's model and Moore and Glasberg's model (Moore et al., 1997). Adapted from Moore and Glasberg (2004).

	Zwicker et al	Moore and Glasberg
Outer and middle ear transfer function	Below 2 KHz: Uniform transmission Above 2KHz: reflects absolute threshold curve but inverted in shape	Filter below 1 KHz reflects absolute threshold curve
Method used to calculate excitation patterns	Based on critical bands theory	Based on equivalent rectangular bandwidth (ERB) as a unit of frequency scale (Glasberg & Moore, 1990)
Specific loudness transfer function for partially masked sounds	Developed correction factor through subjective testing	Derived mathematically
Loudness at threshold	Assumes zero loudness at threshold	Predicts small but finite loudness at threshold

Moore and colleagues' loudness model formed the basis of the ANSI (2007) standards and part of the ISO standards (Scheuren, 2014). In 2002, the model was modified to account for non-stationary sounds (Glasberg & Moore, 2002).

In addition, the model by Moore and colleagues has had successful industrial applications. It has been used to predict the loudness of noises produced by airplanes, cars, ventilation systems, musical instruments and mobile telephones.

Modifications to Moore and colleagues loudness model to account for hearing loss

Modifications to Moore and colleagues' loudness perception model have been made to account for cases of hearing loss (Moore & Glasberg, 1997; Moore & Glasberg, 2004). First, the model's modification for hearing loss assumes that the amount of hearing loss at a specific frequency as measured in audiometric testing in dB HL (HL_{TOTAL}) is divided into a component due to OHC damage and a component due to IHC (and neural) damage ($HL_{OHC} + HL_{IHC}$). The modified model accounts for the main changes in cochlear functions that occur as a result of OHC and IHC loss. Table 4 lists these changes, indicates what type of hair cell loss is responsible for the change, and briefly describes how these changes were accounted for in the modified model.

Table 4: Summary of Moore and Glasberg model's modifications to account for hearing loss. The model accounts for consequences of outer hair cell loss and inner hair cell loss on loudness perception separately.

The dysfunction modeled	In HL_{OHC}	In HL_{IHC}	How it was modeled	Stage in model as in Figure 11
Threshold elevation	▪	▪	• $HL_{Total} = HL_{OHC} + HL_{IHC}$ • Model incorporates default values for HL_{OHC} and HL_{IHC} (90% for HL_{OHC}) • Elevated threshold is modeled by a linear attenuation of excitation level per critical band • Normal hearing is a special case in which HL=0	e
Reduction of compressive nonlinearity	▪		The higher the HL_{OHC}, the steeper the function relating specific loudness to excitation level (Fig. 13).	f
Reduction of frequency selectivity	▪		The higher the HL_{OHC}, the broader the auditory filter, the broader the excitation pattern.	c and e
Reduction of IHC/neural function		▪	By attenuating the excitation level at the frequency in question in a progressive manner	e

Table 4 (continued).

			(the higher the signal level above threshold, the less attenuation applied).	
Complete loss of IHC/neural function (dead region)		▪	Setting excitation to almost zero, evoking zero specific loudness from that region.	e

The only variable that is adjusted to account for hearing loss is the variable HL_{TOTAL} that in turn adjusts the variable HL_{OHC}. HL_{IHC} is simply equal to HL_{TOTAL} minus HL_{OHC}. The value for HL_{OHC} is either chosen on the basis of psychoacoustic data (Moore, Vickers, Plack, & Oxenham, 1999) or adjusted randomly (within a limited value of 57.6 dB) to best fit empirical data (Moore & Glasberg, 2004). Good fits were generally obtained when HL_{OHC} was 0.9 HL_{TOTAL}, with HL_{OHC} not exceeding the maximum value allowed at that frequency (Moore & Glasberg, 2004). The model also requires specifications of any dead regions within the hearing loss.

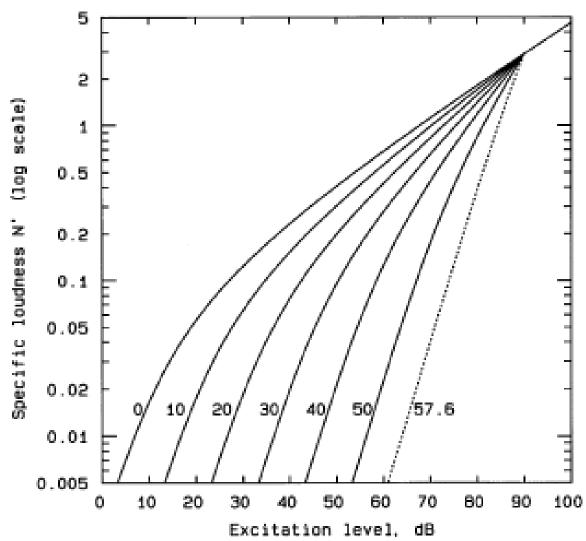


Figure 13: The transformation from excitation level to specific loudness for several peak excitation levels at absolute thresholds. The numbers on the curves indicate the relative values of excitation to the value of normal hearing (for frequencies above 500 Hz). Graph reprinted from Moore and Glasberg (2004) with permission.

The modified loudness models that account for hearing loss have been used in hearing aid design to set amplification algorithms and in the audiology literature to set theoretical loudness goals for hearing aid users. Currently, the loudness model by Moore and colleagues (Glasberg & Moore, 2002; Moore & Glasberg, 2004) is the one widely used in hearing rehabilitation design and research.

Using the loudness model for listeners with hearing loss by Moore and colleagues has adequately predicted loudness of sound within the reduced dynamic range of hearing aid users. An exception to this is the prediction of loudness of soft sounds by this group, which are regularly perceived louder than predicted by the model. A newly introduced concept that describes loudness growth near elevated thresholds for listeners with hearing loss supports this long lasting clinical research observation in hearing aid users, the concept of softness imperceptions (Buus & Florentine, 2002).

Softness Imperception (SI)

Softness imperception (SI) can be defined as the inability to perceive a range of low loudness as soft that can be perceived by normal hearing listeners as soft. In other words, and in terms of describing the loudness function, SI is an abnormally large perceived loudness at an elevated threshold due to cochlear hearing loss (Buus & Florentine, 2002; Florentine et al., 2005).

Currently, there is not enough evidence to confirm one loudness function near threshold for listeners with elevated thresholds. It is constantly mentioned in the literature that loudness growth function near elevated threshold shows large variability from RG to SI (Marozeau and Florentine, 2007; Silva and Epstein, 2012).

Figure 14 displays the current loudness functions for listeners with normal hearing and the currently proposed loudness functions for listeners with hearing loss, the classical steeper power function (rapid growth) and the newly proposed softness imperception (SI). Table 5 compares and contrasts the two views of loudness growth for listeners with hearing loss.

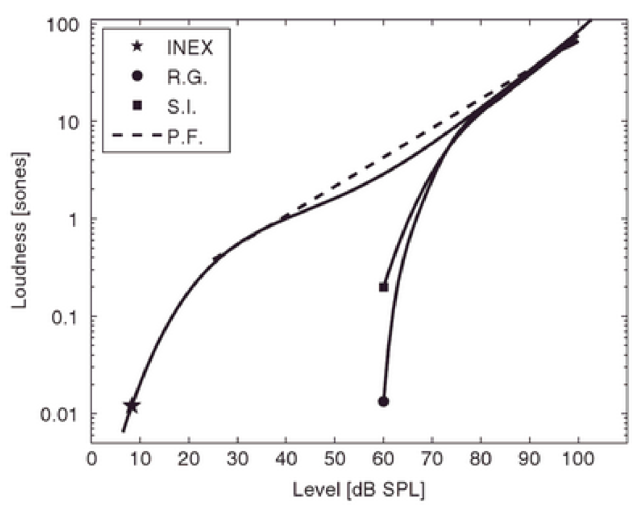


Figure 14: Loudness growth models for normal hearing listeners, power function (PF) (dashed) and INEX function (star), and for listeners with hearing loss, rapid growth (RG) (circle) and softness imperceptions (SI) (square). Graph obtained from (Marozeau & Florentine, 2007) with permission.

Table 5: Comparison between the two existing loudness growth functions for listeners with hearing loss, the long lasting rapid growth function and the new softness imperceptions function.

	Rapid Growth	**Softness Imperception**
Loudness at elevated threshold **(primary assumption)**	Same as loudness at normal threshold	Higher than loudness at normal threshold in *some* listeners
Growth of loudness near threshold	Same as normal growth up to 4-10 dB SL followed by rapid growth (recruitment) (larger slope than normal)	Same (not significantly greater) as normal growth at any SL
Loudness at mid levels	Grows more rapidly than normal	Grows more rapidly in *some* listeners
Loudness at high levels	Same or approaches that of normal hearing	Approaches that of normal hearing
Defines recruitment as	Classical definition: "Abnormally rapid growth of loudness above an elevated threshold"	"Abnormally large loudness at elevated threshold"
Dynamic Range	Reduced in SPL	Reduced in SPL and loudness
Acceptability	Incorporated in loudness models for complex sounds	Referred to in the literature as an alternative loudness growth function for listeners with hearing loss but never incorporated in loudness models

Table 5 (continued).

Loudness growth individual variability	Similar elevated thresholds may exhibit different loudness growth due to etiology differences	Similar and different elevated thresholds may have different loudness at threshold

The way the current loudness models for complex sounds (Moore & Glasberg, 1997; Moore & Glasberg, 2004) transfer excitation level to specific loudness follows a fixed change in slope as a function of increased excitation at threshold (i.e., loudness function slope increases with increasing threshold) (Fig. 13). Although using a fixed shape of loudness function for all listeners with the same audiometric threshold is feasible and in line with the main purpose of modeling, it obscures individual variability that is known to be more prominent in listeners with hearing loss than individuals with normal hearing (Hellman & Meiselman, 1993). The literature finds no strong correlation between audiometric elevated thresholds and predicted psychoacoustic tasks in general including loudness functions obtained from listeners with hearing loss (Elberling, 1999; Launer, Holube, Hohmann, & Kollmeier, 1996).

Other models that account for hearing loss

The Dynamic Loudness Model (DLM) is an adaptation of Zwicker's model by Chalupper and Fastl (2002). The DLM predicts loudness of both stationary and non-stationary sounds for listeners with normal and impaired hearing. For modeling loudness of listeners with hearing loss, only the specific loudness function is modified to reflect loss of compression in loudness growth and is based on audiometric threshold and categorical loudness scaling. Loudness prediction by the DLM was found to be largely in agreement with loudness data for strong spectral time-varying sounds compared to other models for complex sounds (Rennies, Verhey, & Fastl, 2010).

Unlike the psychophysical and physiological models that represent cochlear frequency analysis as a separate linear filter bank followed by a nonlinear compressive stage (Patterson & Holdsworth, 1996), the Time-Domain Loudness Model (TDLM) is a more recent model that combines these two stages in a single nonlinear filter section. The TDLM design resembles the psychophysical model by Moore and Glasberg (1997) more than any other physiology-based model. The TDLM, however, is only an auditory signal processing model that does not predict loudness of stimuli as an output of the model (Neely, Rodriguez, Liu, Jesteadt, & Gorga, 2009).

In terms of modeling low-level sounds in the case of elevated thresholds, both the DLM and TDLM assume zero loudness level at threshold and assume that the slope of loudness function at threshold is twice the slope of the normal hearing loudness function. Empirical evidence has not supported these assumptions.

Summary

As discussed in this section, loudness models have gone from simple predictors of loudness relying on a single sound parameter, i.e., intensity, to multi-leveled models that include multiple sound parameters and account for the contextual and physiological aspects of hearing. A shared stage in all the models for complex sounds is one that uses the loudness function to transform the excitation level into specific loudness. Loudness functions for normal hearing listeners are similar in shape and slope and are agreed upon in the psychoacoustic literature. The loudness functions for listeners with hearing loss of cochlear origin were established decades ago with only minor variations among studies in the value of exponent and were only recently challenged by the concept of softness imperception.

The loudness functions currently employed in all the available loudness models for complex sounds for hearing loss are rapid growth functions that assume equal loudness perception at elevated and normal thresholds. The fact that this assumption has not been sufficiently tested by different research groups, and the fact that there are ongoing clinical loudness outcomes from hearing aid users that do not match predicted loudness by the current models, make this new and controversial concept of greater loudness at threshold (i.e., SI) intriguing and worthy of further investigation.

In an upcoming section (section 3), the studies that have either supported or refuted the concept of SI will be reviewed. The possible reasons for contradicting findings among studies will be proposed. In addition, the possible theoretical physiologic and psychoacoustic rationales for the SI concept that have not yet been investigated will be discussed.

2.1.5 Measurement of Loudness

Loudness, as with any other psychophysical attribute, is very difficult to measure because it is dependent on the interpretation by listeners of what they hear. Since loudness is a subjective judgment, the ideal tool to measure it would require human listeners to listen to sounds and describe the magnitude of their sensation. A valid scientific method to measure loudness would enable the listener to provide a repeatable, reliable, and precise description of the perceived loudness with minimum experimental bias. While in principal researchers have agreed that there is no objective way to directly measure any psychological phenomenon, several methods have been used to quantify the subjective loudness of sounds in human subjects.

Throughout the loudness literature, many different methodologies have been proposed and used to measure the perceived loudness of sounds. This section will review these methodologies along with their strengths and limitations.

2.1.5.1 Perceptual correlates: psychoacoustic measures

Loudness matching

The loudness matching procedure was first proposed by Fletcher and Munson (1933) as a novel approach to create a scale of loudness. In loudness matching, the listener is asked to match the loudness of one stimulus to the loudness of another. The result of this procedure determines the sound's *loudness level*, the level at which the test stimulus and a standard stimulus are perceived as equally loud. As Fletcher and Munson (1933) did, listeners are typically asked to adjust the sound level of a tone with a given frequency to match the loudness of a 1000 Hz tone with a given level. Such experiments were used to obtain equal loudness contours (Fletcher & Munson, 1933) to show the dependence of loudness on frequency. However, loudness matching procedure is also useful to compare loudness of sounds that differ in other physical parameters, such as duration (Florentine et al., 1996) and bandwidth (Zwicker et al., 1957).

Loudness matching measures do not construct a direct loudness scale on which loudness can be measured (i.e., do not provide direct information about how loud a stimulus sounds); however, if the loudness function of the standard stimulus is known, the loudness function of a test stimulus can be obtained from its equal loudness contours. Psychoacousticians consider loudness matching measurement the gold standard by which loudness results obtained using other methods must conform (Florentine, 2011).

Modern experiments of loudness matching employ the adaptive two-alternative forced-choice procedure for listeners to adjust the test stimuli (Silva & Florentine, 2006). The adaptive method, in contrast with traditional method of adjustment described previously, overcomes some of the biases listeners might make in the stimulus adjustment.

Loudness scaling

Loudness scaling can be defined as relating the physical magnitude of sounds to their subjective loudness. The development of loudness scales was pioneered by Stevens, not only for loudness but for other sensory modalities (Stevens, 1955, 1956, 1957). Loudness scaling techniques have been primarily used to obtain loudness growth information. The most common methods that have been used to derive loudness scaling are the following: magnitude estimation, magnitude production, cross modality matching, and categorical scaling.

Magnitude estimation

The guidelines of the loudness magnitude estimation (ME) procedure were first proposed by Stevens (1956). Listeners are asked to assign a number to each sound they are presented with at different levels according to loudness. Listeners are typically instructed to assign a positive number but allowed to assign decimal and fractional numbers as long as they are within a ratio scale. Sometimes a standard or a reference sound is presented along with a given value (for example, 10) so subjects can judge the loudness of the test sounds relative to that of the standard. Researchers suggested that it would be better to not use a standard tone, and allow subjects to choose any numbers to make their judgments, so they would not be influenced by the level of the comparison tone (Hellman, 1976; Stevens, 1955). Poulton (1989), in particular, argued that with the use of a reference tone, the judgment of loudness becomes an invalid judgment of loudness ratios, as opposed to loudness differences or intervals, which more accurately reflect the listener's inner subjective units.

Regardless of these differences, inexperienced listeners can perform the magnitude estimation procedure with reliable responses (Logue, 1976). Magnitude estimation is not typically used in clinical settings since it has been suggested that this measurement is more suitable for group mean loudness data as opposed to individual data (Epstein and Florentine, 2006).

Magnitude production

Magnitude production (MP) is a method that inverts magnitude estimation. Listeners are presented with a number and asked to adjust the level of a test sound until it has a specified loudness that matches that number, either in absolute terms or relative to that of a standard (e.g., twice as loud, four times as loud, half as loud, etc). In such cases the method is sometimes called ratio production. Averaging results from ME and MP may compensate for the biases of each individual method (Hellman & Meiselman, 1993).

Cross-modality matching

Cross modality matching (CMM) was proposed by Stevens (1959) as an alternative to typical loudness matching. In this method, the loudness continuum is matched with a normally functioning sensory continuum. Cross modality matching can be thought of as an extension of ME and MP. Instead of controlling the loudness of a tone to match an assigned numerical value (or a fraction of it) in MP, listeners are asked to control the stimulus loudness to match the magnitude of a percept in another modality, such as the length of line or string, the brightness of a light stimulus, or the tactile sensation of vibration. Same as in the case of ME, instead of providing a numerical value that matches the loudness of a test tone, listeners in CMM are asked to match the magnitude of the other percept (e.g., length of a line) to the loudness of a test sound. Loudness data obtained from CMM have been shown to be in agreement with data obtained by ME and MP from both listeners with normal and impaired hearing (Hellman, 1999; Hellman & Meiselman, 1988, 1993).

Categorical scaling

In categorical scaling (CS), subjects are presented with a test stimulus and asked to rank it on a perceptual scale using one of several verbal categories (Cox, Alexander, Taylor, & Gray, 1997). The category labels may vary between studies, but typically they include: ‘cannot hear’, ‘very soft’, ‘soft’, ‘comfortable’, ‘loud’, ‘very loud’ and ‘too loud’. Sometimes the verbal categories are transferred to numerical scales, for example from 0 through 50 or 1 through 9. Recently, the use of CS has been updated and converted into a computerized adaptive procedure (Brand & Hohmann, 2002; Keidser, Seymour, Dillon, Grant, & Byrne, 1999).

The CS method has been widely criticized by the loudness research community for two reasons. First, since CS uses linguistic categories to describe perception, it has been argued that language is inadequate to transmit perceptual experience (Florentine, 2011). For example, choosing the category "very soft" to describe a sound by a listener with hearing loss, might not necessarily indicate that the sound is perceived as very soft, but rather indicate that it is the softest sound the listener can hear, or that the sound is just softer than what would be described with the higher available category in the scale provided, "soft". Hence, the same word "soft" can refer to different perceptions when used by different people, setting different criterion (Florentine, 2011).

The other criticism of the CS method is that it induces a bias by setting limits to the perception within the categories provided, and that loudness must only grow within those categories. Moreover, in contrast to the continuous value ratio scale of sones described earlier, an ordinal or categorical scale implies a perceptual difference between neighboring categories, which might not be equal.

Despite these shortcomings, categorical scaling has been extensively used in audiology and hearing aid research to assess loudness perception in listeners with hearing loss. This method might be an appropriate way to characterize a listener's loudness growth in a clinical setting due to its feasibility, ease of instruction and time efficiency. In addition, its reliability within session and between sessions has been documented (Al-Salim et al., 2010; Ellis & Wynne, 1999). Using CS as a measure of loudness might be appropriate for clinical use as long as the clinician compares its results within patients and not between patients (Elberling, 1999). Some hearing aid fittings aim to achieve a preferred listening loudness range and as a result, clinical procedures using CS to quantify the growth of loudness have been proposed (Cox 1997).

Nevertheless, for the reasons discussed above, CS might not be appropriate for the purpose of drawing inferences and conclusions about loudness perception characteristics and loudness growth in listeners with hearing loss, such as it is used now in the hearing aid literature. It should be noted that CS loudness functions obtained from listeners with normal hearing have an upward concave shape (Brand & Hohmann, 2001b, 2002; Elberling, 1999; Keidser et al., 2012), which does not agree with currently accepted loudness functions obtained by magnitude estimation and magnitude production methods.

Reaction time

Reaction time paradigm was first used in 1940 as a measure of loudness (Chocholle, 1940). In this paradigm, the listener is asked to physically react to a sound, such as pressing a button, as soon as he hears it. Simple reaction time is latency between the onset of stimulus presentation and the onset of the listener's response. The underlying assumption is that louder sounds generate shorter reaction times (RT) and therefore, sounds with the same RT are interpreted to be equally loud. Loudness data obtained by using reaction time were found to be in agreement with data obtained using loudness matching and magnitude estimation procedures from both listeners with normal and impaired hearing (Humes & Ahlstrom, 1984; Pfingst, Hienz, Kimm, & Miller, 1975).

This method has been used to examine loudness near threshold due to the difficulty of using other perceptual correlates of loudness scaling at such low-level sounds, such as magnitude production and estimation. Reaction time as a correlate of loudness is, however, frequency dependent (Humes & Ahlstrom, 1984; Pfingst et al., 1975); a fact that should be taken into account when employing it.

Equal latency contours were obtained from both animals and human listeners using the reaction time paradigm and successfully replicated the equal loudness contours of humans (Pfingst et al., 1975; Stebbins, 1966).

Psychacousticians agree that there is no one best method to measure loudness. Each method has its strengths and limitations and researchers have to choose the method that best serves the purpose of their experiment while acknowledging the limitations.

2.1.5.2 Physiologic correlates

In contrast to the subjective measures of perceptual correlates of loudness reviewed above, a number of objective physiological measures have been examined or used by researchers as correlates of loudness. Physiological measures do not require the listener's cognitive participation and therefore, are often suitable to use with children, adult listeners with cognitive challenges, and in studies with animal models.

Basilar membrane velocity

Loudness, at least for pure tones, corresponds linearly to the squared basilar membrane (BM) displacement amplitude – up to 60 dB SPL (Buus & Florentine, 2001; Schlauch et al., 1998; Yates, Winter, & Robertson, 1990). In the study by Ruggero, Rich, Recio, Narayan and Robles (1997), loudness growth data obtained from human listeners corresponded with BM velocity measures obtained from a chinchilla. The psychoacoustic three-stage loudness function (INEX) that describes the loudness growth of low, moderate, and high level sounds is in agreement with the three physiological stages of BM input-output function that describes the growth of BM displacement amplitude with increasing sound level (Oxenham & Plack, 1997).

Otoacoustic emissions

As mentioned earlier, BM displacement amplitude is closely related to loudness perception. It is not possible to measure BM displacement directly in humans; therefore, the relationship between otoacoustic emissions (OAEs) - acoustic signals recorded in the auditory canal that originate within the cochlea as a by-product of its normal function - and BM displacement amplitude in humans (Epstein & Florentine, 2005a) and in animal models (Withnell & Yates, 1998) has been examined based on the assumption that OAEs are an indicator of BM velocity, and that they are closely related at least at low and moderate levels.

More recent studies have found that estimated loudness from transient evoked OAEs (tone burst OAEs) and Distortion product OAEs are proportional to both calculated (Epstein & Florentine, 2005a) and measured (Epstein, Marozeau, & Florentine, 2006; Rasetshwane, Neely, Kopun, & Gorga, 2013; Silva & Epstein, 2010; Thorson, Kopun, Neely, Tan, & Gorga, 2012) loudness functions.

In listeners with hearing loss, the estimation of loudness growth using tone burst OAEs has been found to be poor compared to using other physiologic correlates such as evoked potentials (Silva & Epstein, 2012). The use of evoked potentials (Auditory Brainstem Response) will be discussed in the next section.

Auditory brainstem response

A number of studies investigated the correlation between loudness and specific features of auditory brainstem response (ABR), the recorded electrical activity from the auditory system (auditory nerve and brainstem nuclei) generated as a response to acoustic stimulus. Researchers initially investigated the use of click evoked ABRs (Davidson, Wall, & Goodman, 1990; Pratt & Sohmer, 1977; Serpanos, O'Malley, & Gravel, 1997), which are not frequency specific. Studies have found a correlation between psychoacoustic measures of loudness and click evoked ABR measures such as wave V amplitude (Davidson et al., 1990; Pratt & Sohmer, 1977) and wave V latency (Serpanos et al., 1997) obtained from the same listeners, both with normal hearing (Pratt & Sohmer, 1977) and hearing loss (Davidson et al., 1990). Wave V amplitude grows with increasing intensity, while wave V latency decreases with increasing intensity.

On the other hand, other studies failed to find such a correlation between the same above mentioned ABR measures and perceived loudness (Thornton, Yardley, & Farrell, 1987; Wilson & Stelmack, 1982).

More recently, tone burst ABR has been investigated as an estimate of frequency-specific loudness growth. Instead of looking at a correlation relationship, Silva and Epstein (2010, 2012) used the mean square-error as an indicator of the similarity between psychoacoustically measured loudness and some features of tone-burst ABR obtained from listeners with normal and impaired hearing. The researchers used a signal processing scheme to account for the previous click ABR studies’ shortcomings. The results showed close similarity between ABR estimated loudness and perceived loudness.

Auditory nerve firing rate

The rate of auditory nerve (AN) response has been extensively investigated in animal models as a correlate of loudness of sound in both ears with normal and impaired hearing. In one of the loudness coding theories, loudness is based on the total AN discharge rate. Loudness growth may reflect the slope of rate-level functions for individual auditory nerve fibers (Harrison, 1981); however, mixed results have been reported in the agreement of rate-level function and loudness growth function in cases of normal (Bilecen, Seifritz, Scheffler, Henning, & Schulte, 2002; Jeng, 1992; Relkin & Doucet, 1997) and impaired hearing (Heinz et al., 2005; Heinz & Young, 2004) in animal models.

Functional magnetic resonance

Beginning in the last decade, brain scans have been employed to examine the representation of loudness in the central auditory system. Functional magnetic resonance imaging (fMRI) studies have shown a growth of activation with increasing stimulus level (Langers et al., 2007; Sigalovsky & Melcher, 2006). The growth of activation observed in these studies was more related to the growth of loudness, than to the growth of intensity level (Langers et al., 2007). Studies that investigated amplitopicity, specialized intensity-encoding patterns in the human auditory cortex, a similar principle to tonotopicity, found a change in both activation volume and location with increasing level (Bilecen et al., 2002). The volume of activation might indicate the magnitude of loudness perception.

Pupil dilation

A recent study conducted by Liao, Kidani, Yoneya, Kashino, and Furukawa (2016a) looked at pupil dilation as a function of sound level. The purpose of the study was to examine the effect of salience on pupil dilation response by manipulating loudness. The study recorded pupil dilation from eight listeners with normal hearing while listening to randomly presented 1000 Hz tone bursts and white noises of different levels ranging from 45 to 85 dBA. The loudness perception of those sounds was calculated for each presentation level from Glasberg and Moore's (2002) loudness model. A regression analysis was conducted to predict the pupil dilation responses by the calculated loudness and showed that pupil dilation can be predicted by the loudness of sounds. The use of pupil dilation as an auditory-evoked response and as a measure of perceived loudness is discussed in further detail in section 3.0 in this document.

Table 6: Summary of the psychoacoustic and physiologic measures of loudness along with their strengths and limitations. The last column at right indicates the appropriateness of each method to quantify loudness at very low-levels (near threshold) based on strengths and limitations of applying the measure to near-threshold level sounds.

	Psychoacoustic method	Description	Advantages	Limitations	Sensitivity at low-level sounds
Loudness scaling	Categorical scaling (CS)	Listener ranks presented stimulus on a perceptual scale using one of several verbal categories	• Easy instruction • Time efficient • Feasible for clinical use • Provides descriptors of perception that can be directly use by clinicians • Can be used with young children	• Biased by linguistic categories • Implies equal perceptual differences between categories • Constrains perception within categories provided.	Low
	Magnitude estimation (ME)	Listener assigns a number that matches the loudness of presented stimulus	• Reliable measure	• Difficult instructions • Requires training • Time consuming • Not practical for routine clinical use	-

Table 6 (continued).

				• Lack direct descriptors to perception that can be used clinically.	
	Magnitude production (MP)	Listeners are presented with a number and asked to adjust the level of a test stimuli until it has a specified loudness that matches that number, either in absolute terms or relative to that of a standard	• Combining ME and MP compensates for the biases of each individual method	• Same as ME	-
	Cross modality matching (CMM)	Listener is asked to match the loudness of presented stimulus to the magnitude of another percept, e.g., length of line			-
	Loudness matching	Listener is asked to match the loudness of one stimulus to the loudness of another stimulus	• Gold standard method to obtain loudness level	• Time consuming • Do not provide direct scale of loudness	Medium
	Reaction Time	Latency between onset of stimulus and onset of reaction (pressing a button or resale a button)	• Reliable measure for tone near threshold and suprathreshold	• Non-linguistic task • Can be used in trained animal models • Sensitivity to low and high SLs showed mixed results	Medium

Table 6 (continued).

	Physiological correlates	Loudness estimated from:	Advantages	Limitations	Sensitivity at low-level sounds
	Basilar membrane displacement (BM)	Squared BM velocity/vibration amplitude	-	• Only suitable for animal models	-
	Auditory brainstem response (ABR)	Click ABR: wave V amplitude and latency. Tone burst ABR: wave V amplitude using a three stages signal processing procedure (Silva and Epstein, 2010)	• Objective measure	-	Not sensitive near threshold
	Otoacoustic emissions (OAE)	Tone burst OAEs: Cross spectrum response amplitude	• Non-invasive, quick, objective measure	• Poor estimate of loudness in listeners with hearing loss	-
	fMRI	Activation volume at central levels	• Objective	• Acoustic noise	Not tested
	Auditory nerve firing rate	Firing rate	-	• Only suitable for animal models, invasive procedure	-
	Pupil dilation	Peak, latency and average of pupil dilation	• Objective	• Newly introduced method • Incorporates both physiologic and psychologic components of perception	Not tested

2.2 LOUDNESS AT AND NEAR ELEVATED THRESHOLDS

The topic of loudness perception in listeners with hearing loss, and how it compares to that of normal hearing listeners has been investigated in the psychoacoustic and audiology literature for a long time. Most studies have focused on the rate at which loudness grows within the dynamic range as signal intensity increases (Buus & Florentine, 2002; Heinz & Young, 2004; Hellman & Meiselman, 1993; Silva & Epstein, 2012), and how loudness perception differs as a function of other factors, such as stimulus frequency (e.g., Fletcher & Munson, 1933), duration (e.g., Ekman et al., 1966; Stevens & Hall, 1966), and the psychological status of the listener. By contrast, the loudness perception of soft sounds has received less attention from the psychoacoustic and audiologic research community. Listeners with hearing loss lose sensitivity to low-level sounds, which are perceived as being very soft by normal hearing listeners, and the softest sounds they can hear with the hearing loss are no longer low-level sounds. Examining how listeners with hearing loss perceive sounds around their elevated thresholds is worth further attention

Given what we know about the changes in cochlear function and the changes in the auditory pathway's behavior with the presence of cochlear damage, one could hypothesize that the loudness (or perception in general) of low-level sound would be altered once the stimulus is perceived. The psychoacoustic literature to date has overlooked this initial stage of perception (at threshold) and instead focused on describing the change in perception at suprathreshold levels, assuming that perception at threshold remains unchanged.

Examining loudness at and right above threshold was not foremost on the agenda of earlier researchers partly because available methods could not accurately measure loudness at such low levels (Hellman & Zwislocki, 1961; Poulton, 1989). This is not surprising because measuring loudness at threshold is basically an attempt to measure the loudness of a stimulus that is only audible half or three-quarters of the time. In the absence of reliable methods to measure loudness perception near threshold, it was assumed that sounds judged as "very, very soft" at and near threshold by normal hearing listeners would be perceived in a similar way at and near elevated threshold by listeners with hearing loss if they are at the same sensation level. This assumption became the basis for modeling loudness perception and its growth in the case of hearing loss, and was accepted for decades without being tested. This assumption is plausible because one would not question whether a signal at a level that is just audible to any listener would be perceived as anything but the smallest amount of loudness possible.

Loudness perception of listeners with hearing loss has been described in the literature by the well-documented phenomenon known as recruitment. The long-standing notion of recruitment initially assumed that loudness at threshold for these cases is zero, similar to listeners with normal hearing, but grows at a higher rate above elevated threshold compared to normal hearing listeners. Afterward, based on psychoacoustic data, researchers agreed, that loudness at threshold is *not* zero and must have a finite amount (Buus & Florentine, 2002; Buus et al., 1998). Based on the findings of the study by Buus et al. (1998), the concept of softness imperception (SI) was proposed (Buus & Florentine, 2002). Buus and Florentine (2002) examined the growth of loudness with increasing intensity near threshold that is elevated due to cochlear HL. Five listeners with hearing loss were asked to adjust the level of a 1600 Hz tone to match the loudness of a complex tone centered around 1600 Hz. Individual pure tone loudness functions were derived from the matching data. The method used to extract the functions allowed them to estimate loudness functions at and near threshold. The derived loudness functions for the five listeners revealed that the local exponent of loudness function near threshold (10-15 SL) is the same (somewhat larger than unity) as in listeners with normal hearing near threshold (Buus et al., 1998). This finding contradicted the rapid recruitment growth concept that the exponent is larger than normal near elevated threshold. The authors explained this finding of normal rate of loudness growth near elevated threshold, and the fact that loudness reaches normal levels at high-level sounds, by introducing the concept that loudness at elevated threshold due to cochlear hearing loss is larger than normal; a concept that later came to be called softness imperception (SI) (i.e., the inability to perceive the signal as soft but rather perceive it as something greater than soft even though the sound is at or near threshold where normal listeners would label sounds as soft). Thus, the larger loudness at elevated threshold, along with the steeper slope of loudness function around 20 dB SL, contribute to reaching normal loudness at high levels.

The classical view of recruitment contends that there is a rapid growth of loudness with increasing level above elevated threshold, but loudness reaches normal at high levels (above 90 dB SPL). The level at which loudness starts to grow more rapidly than normal has not been well specified. However, the most current description of recruitment states that loudness grows at a normal rate up to 3-4 dB above threshold, and from that point on starts to grow more rapidly than normal, which is sometimes referred to in the literature as rapid growth (RG) when it is compared with the concept of SI. With the introduction of SI, there are now two different theories available to describe the perception of loudness at and near elevated threshold in listeners with hearing loss. A list of the main differences between the two views can be found in table 5.

An extensive number of studies have examined loudness growth with increasing level in listeners with hearing loss either psychoacoustically (e.g., Hellman, 1993; Hellman, 1999; Hellman & Meiselman, 1993; Launer et al., 1995; Moore et al., 1985) or physiologically (e.g., Heinz et al., 2005; Silva & Epstein, 2012). However, only a few of them have measured loudness at low sensation levels that are close to threshold. The methods used to measure or obtain loudness growth functions vary across these studies. In addition, their findings vary in their support of the SI model. The studies that included loudness at low sensation levels are summarized in Table 7.

For example, Whilby et al. (2006b) derived loudness functions of a 1000 Hz tone from eight listeners with normal hearing, and six listeners with bilateral 40-60 dB hearing loss. The loudness functions were derived from binaural level difference for equal loudness (BLDEL) data obtained across a wide range of levels from 10 to 90 dB SL. In a BLDEL task, the listener is asked to adjust the level of a monaural tone to match the loudness of a binaural tone and vice versa. The derived loudness functions for listeners with normal hearing were adjusted to give a value of one sone at 40 dB SPL, and were slightly steeper, but consistent with directly measured loudness functions from the same population. The derived loudness functions for listeners with hearing loss were adjusted so that the loudness of an 85 dB SPL tone matched the overall loudness of an 80 dB SPL tone for the normal hearing listeners. The derived loudness functions for the listeners with hearing loss were widely variable, which also was observed in direct measured loudness functions from the same population. The authors reported that one listener with hearing loss showed an overall steeper than normal slope across all sound levels consistent with the shape of many direct measured individual loudness functions in the literature which describe the classical concept of recruitment. The other listeners showed loudness functions that may be consistent with the concept of SI with loudness being greater than normal at and near threshold in some listeners with hearing loss (Buus & Florentine, 2002). Upon closer examination, the derived loudness functions for listeners with hearing loss in this study showed that they all started at either 10 or 20 dB SL, a level that exceeds the range at which the main controversy lies between SI and RG. Therefore, contrary to the author's report, the findings of this study neither support nor refute the concept of SI.

Another study by Silva and Epstein (2012) obtained loudness growth data from eight listeners with hearing loss using magnitude estimation (ME) and cross-modality matching (CMM). The perceptual ranges for the listeners with hearing loss, which were calculated by subtracting the log of minimum loudness value (the average number or string length a listener entered for a given threshold level) from the log of the maximum loudness value (the largest number or string length a listener entered at 100 dB SPL), and listeners with normal hearing obtained in a previous study (Silva & Epstein, 2010), were plotted as function of threshold. The data showed that listeners with hearing loss demonstrated greater variability in their perceptual range compared to listeners with normal hearing (from normal range to two standard deviations below the normal range). The variability in perceptual range was observed in some listeners including those with similar thresholds. The authors attributed these results to the possibility that listeners with hearing loss can exhibit either recruitment-type loudness growth with a perceptual range that is similar to the range of listeners with normal hearing, or softness imperception-type loudness growth with a perceptual range that is smaller than the range of listeners with normal hearing (i.e., missing the low loudness range).

Table 7: Studies that reported loudness at levels near threshold and whether they support the concept of softness imperception (agreement with SI is defined as larger loudness at threshold compared to normal, or similar slope of loudness growth above threshold (10-15 SL) as normal slope (close to unity).

Study	N / hearing thresholds	Stimuli	Lowest level measured	How loudness function data were obtained	Loudness at threshold	Slope of loudness function 10-15 dB above threshold	Study supports SI?
Silva and Epstein (2012)	8/ 30-70 dB	1000 Hz and 4000 Hz tone burst	5 dB below threshold (loudness obtained from 5-10 SL)	ME and CMM (string length)	Not obtained	Median power function slopes were within the range or slightly shallower than normal	Yes - Loudness growth function near threshold is not steeper than normal
Whibley et. al. (2006)	8/40-60 dB	1000 Hz tone	10-20 dB SL	Derived from BLDEL data	Not obtained	Not obtained	Did not address

Table 7 (continued).

Heinz & Young (2004)	Mild loss: 25-60 dB loss	1000 and 2000 Hz tones, broad band noise, and speech token	10-20 dB below fiber threshold	AN fiber rate-level function (correlate of loudness)	Not specified. No change in firing rate at threshold was noted in study graphs	Low-level slopes were not steeper than normal	Yes - Loudness growth function near threshold is not steeper than normal
Buus & Florentine (2002)	5/10-70	1600 Hz tone	0 dB SL	Derived from loudness matching data	Estimated. Larger than normal	Estimated to be same as or shallower than normal	Yes – Larger loudness at elevated threshold and Loudness growth function near threshold is not steeper than normal
Brand & Hohmann (2001)	8/ 15-90	NB noise (centered at 250 Hz and 4 KHz) and	0 dB SL (above threshold in dB HL).	Obtained by categorical scaling	Similar to rating obtained by control	Steeper than normal	No

Table 7 (continued).

		broadband signals	Exact levels not provided		listeners with normal hearing		
Hellman & Meiselman (1990)	100/40-86 dB	Tone burst, range from 500 to 4K Hz	4 SL	Obtained by ME, MP, and CMM then adjusted by computer program	Not obtained. Suggested that loudness at elevated threshold due to cochlear hearing loss is greater than loudness at normal threshold.	Assumed that slope below 4 SL has same limiting value as in normal. Slopes from 4 SL to levels where function deviates from a linear power function were steeper than normal in 4 listeners. (no other individual data reported)	Did not address but suggested larger loudness at elevated threshold

[1] Hellman and Meiselman (1990) calculated theoretical loudness functions based on a mathematical expression for loudness functions proposed by Zwislocki (1965) that assumes that the loudness of a tone in quit is equivalent to the tone and the intrinsic physiological noise. It also assumes that the effect of cochlear hearing loss on loudness function is analogous to the effect of masking noise on loudness function.

The above-mentioned studies obtained loudness functions at near threshold levels, but were not specifically conducted to test the concept of SI. In addition, the lowest level at which loudness was measured in all of the studies ranged from 4 to 10 SL. In studies were loudness at threshold (at 0 SL) was reported, they were either extrapolated or calculated data. None of the data included *measured* loudness *at* elevated threshold for listeners with hearing loss.

Only three studies have been specifically designed to test the concept of SI. The first study conducted by Florentine et al. (2005), used a reaction time (RT) paradigm as an indirect measure of loudness at and above threshold in six listeners with high frequency sloping hearing loss of cochlear origin. In order to overcome inter-subject variability, RT data were obtained at two frequencies from each listener: 1) a frequency with a normal or near normal threshold (lower frequency); and 2) a frequency with an elevated threshold (higher frequency), so each listener served as his or her own control. In order to eliminate the possibility of the frequency effect on RT, RT data were obtained at different frequencies from two listeners with normal hearing as a control group. The results showed that median RT was faster at elevated threshold than normal or near-normal threshold in all the listeners with hearing loss. This effect on RT was not observed between frequencies in the control group of normal hearing listeners. The authors concluded that their results are consistent with the concept of SI that supports larger loudness at elevated thresholds.

In a study conducted by May, Little, and Saylor (2009), RT data were obtained from six behaviorally trained cats to respond to tone bursts at different SPLs before noise exposure. Recruitment effect was also evaluated in four of those cats post noise exposure. The RT data before noise exposure showed great variability within trials in each cat at levels near threshold, but more consistent data were observed at suprathreshold levels. Therefore, the authors suggested that RT paradigm is more prone to errors at near threshold levels.

May, et al. (2009) reported that their RT data post-noise exposure were not in agreement with the concept of SI. Their results did not show faster RT at threshold when thresholds were elevated post-exposure and cannot be directly compared to Florentine et al's (2004) results of loudness at and near threshold. The RT data in this study pre- and post-exposure were compared at the same SPL. Presenting RT data at the same SPL levels in this manner means that RT at moderate SL pre- exposure were compared with RT at low SL (or even inaudible levels) post-exposure. Thus, RT data post-exposure were expected to show lower loudness. RT data for tones that were 10 dB SL (above normal threshold) and greater were the only SPL levels presented and therefore, a comparison of loudness at the same low SL (from normal and elevated thresholds) was not provided.

Epstein and Florentine (2006c) conducted a study to test the assumption that RT near threshold is not affected by stimulus frequency. RT data that they obtained from 16 listeners with normal hearing at 1 and 4 KHz, showed that there is an effect of stimulus frequency on RT, but it was not large enough to account for the difference in RT observed at thresholds of listeners with hearing loss reported in the study by Florentine et al. (2005).

Although the majority of research suggests that RT is an indirect loudness measure for tones (Buus, Greenbaum, & Scharf, 1982; Chocholle, 1940; Florentine et al., 2005; Marshall & Brandt, 1980; Pfingst et al., 1975), others have suggested that simple RT as a signal detection task requires a different perceptual process than loudness perception as a stimulus discrimination task (Kohfeld, Santee, & Wallace, 1981) and this may lead to ambiguous results (Humes & Ahlstrom, 1984).

The second study that was conducted to test the concept of SI was by Moore (2004). Loudness matching data were obtained from four listeners with bilateral sloping hearing loss. The listeners matched the loudness between a tone at a frequency where hearing was normal, and a tone at a frequency where there was a significant hearing loss. In two of the listeners, loudness matches were obtained at the same frequency across ears. In one listener, loudness matches were obtained at two different frequencies across ears. In another listener, loudness matches were obtained at two different frequencies in the same ear. Loudness matching data were obtained at SLs as low as 2 dB from three listeners and at 4 dB from the fourth listener. It was hypothesized that if the listeners exhibited SI, then a tone at very low SL from an elevated threshold would be matched by a tone at a higher SL from a normal threshold. The resulting loudness matching data showed that tones at very low SLs (between 0 and 4-10 SL) presented at frequencies with elevated thresholds were matched to tones at frequencies with normal thresholds at the same SL, which is inconsistent with the concept of SI. The data also showed that above 4-10 dB SL, loudness in impaired ears/frequencies grew faster than normal with increasing level. Both of these findings - normal loudness at threshold and faster growth of loudness above threshold - are consistent with the traditional view of loudness recruitment.

One of the possible limitations in the study by Moore (2004), is that the accuracy of the bi-frequency loudness matching method depends on the assumption that listeners, especially those with hearing loss, are able to perform loudness matches between widely different frequencies. This task is often difficult (Hellman, 1976; Zwicker et al., 1957) and becomes more difficult with the presence of hearing loss (Schuknecht, 1970 and Penner,1984). The authors of the study reported that their listeners were trained in a two-hour session until their performance appeared to be stable and consistent. One could argue that such lengthy training might impact the listener's matching criterion in order to coincide with the expectations of the experimenter.

The two studies mentioned above are the only ones that aimed to test the concept of SI in a controlled laboratory. Table 8 compares the main differences between the two studies and lists possible limitations in each of them. While each study might have its own particular limitations, several are shared. For example, both studies included a very limited number of listeners, which might not represent the wide variability of loudness perception that exists among the population of listeners with hearing loss. However, small sample size is widely accepted in psychoacoustic studies.

Table 8: Studies testing the concept of SI.

Study	# of listeners with HL	Etiology of HL	Loudness measure	Strengths	Limitations
Florentine et al (2005)	6/sloping loss	Not specified	Reaction time to 200 ms tone at 2 frequencies in same ear (one at normal and one at elevated threshold)	Sensitive to low SLs	• RT is an indirect measure of loudness. • RT data at suprathreshold did not match empirical loudness data obtained with other methods. • RT to tones is frequency dependent to some extent (Epstein & Florentine, 2006b). Study has only 2 listeners to demonstrate absence of frequency effect on RT data. • Validity of RT paradigm to assess loudness has mixed views
Moore (2004)	4/ bilateral sloping loss with significan t	Unknown (2)/Cong enital/noi se induced	Loudness matching of 500 ms tones at 2 different or same frequencies across ears (one	Direct measure of loudness level Sensitive to low SLs	• Listeners have asymmetrical loss. Listeners with this type of loss has steeper loudness functions than listeners with symmetrical loss (Knight and Margolis, 1984); therefore, this conclusion should be restricted to this specific type of listeners (Marozeau & Florentine, 2007).

Table 8 (continued).

	asymmetr y		at normal and one at elevated threshold)		• All listeners in the study have severe hearing loss (thresholds of 63, 70, 71, and 74.5 dB SPL) who are more likely to show steeper loudness functions than listeners with loss of less severity (Hellman & Meiselman, 1990). • Questionable accuracy of loudness matching of tones with different frequencies.

The third study that was conducted to test the concept of SI was a review by Marozeau and Florentine (2007). In this study, the individual loudness functions for four listeners with hearing loss were extracted from five different studies using five different methods to measure loudness. If loudness at and near threshold was not measured in the study, the loudness was extrapolated from the threshold SPL and the slope of the loudness function at the lowest level measured. The criteria in the study was that listeners showing rapid growth will exhibit a slope near threshold that is steeper than the comparable slope in listeners with normal hearing, and that listeners showing softness imperceptions will exhibit a slope near threshold that is similar to listeners with normal hearing.

One of the studies reviewed in Marozeau and Florentine (2007) was conducted by Hellman and Meiselman (1990) who obtained loudness growth functions from 100 listeners with bilateral hearing loss. Loudness judgments were obtained using ME, MP, and CMM methods of tone bursts across the range from 4 dB SL up to 35 dB SL. Loudness judgments were obtained at frequencies where hearing loss exceeded 40 dB, which ranged from 500 Hz to 4K Hz. Loudness functions were then generated using a computer program to best fit the data points from each listener, assuming that the slope of loudness near threshold (up to 4 dB SL) has the same limiting value in normal and impaired hearing and that it is also the same for levels above 30 dB SL.

Marozeau and Florentine (2007) extrapolated the loudness at and 4 dB above threshold for four listeners in the study by Hellman and Meiselman (1990). The loudness was extrapolated from the threshold SPL and the slope of the loudness function at the lowest level measured (4 dB SL in this study) using ME method. The authors reported that the extrapolated loudness functions from the four listeners showed that loudness growth at threshold is the same or shallower than the INEX function for normal hearing.

The slope values of loudness growth for the four listeners obtained using ME reported in Hellman and Meiselman (1990) were not in agreement with the values reported in Marozeau and Florentine (2007). It is possible that Hellman and Meiselman (1990) reported the overall slope of the loudness function while Marozeau and Florentine (2007) calculated the specific slope for near threshold levels only. Therefore, the individual data from Hellman and Meiselman (1990) cannot be considered to support the SI concept.

All the remaining studies that were reviewed by Marozeau and Florentine (2007) showed individual variability between listeners (Brand & Hohmann, 2001a; Buus & Florentine, 2002; Whilby et al., 2006b). Some exhibited rapid growth and some exhibited softness imperceptions, while others exhibited an intermediate behavior or a combination of both. The authors concluded that loudness functions presented in their review greatly emphasized the individual variability among listeners with hearing loss, a variability that tends to be masked by averaging data in research studies. Nevertheless, the accuracy of their conclusions has been disputed. Therefore, one cannot conclude from Marozeau and Florentine (2007) that there is solid evidence to support the existence of the concept of SI in some listeners with hearing loss.

In summary, based on this review of the literature, there are currently only two studies that specifically tested the concept of SI (Florentine et al., 2005; Moore, 2004) and both revealed conflicting/inconsistent findings. Therefore, there is a need for further testing of models that explain loudness perception at or near elevated thresholds of individuals with hearing loss.

There is enough evidence to show that listeners with hearing loss show larger individual differences in their loudness function slopes and shapes compared to variations in listeners with normal hearing (Hellman & Meiselman, 1990). Within these variations, both models of loudness growth have been reported to explain findings in the literature (Florentine et al., 2005; Marozeau & Florentine, 2007; Moore, 2004). Investigations that aim to correlate cochlear hearing loss etiologies to different loudness growth profiles are often proposed in the literature.

In the following sections, previous and current findings in audiological clinical research, as well as current clinical practices in hearing aid fitting that support the concept of SI in some listeners with hearing loss will be reviewed. In addition, possible physiological rationales behind the concept of SI will be proposed.

2.2.1 Clinical rationale for softness imperception

Before the introduction of non-linear hearing aids, linear hearing aid fitting aimed mainly to provide audibility to as many sound levels as possible, while maintaining as much comfort as possible. To do this, linear hearing aids basically provided audibility of moderate-level sounds. By definition, linear fitting provides a frequency dependent amount of gain that is equal across all input levels. Soft sounds that need the greatest amount of gain to reach audibility were not always amplified with linear fitting strategies due to poor feedback control at that time, and loudness discomfort since providing the same increased amount of gain to moderate-level sounds, over amplifies them. High-level sounds do not get increased gain as a function of further increases in intensity in linear fitting due to the use of peak clipping with linear fitting strategies.

With wider adoption of non-linear hearing aids in the late 1990s, it became possible to provide frequency and level dependent gain, such that low-level sounds receive the greatest gain and moderate-level sounds receive more gain than high-level sounds, through the use of wide dynamic range compression (WDRC) algorithms and the introduction of enhanced feedback management technology. This change in fitting strategy introduced the possibility of returning audibility to all sound level categories for a listener with mild to moderately severe hearing loss. In turn, several available prescription formulas at that time, i.e., Independent Hearing Aid Fitting Forum (IHAFF) (Cox, 1995), FIG6 (Killion & Fikret-Pasa, 1993), the National Acoustic Laboratories – nonlinear version 1 (NAL-NL1) (Byrne, Dillon, Ching, Katsch, & Keidser, 2001), and the desired sensation level [input/output] (DSL [i/o]) (Cornelisse, Seewald, & Jamieson, 1995), were modified to account for this new goal of hearing aid fitting.

It was not until the introduction of WDRC fitting strategies that soft sounds started to be re-introduced to the dynamic range of hearing aid users; an advantage that has not always been well received. Despite the use of most current hearing aid algorithms employing loudness models that account for hearing loss and following hearing aid fitting best practices, amplified soft sounds are often described by hearing aid users as loud and unpleasant or loud but satisfactory (Blamey & Martin, 2009; Johnson & Cox, 2013; Mueller & Powers, 2001; Mulla, 2013; Shi et al., 2007; Zhang, 2012). This is the case even after the period of time required to adapt to their new loudness profiles. Table 10 summaries some hearing aid studies that found hearing aid users rate amplified soft sounds louder than the target loudness (very soft), yet show high satisfaction with this loudness level.

Table 9: Loudness rating of soft sounds by hearing aid users.

Study	N	Age	Method	Loudness Measure	Result
Mueller & Powers (2001)			Looked at loudness rating by new and experienced HA users and compared them to norms	Categorical scaling	By week 8, new users report soft sounds are comfortable, but slightly soft (closer to a normal but not completely normal)
Shi et al. (2007)	10 normal hearing 12 HA users	Normal: 19-25 HL: 66-84	Looked at relationship between loudness growth function and PAL using participants' current HA in "as is" condition.	• Absolute magnitude estimating (500 and 2000 Hz warble tones) • PAL	Soft sounds were the least normalized, but rated with highest satisfaction
Blamey & Martin (2009)	61 (from 4 different studies)	28 - 92	Real life loudness and satisfaction ratings of 18 environmental sounds (unaided-aided WDRC and aided ADRO) after 4 weeks of HA use	PAL (real-life version)	All input sound levels were rated higher than normal but only soft sounds got higher satisfaction ratings compared to average and loud. Sounds that were very soft or inaudible also received low satisfaction ratings
Zhang et al. (2012)	18 Control 18 experiment	X = 64	Control group: fit to target and activated training (learning) one month later. Exp. Group: fit to target and activated training in day one. Collected measures after 8 weeks.	PAL	"Even with the adaptation period, the control group still wanted less amplification for soft sounds – but not for moderate or loud sounds." At end of experiment: soft sounds rated louder and rated higher satisfaction by both groups

Table 9 (continued).

Johnson & Cox (2012)	15 new HA users	62-82 (X=71)	All 15 were fitted with four different HA pairs (fitted with a 5-step structured fitting and verification process "following best fitting practices") and completed the PAL after one month of each pair use in the field.	PAL	Soft and average sounds rated louder than normal. Loud sounds rated softer. Participants showed "acceptable self-reported loudness perception" for all input levels (high satisfaction in PAL). Again soft sounds showed the highest satisfaction.
Mulla et al. (2013)	30 HA users	X = 63	Obtained aided (as is) ratings	• Loudness contour (500, 2000, speech) • PAL	Normal loudness perception for average and loud sounds (as measured in both measures) was returned to >70% of participants. Soft sounds were the least normalized. However, all levels rated with high satisfaction.

Audiologists have tried to address the problem of soft sound perception in various ways. One approach has emphasized counseling new hearing aid users on the importance of the adjustment period to new soft sounds and setting realistic expectations.

Another approach involves the use of expansion in programming hearing aids. Like compression, expansion delivers a different amount of gain to different input levels. However, unlike compression, with expansion the amount of gain increases with increasing level. To be able to apply expansion in a hearing aid's input/output (I/O) function, the point where gain application begins is increased, so a certain range of very low-level sounds are not amplified.

This phenomenon of louder perception of soft sounds in new hearing aid users can be explained in two ways: First, this phenomenon might support the concept of SI that proposes abnormally large loudness at elevated thresholds. Accordingly, once low-level sounds are amplified to be audible right above the elevated threshold in listeners with hearing loss, these sounds are expected to be perceived louder than how the same low-level sounds are perceived by a listener with normal hearing. If the loudness growth function for listeners with cochlear hearing loss is linear and modeled accurately in a hearing aid algorithm, there will not be a need to use excessive expansion. The current clinical practice of frequent need to use expansion is compatible with the concept that low-level sounds that are amplified to the audible range right above their elevated thresholds for listeners with hearing loss are disturbingly loud. It also may indicate an aggressive use of compression to compensate for the nonlinearity loss. The SI model, however, suggests that normal nonlinearity is not totally lost and that the loss of nonlinearity starts at a higher-level point.

Second, this phenomenon might be explained by the fact that these soft sounds are the category of sounds that were mostly inaudible to the listener with hearing loss prior to amplification. Although moderate and loud sounds are perceived differently after amplification (moderate sounds that were soft are now moderate and loud sounds that were moderate are now loud), they were always audible and within the dynamic range of a listener with mild-to-moderately severe hearing loss. Soft sounds, however, present a different scenario. When they become audible to a listener with hearing loss, despite the fact that they will occupy a previous category of sounds that were soft before amplification, they are still unique in nature. Previous sounds that were perceived as soft for a listener with hearing loss are actually moderate sounds for a normal hearing listener, and those types of sounds have different acoustical properties than soft sounds. When real soft sounds are amplified to be perceived as soft for a listener with hearing loss, their unique features might make it hard for the auditory system to accept them. Further research is needed in this direction.

Hearing aid research has just started to incorporate the two models of soft sound perception in listeners with hearing loss. One of the new fitting rationales of hearing aid manufacturers is to adjust access to soft speech based on the individual's loudness perception. Their idea is that instead of providing full audibility to all sounds available to listeners with normal hearing, soft sounds should be made audible only for those listeners who will naturally perceive them as soft (Beck, 2015).

Besides the change in hearing aid fitting rationales that are expected to occur if individual embrace the concept of SI; the goal of a hearing aid fitting in terms of returning normal loudness perception across all loudness categories will shift as well. The current paradigm is to counsel new hearing aid users on the importance of the adaptation period in which the auditory system gradually accepts loud amplified soft sounds and perceives them normally as soft. Acknowledging the concept of SI might alter this approach and encourage clinicians to counsel new hearing aid users differently, with more realistic expectations. Perhaps a new hearing aid user will be told that sounds that are perceived as soft by normal listeners around you will be perceived as louder by you.

2.2.2 Physiologic and psychoacoustic rationales for softness imperception

While the relatively new concept of SI has not been sufficiently tested, the literature review in the previous sections of this paper describes the physiologic and psychoacoutic rationales that have been proposed to explain the existence of the concept of SI (i.e., mainly the abnormally large loudness at elevated threshold due to cochlear hearing loss). These rationales are listed below:

The concept of SI aligns with the mechanics of the basilar membrane and IHC/OHC physiological function (Buus et al., 1997; Schlauch et al., 1998)

The BM gain provided by the function of OHCs is linear and constant within 10-15 dB of threshold (Ruggero et al., 1997; Yates, 1990). The nonlinear gain to BM starts at levels that are higher than 15 dB and therefore, the loss of OHC and the loss of BM nonlinearity should not impact the linear BM growth of displacement at low, near threshold levels. Because loudness has been strongly correlated with BM response (Schlauch et al., 1998), one might expect loudness growth near threshold (up to 15 dB SL) to be the same for listeners with normal hearing and those with hearing loss.

The concept of SI is consistent with the increased AN fibers involved in sound detection

Cochlear excitation is coded by a combination of AN fibers. The small population of high spontaneous firing rate fibers (High-SR) with low threshold is responsible for coding low-level sounds at and near threshold. The larger population of medium and low spontaneous firing rate fibers (medium-SR and low-SR) that have higher thresholds is responsible for coding moderate- and high-level sounds. In the case of cochlear damage, OHCs are the first type of sensory cells that are affected in typical etiologies of cochlear hearing loss. High-SR neurons have the largest synapses and are located on the side of IHC facing OHCs. Therefore, they are more vulnerable to damage than low- and medium-SR fibers. As suggested by Florentine (2011), elevated threshold due to cochlear damage will be determined mostly by lower spontaneous, higher threshold fibers that normally code for higher loudness. If the activity of these neurons is still interpreted by the brain as high loudness, loudness at elevated threshold might be higher than normal.

Assuming that all types of AN fibers stay relatively intact with cochlear hearing loss, another proposed explanation for increased loudness at threshold is that with the increase in absolute threshold, the response range of these different fibers overlaps. High spontaneous-low threshold neurons will have an increased threshold that is similar to other types of neurons' thresholds. If these fiber populations contribute in different ways to the neural coding of loudness, and there is an increase in the number of neurons responding to sounds at threshold, a larger loudness might be perceived at threshold.

To a certain extent, current models of loudness perception incorporate physiological processes within their stages that do not include separate processing for different populations of AN fibers. There are existing models that aim to simulate physiological processes in the auditory periphery for reasons not related to intensity coding per se. For example, the MATLAB Auditory Periphery (MAP) model by (Meddis et al., 2013) is a computational model that simulates the physiological processes at different stages within the auditory system including the middle ear, the inner ear, and parts of the brainstem. What is unique about this model is that it specifically accounts for activities driven by low and high spontaneous firing rate fibers in separate channels, as well as their synapses with the cochlear nucleus. The absence of these stages in loudness models might obscure the variability of intensity coding at this critical stage. Meddis et al. (2013) elegantly concluded their chapter that described this computerized model as: "Computer models of the physiology of the auditory periphery can be used to explore normal and impaired hearing and sometimes spring surprises" (Meddis et al., 2013, p. 19).

The concept of SI agrees with increased internal noise with HL

Absolute threshold is "the minimum detectable level of a sound in the absence of any other external sounds" (Moore, 2012, p. 11). Signal detection theory defines the absolute threshold as the level of intensity at which the amount of noise that is the sum of the stimulus noise and the internal physiological noise can be distinguished by the listener as different from the internal noise alone. Some loudness growth models have accounted for hearing loss as an increase in internal noise, and its effect on loudness growth has been modeled similarly to the effect of the presence of external noise (Hellman & Meiselman, 1990). The increase in internal noise as a result of hearing loss is thought to require larger than normal loudness at threshold to attain the detectability (d") corresponding to threshold (e.g., d"=1) (Florentine et al., 2005). In other words, with cochlear hearing loss, neural noise increases and requires greater stimulus intensity which might then produce a larger loudness perception.

A recent growing body of literature is showing that continuous proper excitatory input from the cochlea to higher auditory pathways is crucial for normal sound processing. Animal model research has shown that failure to receive enough input due to hearing loss results in hyperexcitability at higher auditory pathway stations, such as the cochlear nucleus (CN) (Gröschel, Ryll, Götze, Ernst, & Basta, 2014; Rubio, 2006; Vogler, Robertson, & Mulders, 2011; Whiting, Moiseff, & Rubio, 2009) and inferior colliculus (IC) (Gröschel et al., 2014; Sun et al., 2011; Wang et al., 2013), and this sometimes results in changes to the functional properties of the auditory cortex (Syka & Rybalko, 2000; Xu, Kotak, & Sanes, 2007).

Despite the fact that some of these studies employed the conductive hearing loss paradigm (Rubio, 2006; Sun et al., 2011; Whiting et al., 2009; Xu et al., 2007), results can still be generalized to the effects of cochlear hearing loss because both types of losses result in a reduction of neural output from the cochlea. This is unlike the noise exposure paradigm, which results in a sensory loss. One might think that the noise exposure paradigm is more comparable to cochlear hearing loss in humans since both occur in the cochlea, but that might not be the case. Noise exposure studies use high levels of continuous noise that result in the total death of IHCs, which is rarely the case in human cochlear loss, where it is more progressive in nature. Other causes of cochlear hearing loss in humans are age related cochlear damage (presbycusis) and damage from ototoxic drugs, which result in reduced output from the cochlea rather than total death of the IHCs and absence of output from the cochlea.

Hyperactivity in the central auditory pathways, mediated by a reduction in inhibition, recently has been studied and linked to the mechanism of tinnitus, a perception of sounds in the absence of any stimulus often developed following a cochlear hearing loss (Middleton et al., 2011). In further clinical evidence, tinnitus have been linked as well to OHC dysfunction (Modh, Katarkar, Alam, Jain, & Shah, 2014). Therefore, if OHC dysfunction is known to be responsible for loudness perception changes in listeners with hearing loss and in the same time related to the generation of tinnitus in the same listener, a larger loudness at threshold might be expected due to the higher stimulus level needed in this listener to elicit a noticeable increase in noise above the increased internal noise (that is exhibited as tinnitus).

Moreover, hyperactivity in the central auditory pathways is often exhibited as an increase in the spontaneous firing rate across different stations along the auditory pathway (i.e., CN, IC, and auditory cortex). Increased spontaneous firing rate has been found to be proportional to the degree of hearing loss; that is the greater the hearing loss, the greater the increase in spontaneous firing rate (Mulders, Ding, Salvi, & Robertson, 2011). If firing rate is known to be one of the loudness coding theories, an increase baseline firing rate in listeners with cochlear hearing loss might be interpreted by the brain as an increased baseline of loudness and therefore, a larger loudness for sounds at their elevated thresholds and above (i.e., at threshold and throughout the dynamic range afterward). This expected loudness profile is consistent with SI.

Summary

The existing literature addressing the perception of loudness and loudness growth at and near threshold in listeners with hearing loss describes how listeners with cochlear hearing loss perceive sounds near their elevated threshold, as either exhibiting rapid growth of loudness or softness imperceptions (SI). This variability in loudness perception has always been attributed to possible individual differences in cochlear damage etiology, which have never been addressed. Despite frequent references to the two possible models, very few studies actually tested the existence of the newer model, SI. The concept of SI has possible theoretical and physiological rational behind it, which were reviewed in this section. Although not enough existing data are available to support the existence of SI in some listeners with hearing loss, unexplained clinical loudness measures of soft sounds obtained from hearing aid users which have not be adequately explained by the concept of rapid growth, make this new concept worth testing.

2.3 CONCLUSION AND POSSIBLE RESEARCH DIRECTIONS

In light of the preceding literature review of loudness perception in listeners with cochlear hearing loss, the current evidence regarding how listeners perceive sounds at or near their elevated threshold is controversial. Unlike the extensive research conducted on other aspects of loudness perception in listeners with hearing loss, very limited research has focused on this issue, and the few studies that have done so have not yielded any strong evidence to support one view or another. The same question remains unanswered: Can loudness at elevated thresholds in listeners with hearing loss due to cochlear pathology be larger than loudness at normal thresholds for healthy controls?

Figure 15 schematically summarizes the current status regarding loudness at elevated thresholds based on the results of this literature review and the research questions it explored.

There are possible reasons that the question of loudness perception at elevated thresholds has remained unanswered including the following:

1) The uncertainty of the accuracy/validity of the loudness measures that have been used to measure/examine loudness at near-threshold sound level.

2) The sample size recruited in the existing studies conducted to examine loudness at near-threshold sound level is very limited. A small sample size does not represent the individual variability that is more apparent within listeners with hearing loss.

3) Studies that examined loudness perception in listeners with hearing loss have not employed designs with stratified cochlear pathology based on the etiology or location of damage; a confound that is worth addressing.

4) Psyhcoacoustic measures of loudness that are sensitive to sound levels near threshold (e.g., reaction time) have not been combined with another objective physiological measure to validate the perception's magnitude and nature of participants' responses.

5) Loudness studies that examined loudness at near-threshold levels did not include a control group of normal hearing listeners across all test frequencies to control for psychoacoustic effects on the study tasks that are not related to loudness (e.g., frequency effect, neural sound deprivation effect, etc.). Studies mainly used normal hearing at some frequency or contralateral normal hearing ears as control.

Addressing the previous literature limitations will facilitate answering these questions which in turn will clarify some critical components that should be accounted for in the theoretical framework of loudness perception in listeners with cochlear pathology, such as the description of loudness recruitment and how to best describe it and or individualize it for different listeners. Answers to these questions may also shift current paradigms in hearing aid fitting loudness goals and loudness outcomes.

Based on the previous literature review, to address the controversy of loudness at elevated threshold, the same question is needed to be asked again. Is loudness at elevated threshold in listeners with cochlear hearing loss different/greater than loudness at a normal threshold? This question is indeed questioning an underlying assumption of all loudness models for listeners with hearing loss. An assumption that has never been questioned and therefore, never been addressed. Research studies that overcome some of the limitations of the studies that addressed the controversy of loudness at threshold are needed. A good start is a study that 1) includes a larger sample size that could include a wider range of loudness function profiles among listeners with hearing loss, 2) employs a physiologic correlate of loudness in order to add a different prospective to the argument, and 3) combines the physiologic objective measure to a subjective psychoacoustic measure. The next section (5.0) will discuss using pupil dilation response to sound as a correlate to its loudness. The section will provide rationales for the appropriateness of this method to answer research questions related to loudness at low sensation levels.

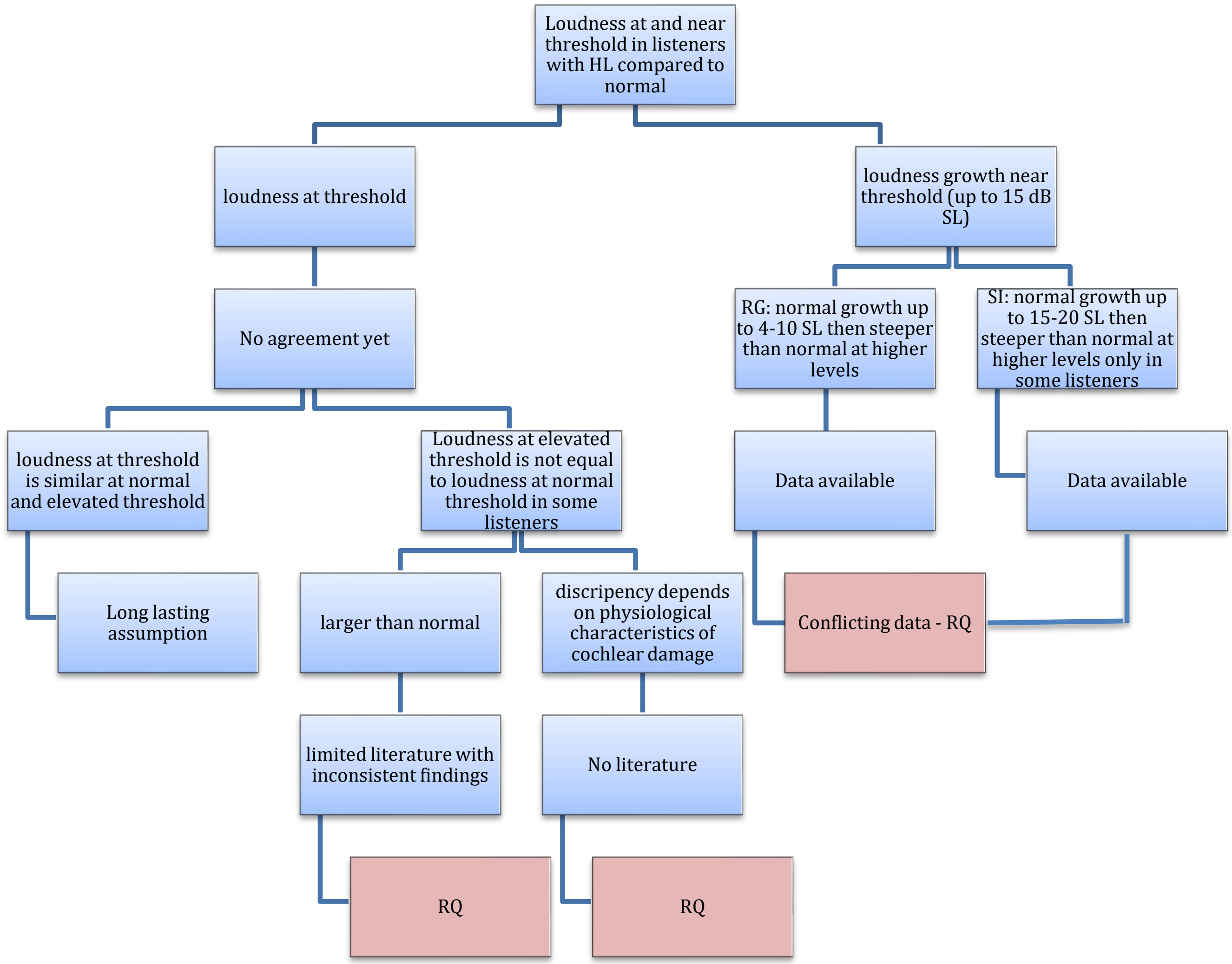


Figure 15: Summary of the status of the literature investigating loudness at elevated thresholds and where research is needed.

3.0 LOUDNESS AND PUPILLOMETRY: CAN PUPIL DILATION RESPONSE TO SOUND BE A CORRELATE TO THE PERCEPTION OF ITS LOUDNESS?

The methods used in the limited literature that has examined loudness perception at low sensation levels in listeners with hearing loss were all psychoacoustical measures, including loudness matching (Moore, 2004) and reaction time in those studies that examined behavioral responses (Florentine et al., 2005). Despite the limitations of exclusively using these psychoacoustic measures, as discussed in section 3.0, the findings of these studies showed conflicting results. Testing the model of softness imperceptions using a physiologic measure will complement the literature based on psychoacoustic measures and add depth to our understanding of loudness perception.

3.1 THE USE OF PUPIL DILATION IN THE LITERATURE

William Shakespeare once said that the eyes are the windows of the soul. Measurement of eye movements has been used in cognitive psychology research for decades. More recently, the measurement of pupil dilation within the eye has been used as a correlate to a wide range of mental states, from emotion (Bradley, Miccoli, Escrig, & Lang, 2008; Bradshaw, 1967; Partala & Surakka, 2003), to attention (Eldar, Cohen, & Niv, 2013; Gabay, Pertzov, & Henik, 2011; Privitera, Renninger, Carney, Klein, & Aguilar, 2010), decision making (Einhauser, Koch, & Carter, 2010; Einhäuser, Stout, Koch, & Carter, 2008; Lavín, San Martín, & Jubal, 2013; Preuschoff, Marius't Hart, & Einhäuser, 2011), memory (Goldinger & Papesh, 2012; Naber, Frässle, Rutishauser, & Einhäuser, 2013), and even effort and processing load (Kahneman, Beatty, & Pollack, 1967; Koelewijn, de Kluiver, Shinn-Cunningham, Zekveld, & Kramer, 2015).

The pupil constricts in response to light in what is known as the pupillary light reflex. The primary function of the pupillary light reflex is to regulate the amount of light entering the eye in order to maintain visual acuity when illumination changes in the environment (Lowenstein & Lowenfeld, 1962). In conditions where the illumination stays constant, the pupil has been observed to change in size under physiological and psychological factors (Goldwater, 1972; Tryon, 1975) and also in relation to cognition, which has been investigated for over 100 years (Bumke, 1911, as cited by Beatty, 1982b).

3.1.1 The physiology behind pupil dilation

Pupil diameter change is controlled by two iris muscles: the pupil dilator muscle and the pupillary sphincter. The activity of the dilator muscle is mediated by the sympathetic pathway and the pupillary sphincter is under the control of the parasympathetic pathway, mediated by the Edinger–Westphal complex of the oculomotor nucleus (Steinhauer, Siegle, Condray, & Pless, 2004). Therefore, the pupil dilation response reflects the contribution of the sympathetic (SNS) and parasympathetic (PNS) nervous systems as they act in a reciprocal manner (Loewenfeld, 1999). When the pupil dilates, a sympathetic activation occurs and stimulates the dilator muscle and a parasympathetic inhibition occurs to the Edinger-Westphal complex that relaxes the sphincter muscle and vice versa in the case of pupil restriction.

Current evidence suggests that cognitive-, task-evoked pupillary response is associated with the activation of the locus coeruleus-norepinephrine (LC-NE), that is modulated by the sympathetic system (Sara, 2009) which in turn modulates prefrontal attentional control (Unsworth & Robison, 2016). However, this specific chain of neural activation is not yet fully understood (see discussion section in (Ayasse, Lash, & Wingfield, 2017; Kuchinsky et al., 2014)

In auditory stimulation tasks, the activation of the superior and inferior colliculus within the auditory pathway have been found to evoke a pupillary response (Netser, Ohayon, & Gutfreund, 2010; Wang, Boehnke, White, & Munoz, 2012). Therefore, different auditory activations of the superior colliculus might cause different pupillary dilation responses.

3.2 THE USE OF AUDITORY EVOKED PUPILLARY DILATION IN THE LITERATURE

The *auditory* evoked pupillary dilation is a promising indicator of auditory discrimination. It has been used as an indicator for frequency, spatial and intensity discrimination in animals (Bala & Takahashi, 2000; Khahnemann & Beatty, 1967; Wang, Boehnke, Itti, & Munoz, 2014) and in humans (Liao, Yoneya, Kidani, Kashino, & Furukawa, 2016b; Wang & Munoz, 2014). The majority of pupillometry studies that have recorded pupillary response to auditory stimuli have shown a response when the property, probability, or intensity of the stimulus was manipulated.

When an *auditory*-evoked pupillary dilation is obtained as a response to a property change in a continuous auditory stimulus presented in a repetitive manner without tasking the observer to do anything, this phenomenon is referred to as the *orienting reflex*. The orienting reflex is defined as the unit of attentional processing responsible for directing an organism towards changes in its environment (Barry, 2009). The reflex is sensitive to the slightest change in environmental conditions and thought to be essential for organisms' ability to detect novelty in their environment. A major determinant of this reflex is the novelty of the stimulus, which is operationalized by the habituation of the response with stimulus repetition. In psychophysiological research, pupil dilation is looked at as one of the physiologic changes in the orienting reflex context. Other physiologic correlates include change in heartbeat, skin conductance, etc. Orienting reflex can manifest at different levels: behavioral, muscular, sensory, cortical, and autonomic (Deckers & Hricik, 1984). By contrast, the *defense reflex* (DR) is a different class of response to high intensity stimuli. Research conducted on pupillary response to sound has mostly looked at the pupillary response as an orienting reflex.

This body of literature shows that the pupil dilates in response to a novel sound; however, this pupillary response in the orienting reflex context does not necessarily correlate with the magnitude of the perception such as the perception of loudness, which is what the Primary Investigator (PI) is interested in. According to Groves and Thompson's (1970) Dual-processing Theory that models the preliminary processes of the orienting reflex, the pupillary dilation response to sound as an orienting reflex reflects a process that precedes any perceptual process. Therefore, it does not reflect the magnitude of loudness perception of an auditory stimulus.

In contrast with the pupillary orienting reflex, a *task*-evoked pupillary response can manifest perception properties of the evoking signal. Because the listener is tasked to attend and reflect to the signal, the evoked response does not habituate, and indeed reflect aspects from the evoking signal as well as the perception and cognitive processing of the signal (Beatty, 1982b). The task evoked pupillary response occurs at short latency following the onset of processing and fades quickly once processing is terminated.

As mentioned earlier, *auditory* task-evoked pupillary responses have been obtained for a range of cognitive processing mostly for language processing, attention, memory, and complex reasoning. There was limited use, however, for sensory detection, in which the PI is interested to answer the research question. The details of these studies and their findings will be discussed in the following section in this document (section 5.3).

3.3 PILOT STUDY: ITS MOTIVATION, DETAILS, AND FINDINGS

There has been limited use of the pupillary response as an indicator for perceived loudness and no study has ever looked at pupillary response to sounds presented at low sensation level (<10 SL). In order to answer the research question proposed previously in this document which would investigate the magnitude of perceived loudness to sounds presented at low sensation levels using the pupillary dilation response as a correlate to loudness, three properties of the pupillary response to sound should have been be documented in pupillometry research studies: 1) the pupil shows dilation in response to auditory stimuli in a sound detection task; 2) the pupil dilation response magnitude increases with increasing sound level; 3) the change in pupil response with increasing sound level reflects the change in perceived loudness by listeners; and 4) the pupil shows a dilation response to ***low-level*** auditory stimuli (low sensation level), i.e., the pupil dilation response is sensitive to low-level sound detection.

1) *The pupil shows a dilation response to auditory stimuli detection presented in quiet?*

Pupillary dilation response is sensitive to novel acoustic changes and is an index of the orienting reflex. For example, significant pupillary dilation response has been found to white noise oddballs against a background of repetitive pure tones, and not to pure tone oddballs (Liao et al., 2016b). This is a response to an oddball effect that is not in the intensity domain, but in the spectral sound domain. Therefore, finding a pupillary dilation in response to sound detection in this study does not necessary support the use of pupillary response in a sound detection task in quiet.

Dyback and Wallgren (2016) looked at the sensitivity of pupil dilation response to sound detection in adults. The researchers collected pupillary response to pure tones at frequencies 500, 1000, 2000, and 4000 Hz, presented at four intensity levels: silence, 30, 50, and 70 dB SPL. Twenty listeners with normal hearing were recruited in the study. They looked at the difference in pupil dilation between silence conditions and the different intensity levels within each frequency. Participants fixated on a plus sign on the screen located in front of them. The pure tones were presented at 500, 1000, 1500, or 2000 ms delay. The stimulus duration was 500 ms followed by a 2500 ms inter-stimulus interval. The results showed that pupillary responses occurred between 300-2000 ms post-tone onset, a time window that was earlier than the response time window previously reported with spoken words (Silk et al., 2009). The results also showed that the difference between the silence condition and different sound intensity levels was only statistically significant for 4000 Hz tones.

This study is one of the few that employed pupillometry during a sound detection task, and its findings are promising for collecting a pupillary response to tones of levels as low as 30 dB SPL. However, several limitations of the study were noted as follows:

The test room where the pupil data was collected was a well-lit room as reported by the authors, and the brightness of the eye tracking screen was 148 cd/m^2. The excessive brightness of the environment might have activated the parasympathetic system and restricted the pupils, which may have opposed pupil dilation in response to the evoking signal.

The intensity levels of the signals were not consistent throughout the experiment across subjects. The signals were presented in the sound field and sound level meter measurements that were obtained to verify signal levels varied across subjects and across trials. For example, the measured values for the 30 dB sound level revealed measurements between 33.1 up to 38.5 dB SPL (mean of 35.2). For the 50 dB sound level, measurements varied between 40.2 up to 53.8 dB SPL (mean of 48.5). This variation in sound level might have resulted in variation in pupillary response. This variation in presentation level is expected to evoke different (smaller and/or larger) pupillary dilation for the same intensity level condition across participants, a variation that is not due to the variation in the perception per se, but to the variation in the evoking signal property. Additionally, normal hearing was verified by self-report only and was not verified with any objective hearing screening procedure. Participants were only asked to fixate on a plus sign on the screen and were not involved in any task or asked to respond to any stimuli, which is known to distort the cognitive status and make it harder to interpret the pupillary response as a response to detection only.

In another study conducted in Japan by Liao et al. (2016a), the pupillary responses to 500 ms duration environmental sounds were recorded from eight listeners who stared at a fixation point and listened to the auditory sequence without being involved in any task. The environmental sounds were 1000 Hz tone burst, white noise, beep, bird, chirp, tone, laughter, crying, phone, dog, and scratch, and were presented randomly with 10 seconds inter-stimulus intervals. The results showed that the pupil dilated in response to sounds regardless of the sound type.

Based on the findings of the above studies, the pupillary response to sound in a detection in quiet task is documented in the literature and therefore, supports the use of the pupillary response to sound as correlate to at least, sound detection.

2) The pupil dilation response to sound changes in magnitude as the sound level increases

Studies that have examined the change in pupillary dilation response with increasing sound intensity have shown an increase in magnitude with increasing sound level. These findings have been consistent when using different acoustic signals, such as pure tones (Dybäck & Wallgren, 2016; Liao et al., 2016a; Liao et al., 2016b; Nunnally, Knott, Duchnowski, & Parker, 1967), bursts of white noise (Liao et al., 2016a) and broadband noise (Antikainen & Niemi, 1983b). However, the presentation levels of the signals employed previously were all well above threshold (the lowest level being 30, and the highest being 85 dB SPL/100 dB A) (Antikainen & Niemi, 1983a; Dybäck & Wallgren, 2016; Liao et al., 2016a; Liao et al., 2016b; Nunnally et al., 1967). No studies have examined this relationship when low-level signals were the eliciting signal of the pupillary response. On the other hand, other studies did not find this relationship with similar sound levels in cochlear implant users (Kim et al, 2015).

The findings of the above studies support the use of the pupillary response to sound as a correlate to change in intensity of the signal which is a prerequisite to be a correlate to change in loudness perception.

3) The change in pupil response with increasing sound level reflects the change in perceived loudness

In the study by Liao et al. (2016a), the pupillary dilation response predicted both the salience and loudness subjective judgments of the paired sounds. The authors also concluded that subjective salience of sounds is defined or heavily influenced by loudness. When the loudness of these sounds was calculated using the Glasberg and Moore (2002) loudness model, a regression analysis showed a strong linear correlation between the pupillary dilation response and loudness. When the loudness was the same, the pupil dilated in similar manner.

The findings of this study support the use of the pupillary response to sound as a correlate to its loudness change at least when the signal is presented at a high sensation level.

4) *The pupil shows a dilation response to* ***low-level*** *auditory stimuli (low sensation level)*

The detection of low-level sensory signals is one of the most basic cognitive tasks in humans (Beatty & Lucero-Wagoner, 2000b). To the best of the author's knowledge, only two pupillary response studies looked at the response to sensory signals presented at near-threshold intensities. Only one of them is published. Hakerem and Sutton (1966) collected pupillary responses in a visual perceptual detection. Their data showed a pupillary response to light presented at threshold when it was detected, where the response was absent when the stimulus was not detected. When the participants were not asked to report whether or not a stimulus was presented, the pupillary response was absent unless the stimulus was considerably above threshold. Beatty and Wagoner (1977) in an unpublished study, extended the visual detection findings to an auditory detection task and found similar findings. The task involved signal in noise detection of a 1000 Hz tone presented in background noise. They reported a pupillary response only if the signal was detected. The response was also larger when the listener was more certain that the signal was presented compared to when the listener was uncertain.

Summary

Taken all together, the pupillary dilation response to sound is a phasic response that has been observed in the literature from normal hearing listeners. The previous literature has only employed sounds that are at comfortable or loud levels (well above threshold of hearing). This response grows in magnitude with increasing sound level, and in loudness studies, the magnitude of the response correlated to the perceived loudness.

Because pupillary response to sounds as a correlate to their loudness has only been employed in studies that looked at the response to high sensation level sounds (> 40 dB SPL), it is unclear whether the pupillary response is also sensitive to low-sensation level sounds, i.e., at and near hearing threshold. It is also unclear if the response will be characterized by a similar growth pattern as the intensity level of the sound increases. Therefore, as a first step, a pilot study was required to examine if pupillary dilation response occurs when sounds are presented in quiet at very low sensation levels to listeners with normal hearing, and to examine the behavior of the response at low sensation level.

3.3.1 Pilot Study

Pupil diameter data were collected from four normally hearing young adults, aged from 20 to 25 years, in a sound detection task. Five hundred millisecond pure tones of low (500 Hz) and high (4000 Hz) frequencies were randomly presented at different sensation levels (-4 SL, 0 SL, 4 SL, 10 SL, and 16 SL) as shown in Figure 15. Participants were asked to fixate on a square fixation point on a screen that changed into a plus sign after 3000 ms indicating the beginning of a period where a tone might be presented. The signal was presented at 2000 ms, 3000 ms, or 4000 ms from the onset of the plus period. Three thousand milliseconds post tone onset, the plus sign changed into a square again. When the plus changed back to a square, the participant responded if the tone was heard by pressing a yes/no button (Figure16). Trials of silence where no sound was presented were randomly presented within each block with the same number of occurrences as each sound level. The details of the trials' design, apparatus, and acoustic stimuli generation and presentation used in the pilot study were the same as would be used in the main proposed study, and will be discussed in detail in the methods section.

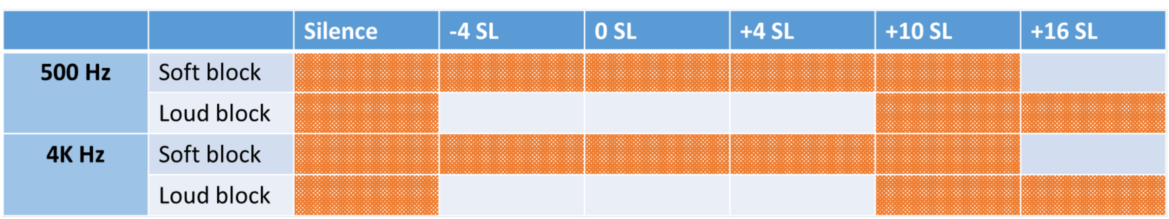


Figure 16: Design of the pilot study

Tones were presented in four blocks, 500 Hz/soft, 500 Hz/loud, 4000 Hz/soft, and 4000 Hz/loud. Soft blocks included the presentation levels: -4, 0, 4, and 10 SL. The loud blocks included the presentation levels: 10 and 16 SL. The order of completing the blocks was counterbalanced between subjects and five-minutes breaks were given between blocks.

The reason for splitting the presentation levels into soft and loud blocks was to observe if the pupillary dilation response to the 10 SL condition differed in magnitude when presented with softer versus louder sounds. It was hypothesized that the perception of a sound would be influenced by the context it which it was presented. The 10 SL sound was hypothesized to be perceived louder when presented in the soft block, yielding larger pupillary response than when it was presented in the loud block. This finding would be supportive of the pupillary response magnitude as a correlate to the perception of loudness.

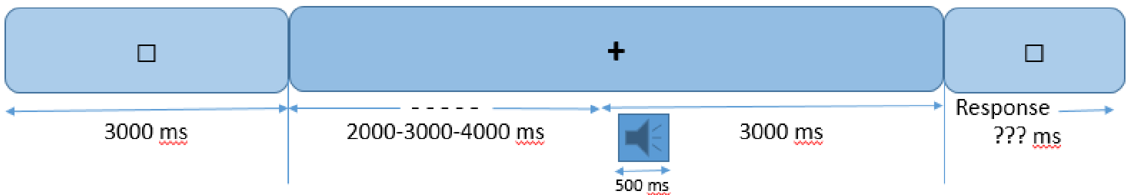


Figure 17: Design of trials in pilot study.

Results

In the soft blocks, pupillary dilation data were obtained from three participants in response to 500 Hz tone and from four participants in response to a 4000 Hz tone. In the loud blocks, pupillary dilation data were obtained from two participants in response to both tone frequencies.

Figure 17 shows the percent of "yes" responses at each presentation level for all participants. On average, tones presented -4 dB below threshold were detected 20% of the time. This percentage is appropriate for a perception of a tone that is slightly below threshold. Tones presented at threshold (0 SL condition) were detected on average 50% of the time (Figure17). This criterion is slightly different from the 75% criteria that is claimed from the threshold procedure employed to determine hearing threshold in the pilot study. Tones presented at 4 SL were detected 80% of the time and tones presented at 10 SL and 16 SL were detected 100% of the time due to their presentation well above threshold. The instrumentation used to set the presentation level might not have produced the exact targeted sensation level presentations due to the use of an audiometer as an attenuator after transforming the obtained threshold from SPL to dB

HL, instead of an attenuator that attenuates the obtained threshold in SPL. However, the percentage of "yes" responses indicated that presentation levels were approximately set at the targeted sensation levels.

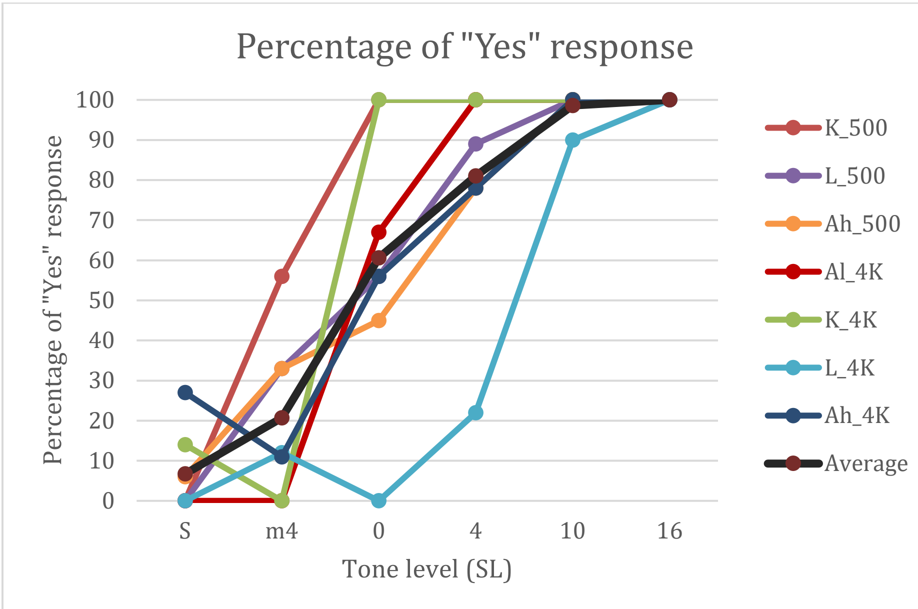


Figure 18: Percentage of "Yes" responses at different levels.

The flow chart below (Figure 18) shows the average pupillary dilation response from all participants post-tone onset. Responses to sounds at each level were separated according to when the participant responded "yes", and when the participant responded "No", and they were graphed separately. Each graph is the average pupillary response as a function of time post-tone onset.

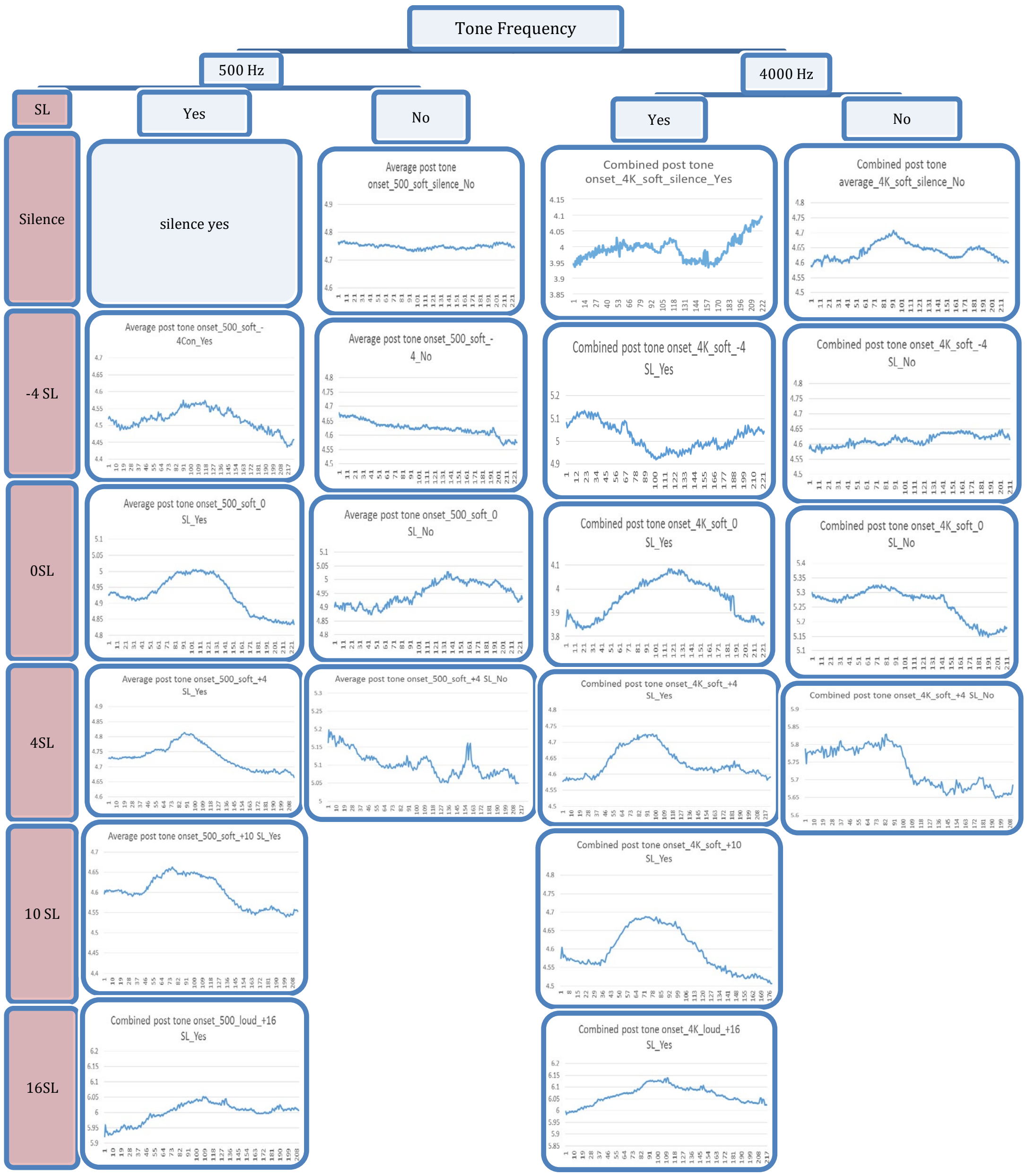


Figure 19: Average pupillary response of all participants post-tone onset. Columns of graphs from left to right are responses to 500 Hz tones when participants responded "yes"; 500 Hz tone when participant responded "no"; to 4000 Hz when they responded "yes', and to 4000 Hz when they responded "no", respectively. Each row shows graphs of one presentation level. The first row is when no tone was presented. Second to sixth rows are for -4, 0, 4, 10, and 16 SL presentation levels, respectively. In each graph, the x-axis represents time (60 units per second) and the y-axis represents absolute pupil diameter in mm.

Pupillary peak dilation (the difference in pupil diameter from the baseline diameter) was then calculated for each condition. For each condition, the highest value in the pupillary response post tone onset was corrected by subtracting the average of the first 500 ms post tone onset. The data of the 500 ms before tone onset were unavailable at the time of the calculation. The pupillary peak dilation was then averaged across participants for each presentation level. Tables 10 and 11 show the average pupillary peak dilation for each presentation level.

Table 10: Average pupillary peak dilation value and latency of peak for 500 Hz tone for different sound levels. The values in the soft blocks are averages of three participants. The values of loud blocks are averages of only two participants.

	Soft blocks				Loud blocks	
	-4 SL	0 SL	4 SL	10 SL	10 SL	16 SL
Peak dilation (mm)	0.072	0.081	0.083	0.061	0.078	0.109
Latency (sec)	1.6	1.7	1.5	1.3	2.0	1.8

Table 11: Average pupillary peak dilation value and latency of peak for 4000 Hz tone for different sound levels. The values in the soft blocks are averages of three participants. The values of loud blocks are averages of only two participants.

	Soft blocks				Loud blocks	
	-4 SL	0 SL	4 SL	10 SL	10 SL	16 SL
Peak dilation (mm)	-	0.152	0.085	0.082	0.057	0.132
Latency (sec)	-	1.9	1.6	1.4	1.8	1.8

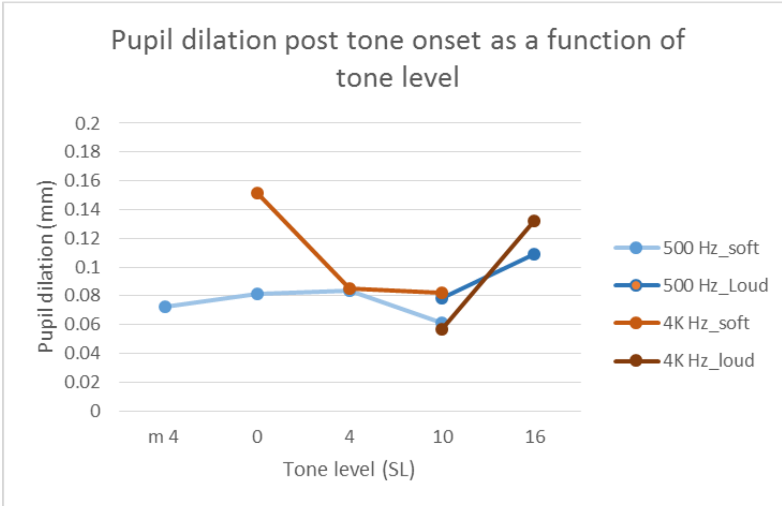


Figure 20: Average pupillary peak dilation for each presentation level.

Conclusion and Discussion

Based on the pilot data, the pupil dilates in response to low sensation level sounds. The pilot data are consistent with the opinion of cognitive psychophysiologists that pupillary response, "Is only a correlate of cognitive intensity, hence the marker is indirect and not casually linked" (Just and Carpenter, 1993, p.312). Pupillary dilation response to auditory stimuli should not be assumed to occur only in response to external sensory events because emotions, mental processes, and attentional effort also contribute to the systematic change in pupillary diameter. Descriptive evaluation of the results of the pilot study showed that the pupillary responses to sounds in the study were larger in amplitude when the tone was reported as heard than for when they were reported unheard. Similar results were seen by Haider et al (1964) when they examined human evoked potentials to visual stimuli close to threshold. Responses were of larger amplitude when the stimulus was seen, than when stimulus was not seen.

It was predicted that the pupillary response to sounds would increase in magnitude with increasing sound level. However, unexpectedly, the pilot data showed either similar (as in 500 Hz condition) or larger (as in 4000 Hz condition) pupillary dilation response to sounds presented around threshold (-4, 0, and 4

SL) and were heard (perceived) by the listener, compared to sounds presented at higher sensation levels (10 and 16 SL).

This finding can be explained by the difference in the task difficulty. Detecting sounds around threshold level is a more difficult task that involves uncertainty, unlike detecting a comfortably perceived sound where the listeners is more certain about the detection. Davis (1964) looked at auditory evoked potentials and reported an increase in amplitude when subjects had to make a difficult sensory discrimination. Sutton, Braren, Zubin, and John (1965) found that auditory evoked potentials increased in amplitude when subjects were uncertain if the stimulus would be presented compared to when subjects knew the stimulus was always going to be presented.

It can be concluded that the characteristics of pupil dilation in response to a sound might depend on the context in which it is presented (i.e., sound detection around threshold of hearing versus sound detection at a comfortable level). Pupil dilation in response to sounds presented at low SLs is greater due to increased difficulty of detection, uncertainty, and the requirement of making a decision. Pupil dilation in response to sounds at higher SLs increases with increasing level, which is the manner previously seen in the literature.

The proposed study aims to utilize the power of pupil dilation response to sound as a physiologic measure of its perceived loudness. Pupil dilation response has many advantages over the previously used psychoacoustic measures to address the question of loudness at low sensation levels. Advantages include that it is a physiologic response that requires no training, and that it is an integrative measure that includes information from perception, as well as cognitive processing of a stimulus; therefore, it could reflect aspects of stimuli detection, as well as higher aspects of perception, such as the magnitude of that perception, (i.e., perceived loudness).

4.0 RESEARCH QUESTION, SPECIFIC AIMS, AND HYPOTHESES

This study aims to investigate whether the loudness perception of sounds at low SLs (SLs at and near threshold) in individuals with cochlear hearing loss is greater than in individuals with normal hearing. In other words, whether hearing loss alters the perception of loudness of low SL sounds (0 to 10 SL) in individuals with cochlear hearing loss and causes these individuals to perceive low SL sounds louder than how normally hearing individuals perceive them. The concept of softness imperceptions (greater loudness of sounds at elevated threshold) is suggested by relatively recent studies (Buus & Florentine, 2002; Buus et al., 1998; Marozeau & Florentine, 2007), but is insufficiently tested (Florentine et al., 2005; Moore, 2004). Because this new concept conflicts with the long-lasting description of recruitment (abnormal growth of loudness above threshold), which describes how individuals with cochlear hearing loss perceive loudness of sounds across levels, research studies are needed to enhance our understanding of the impact of cochlear hearing loss on the loudness perception of low-level sounds. The current knowledge base on the loudness growth function for listeners with cochlear hearing loss implicitly assumes equal loudness for sounds at threshold whether the threshold is normal or elevated (due to cochlear hearing loss).

Sufficient evidence suggests that the magnitude of pupillary dilation in response to sound reflects the magnitude of the level of sound presented (Antikainen & Niemi, 1983a; Dybäck & Wallgren, 2016; Liao et al., 2016b; Nunnally et al., 1967) and there is some evidence to support that the magnitude of pupillary dilation also reflects the perceived loudness (Liao et al., 2016a).

In this investigation, pure tones were presented at five sensation levels around the participants' hearing threshold. Pupillary diameter data were collected from two groups: listeners with normal hearing thresholds (NH group) and listeners with mild-to-moderate cochlear hearing loss, i.e., elevated hearing thresholds (HL group). Pupil dilation response was intended to be used as an indicative measure of the magnitude of the perceived loudness of sounds presented at low sensation level.

Based on the gap in the literature on how cochlear hearing loss alters the loudness perception of low sensation level sounds, the following research questions were addressed:

Research Question: Is perceived loudness at low sensation level (0 to 10 SL) different in listeners with cochlear hearing loss compared to listeners with normal hearing?

According to the loudness model by Moore and Glasberg (2004), the perceived loudness of sounds at elevated hearing threshold by a listener with hearing loss is not different from the perceived loudness of sounds at a normal hearing threshold by a normally hearing listener. This model also proposed that loudness growth right above an elevated threshold (up to 4-10 SL) grows in a similar manner as above a normal threshold, but then grows at a faster rate above that to catch up with normal loudness perception at higher levels. In contrast, in the model proposed by Florentine and colleagues (Buus & Florentine, 2002; Florentine, Buus, & Rosenberg, 2005), the loudness at elevated threshold is different (greater) in some listeners with cochlear hearing loss compared to loudness at normal hearing threshold. This model also proposes that loudness growth above threshold in listeners with hearing loss is unchanged from loudness growth above normal threshold.

Specific Aim 1 (SA1): Determine if there is a significant difference in the pupillary dilation response to tones presented **at threshold** (0 dB SL) between listeners with hearing loss and listeners with normal hearing.

Null Hypothesis (H0): Sound presented at 0 SL (at threshold) in listeners with hearing loss will elicit a pupillary response that is similar in magnitude to a response from normally hearing listeners at normal threshold. This finding will indicate that sounds at threshold in the two groups were perceived as equally loud and elicited the same pupillary response magnitude. Equal pupillary response amplitude indicates equal difficulty in a sound detection task at threshold and equal difficulty in decision making. This finding will be consistent with the loudness model by Moore and colleagues.

Alternative Hypothesis (H1): Sound presented at 0 SL (at threshold) in listeners with hearing loss will elicit a pupillary response that is different in magnitude from a response from normally hearing listeners at threshold.

Alternative Hypothesis (H1a): Sound presented at 0 SL (at threshold) in listeners with hearing loss will elicit a pupillary response that is *smaller* in magnitude than a response from a normally hearing listener.

Because there are no previous studies that have collected pupillary response from normally hearing listeners in an auditory threshold detection task, this hypothesis is based on the finding of the pilot data. The pilot data obtained from normally hearing listeners showed greater pupillary dilation in response to sounds presented at threshold (0 SL), compared to sounds presented at higher level. This finding was interpreted to mean that the pupillary response at threshold incorporated other aspects of the context in which the signal was presented. A signal at threshold is more difficult to detect than a signal presented above threshold, and the context of presenting the signal at threshold requires the listener to make a harder decision than when the signal is presented at a more comfortable level in which the signal can more easily be detected. This interpretation is consistent with the typical finding in the listening effort literature of increased pupillary dilation with increased effort (Beatty, 1982b; Kahneman, 1973) and the findings of electrophysiology studies related to a detection task that requires a decision (Davis, 1964; Sutton, Braren, Zubin, & John, 1965).

Alternative Hypothesis (Hb): Sound presented at 0 SL (at threshold) in listeners with hearing loss will elicit a pupillary response that is *greater* in magnitude than a response from a normally hearing listener at threshold.

While this finding was less likely to occur based on the theoretical rationale for the first hypothesis, it would indicate that listeners with hearing loss perceived the loudness of sounds at threshold differently from normally hearing listeners. However, depending on the pattern of pupillary response change obtained from the normally hearing listeners across other levels included in the experiment, the interpretation might change.

Specific Aim 2 (SA2): Determine if there is a significant difference in the pattern of change of pupillary dilation response with increasing sound level within the range of **low sensation levels** (4 dB SL, 10 dB SL, and 16 dB SL) between listeners with hearing loss and listeners with normal hearing.

Null Hypothesis (H0): The pattern of change in pupillary response magnitude across sound levels above threshold will be *different (steeper)* in listeners with hearing loss compared to normally hearing listeners. This finding would be consistent with the Moore, et al model (Moore & Glasberg, 1997; Moore & Glasberg, 2004).

In the loudness model by Moore and colleagues, it is proposed that loudness at elevated threshold is equal to loudness at normal threshold and grows in a similar manner up to 4-10 SL. However, above that SL (the 10 and16 SL conditions in this experiment), the loudness grows at a faster rate in listeners with hearing loss. If the model holds true, we would expect to find a greater pupillary response magnitude to tones presented at 10 and/or 16 SL, in listeners with hearing loss, compared to normally hearing listeners.

Alternative Hypothesis (H1): The pattern of change in pupillary response magnitude across sound levels above threshold will be *similar* in listeners with hearing loss and listeners with normal hearing listeners across all tone levels (4 SL, 10 SL, and 16 SL). This finding would be consistent with the model by Florentine and colleagues.

In the Florentine, et al model (Buus & Florentine, 2002; Florentine et al., 2005), perceived loudness at an elevated threshold is greater than perceived loudness at a normal threshold, but then grows in a similar rate up to approximately 15 SL.

If this model holds true, we would expect to find one of the following:

1) A greater pupil dilation response to tones presented at 4 SL, 10 SL, and 16 SL in listeners with hearing loss compared to normally hearing listeners. This increase in dilation response across all sound levels will yield a growth rate that is similar to the growth rate in normally hearing listeners. This finding would indicate that listeners with hearing loss perceived sounds presented at those sensation levels louder than listeners with normal hearing. This finding would be consistent with the model by Florentine and colleagues.

2) A similar pupil dilation response to tones presented at 4 SL, 10 SL, and 16 SL in listeners with hearing loss and normally hearing listeners. This finding will yield a growth rate that is similar to the growth rate in normally hearing listeners. Similar pupillary response magnitude to tones above threshold (4, 10, and 16 dB SL) between the two groups would suggest that the pupillary dilation might not be a correlate of absolute loudness of these sounds, but rather a correlate to the relative loudness of the sounds perceived in the same task. This finding would be consistent with the model by Florentine and colleagues.

Specific Aim 3 (SA3): Determine if the dynamic range in terms of perceived loudness as measured by the difference in subjective loudness ratings of two tones presented at different SLs is significantly different in listeners with normal hearing and listeners with hearing loss. This aim will be used to support and complement the findings of Aims 1 and 2.

Null Hypothesis (H0): The dynamic range of perceived loudness as measured by the difference in subjective loudness ratings of two tones presented at different SLs (4 and 16 SL, 4 and 20 SL, 4 and 24 SL, 4 and 28 SL conditions) is *greater* in listeners with hearing loss compared to listeners with normal hearing.

Because the traditional model of recruitment suggests that loudness at an elevated threshold is equal to loudness at a normal threshold, and that loudness grows faster than normal above an elevated threshold, it is expected for the range of loudness between a level that is right above threshold (4 SL) and a level where abnormal growth should already be apparent (16, 20, 24, and 28 SL) to be greater in listeners with hearing loss compared to normally hearing listeners (Figure 21). Therefore, this finding would be consistent with the Moore, et al model.

Alternative Hypothesis (H1): The dynamic range of perceived loudness as measured by the difference in subjective loudness ratings of two tones presented at difference SLs (e.g., 4 SL and 16 SL condition) is *similar* in listeners with hearing loss and listeners with normal hearing.

In the Florentine, et al model, it is proposed that loudness at an elevated threshold is greater than loudness at a normal threshold, but that they both grow at a similar rate up to approximately 15 SL (Buus & Florentine, 2001, 2002). If this model holds true, we would expect the dynamic range in terms of loudness between 4 and 16 SL in listeners with hearing loss compared to normally hearing listeners to shift up, while the range itself remains unchanged (Figure 21). The model by Florentine et al suggests that listeners with hearing loss have reduced dynamic range in terms of SPL, as well as in terms of perceived loudness due to starting the dynamic range with an increased loudness.

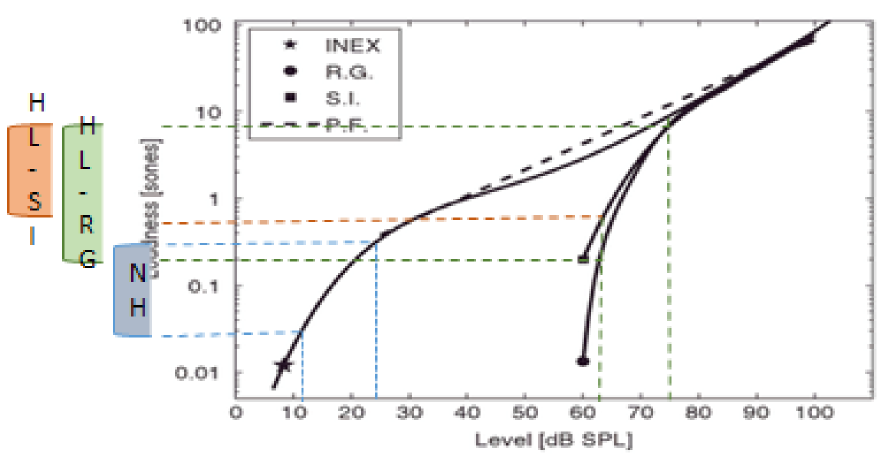


Figure 21: Vertical bars on the left display the expected range of perceived loudness difference between a 4 SL sound and a 16 SL sound in listeners with normal hearing (blue bar), listeners with hearing loss that show rapid growth RG above threshold (green bar); and listeners with hearing loss that show softness imperceptions SI above threshold (green bar).

Significance

No research group that is unbiased to the results of their study has tested the conflicted concept of loudness at threshold and low SLs in listeners with hearing loss. When new challenges arise to long-standing gold standard models, it is necessary to gather new empirical data.

The findings of the proposed study will advance our understanding of loudness perception at and near elevated threshold in listeners with cochlear hearing loss, and may lead to future changes in the definition and concept of recruitment on which psychoacousticians rely heavily in designing and interpreting results in studies involving listeners with sensorineural hearing loss. Thus, understanding loudness perception at elevated threshold has potential significance for understanding sensorineural hearing loss and for hearing aid design.

This study also aims to make a general methodological contribution to the literature. Only a few studies have used pupillary response as a correlate to loudness perception (Liao et al., 2016a) or intensity level (Dybäck & Wallgren, 2016; Liao et al., 2016b). In addition, none of the previous studies has collected pupillary response data with respect to sounds presented at low SLs (near hearing threshold). This study is the first to investigate the nature and pattern of pupillary response as a physiologic correlate to loudness perception at low sensation levels in both listeners with normal hearing and those with cochlear hearing loss. Findings of this study will add a novel contribution to the literatures of pupillometry and psychoacoustics.

5.0 RESEARCH DESIGN AND METHODS

5.1 PARTICIPANTS

5.1.1 Inclusion and exclusion criteria

The study compared the magnitude and behavior of pupillary responses to sounds presented at low SL obtained from adults with normal hearing to responses obtained from adults with sensorineural hearing loss of cochlear origin. The cochlear type of hearing loss is the type responsible for altering the function of loudness perception and for introducing the proposed presence of softness imperceptions (abnormally greater loudness at elevated threshold). All participants met the following inclusion criteria:

1. Language: Native English speakers, which was determined by asking potential participants about their native language.
2. Age: All participants had to be adults (ages 18 to 69 years old). The minimum cut off age (18 years old) is the starting age of adulthood. The maximum cut off age (69 years old) is to exclude participants with increased risk of declined cognitive functions that might impact the performance in the sound detection procedure and/or comprehension of the study instructions. Moreover, including participants older than 69 years of age would make it impossible to match the age between the two groups due to the low prevalence of bilateral normal hearing in individuals older than 69 years of age (Lin, Niparko, & Ferrucci, 2011).
3. Hearing threshold: All participants completed a standard pure tone audiometric testing. For the normal hearing group, participants had a normal air conduction hearing threshold of 25 dB HL (or

better) at 250, 500, 1000, 2000, and 4000 Hz in both ears. Hearing threshold at 8000 Hz did not exceed 50 dB HL.

For the hearing loss group, participants had elevated audiometric hearing thresholds (>25 dB HL) at least at two or more frequencies, including 4000 Hz in both ears. Hearing threshold at 4000 Hz in the right ear was within 35 to 60 dB HL. The upper limit for the hearing threshold was set to exclude hearing losses with inner hair cell loss involvement (Stebbins, Hawkins, Johnsson, & Moody, 1979).

All participants with bilateral hearing loss had symmetric hearing. Asymmetry was determined by using the pure tone average across frequencies (1-8KHz). All participants had pure tone averages that were less than 15 dB different between ears (Hunter, Ries, Schlauch, Levine, & Ward, 1999). Pure tone average is the most effective method for asymmetry calculation (Cheng & Wareing, 2012; Zapala et al., 2012). Asymmetric hearing loss was an exclusion criteria so participants in the study do not have different loudness perception profiles between ears which might impact the listener's perceptual skills.

4. Word recognition testing: All participants scored at least within the 95% confidence interval according to his or her pure tone average (Dubno, Lee, Klein, Matthews, & Lam, 1995) on the Northwestern University Auditory Test No. 6 (NU-6) (Appendix D). The pure tone average is the average of thresholds at 500, 1000, and 2000 Hz. An appropriate proportional score is important for participants in the hearing loss group in order to exclude participants with hearing loss with possible retrocochlear involvement.
5. Vision: All participants had to pass a vision screening involving a Snellen chart at 20/20 or better, with a viewing distance (3 feet= 91 cm) that is equal to that of the viewing of the computer screen in the experiment. Participants can pass the screening with or without vision correction.

Corrections were limited to surgical correction or contact lenses. Participants were not allowed to wear eyeglasses for the experiment since it may affect the clarity of the eye image during pupil recording.

6. Gender: Participants from both genders were included with no requirement for gender balance.

In order to control for factors that might impact the dynamic range of the pupillary response and factors that might result in an atypical loudness perception, the following exclusion criterion were determined:

1. Self-reported history of neurological damage, brain surgery, brain tumor, ear surgery, or ear disease.
2. Self-reported clinically diagnosed or suspected psychological or mental disorder (e.g., schizophrenia, depression, bipolar, and anxiety). Psychological disorders have been found to impact auditory and visual pupillary responses during cognitive tasks (Steinhauer, Hakerem, & Spring, 1979; Steinhauer & Zubin, 1982).
3. Abnormal pupillary light reflex and/or abnormal pupillary miosis/contraction to near vision as determined during a screening task completed by the PI. The task will be described in the screening procedures' section 7.4.1.
4. Participants with self-reported hypersensitivity to sounds (hyperacusis), which was determined by self-report and/or the screening procedure that determines the uncomfortable loudness level (UCL), for 4000 Hz pure tones. All participants UCLs were within 2 standard deviation of Pascoe (1988) normative data for UCL.
5. Self-reported current use of medications that impact the pupillary dilation. Participants had to be free of regular use of anticholinergic drugs that are known to impact the dynamic range of pupil

dilation and influence sympathetic response. These medications include: Parkinson's medications, diphenhydramine (Benadryl), trihexyphenidyl (Artane), Benztropine mesylate (Cogentin), biperiden (Kineton), antipsychotics, clomipramine (Anafranil), chlorpromazine (Thorazine), Clozapine (Clozaril), fluphenazine (Prolixin), loxapine (Loxitane), olanzapine (Zyprexa), perphenazine (Trilafon), pimozide (Orap), quetiapine (Seroquel), thioridazine (Mellaril), thiothixene (Navane), and trifluoperazine (Stelazine).

Participants in the hearing loss group had to comply with the following additional exclusion criteria:

6. Any previous use of amplification devices including hearing aids for more than two months within the last three years.
7. If hearing threshold at 4000 Hz is better than 35 dB HL or greater than 60 dB HL. The lower limit (35 dB HL) is required as a minimal hearing loss to be included in the study in order to increase the chance of enough outer hair cell loss/damage to shows an impact on the perception of loudness. The maximum limit for hearing loss (60 dB HL) is specified in order to exclude any possible effect of dead regions on the hearing loss. Hearing loss due to dead regions implies inner hair cell loss, in addition to outer hair cell loss, which in turns impacts the perception of loudness in a different manner than the manner tested in this study. Response to sounds at a frequency of a dead region is more likely to occur in response to stimulating neighboring frequency hair cells.
8. Presence of a conductive component in the hearing loss defined by an air-bone gap of 15 dB HL or greater.

5.1.2 Sample size

G-power and PASS were used to conduct a power analysis to calculate the sample size needed for the proposed study.

Due to the lack of previous data on means and standard deviations of pupillary peak dilation in response to tasks that are similar to the task of the current study (detection in quiet to low level sounds), the effect size was calculated based on the pilot data, as well as previous data of mean peak dilation in response to speech comprehension tasks. It also was assumed that the dilation response to low-level sound detection would be smaller in magnitude than a dilation response to a speech comprehension task that involves a greater cognitive processing workload.

Assuming a difference in the mean pupil peak dilation between the two groups in this study being as low as 0.02, a standard deviation of 0.025, and an alpha of 0.05, an effect size (Cohen's d) was calculated to be 0.8, and a sample size of 52 participants in total (**26** participants in each group) was required to achieve 80% statistical power for SA 1, which is the primary aim in this study and therefore, appropriate for power analysis.

The standard deviation of the pupil dilation was estimated based on the highest and lowest mean peak dilation values in the pilot data (range of data divided by 4) to be 0.025.

The standard deviations reported in a speech comprehension listening effort study were 0.028 and 0.039, depending on the speech rate condition (Zhang et al, 2017). Therefore, the same power analysis conducted using the same parameter but changing the standard deviation to the one found in a slow speech comprehension task (0.028) revealed an effect size (Cohen's d) of 0.714, and a required sample of 64 participants in total (**32** participants in each group).

An effect size calculation and a power analysis for Aim 2 was difficult to perform given the lack of pervious data on how the pupillary response changes with increasing sound level. It was expected, however, that the statistical model to be used for analyzing Aim 2 will require a larger sample size. Therefore, the sample size calculated based on Aim 1 served as the minimal number of participants for this study.

Using a 5 x 2 within-between-factors repeated measures design with an alpha of 0.05, and an anticipated small to medium effect size (Cohen's f= 0.15), and the correlation between repeated measures is 0.5, a sample size of 56 participants in total (**28** participants in each group) was required to achieve an 80% statistical power for SA 2.

Power analyses of both Aims 1 and 2 revealed similar required sample sizes (total of 52, 56, and 64) for 80% statistical power to detect small differences between the two groups in the experiment. A sample size of **64 (32 in each group)** was chosen for this study. This sample size is conservative compared to sample sizes of previous auditory evoked pupillary response studies that ranged from 8 to 30 participants (Dybäck & Wallgren, 2016; Jackson & Sirois, 2009; Liao et al., 2016a; Liao et al., 2016b; Partala & Surakka, 2003; Schlemmer, Kulke, Kuchinke, & Van Der Meer, 2005).

5.1.3 Recruitment

All participants in the current study were recruited from the following:

1. The UPMC/University of Pittsburgh Research Participant Registry ("Pitt+Me™ Registry"). The Pitt+Me Registry is a voluntary database of people who are willing to consider participation in research studies and share their medical records for the purposes of research and medical care. Any UPMC patient as well as people in the community can be a part of the Pitt+Me Registry.
2. Pitt+Me social media platforms (Facebook and Twitter).
3. Recruitment flyers were posted on buildings of the University of Pittsburgh and the UPMC audiology clinics.

4. One of the co-investigators, who was also the director of UPMC audiology clinics and who had access to the medical records of her patients, generated a list of potential eligible participants for the study. These participants were invited to participate in the study by either direct face-to-face communication by their clinician, by mailing them a dear patient letter of recruitment, by emailing them a dear patient letter, or by initiating a phone call to invite them to the study. Initiating a phone call was performed only if the potential participant preferred and asked for this method.
5. A posted Craigslist's advertisement.

All posts, flyers, ads, letter, and registry invitations were IRB-approved. All participants were monetarily compensated for their participation.

5.2 DESIGN

The research question was investigated using a cross-sectional within subject design with multiple subjects. The study included two experiments.

5.2.1 Sound Detection Experiment (for Specific Aims 1 and 2)

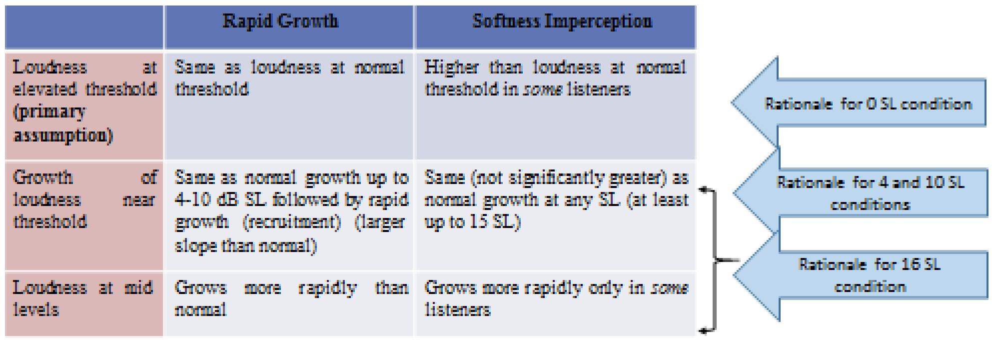
A mixed, 6x2, two-factor, repeated measure, within-between subject design was used to address SA 1 and 2. The main effects were the presentation sensation level (will be referred to as presentation level throughout this document) of the acoustic stimuli (silence= no tone, 0 SL, 4 SL, 10 SL, and 16 SL), and the participant's group (normal hearing vs. hearing loss). The presentation level was the within-subjects factor and the hearing condition was the between-subjects factor (Table 12). The dependent variable (DVs) for SA 1 and 2 was the pupillary dilation (the change in pupil diameter within the pupillary phasic response to sound). This experiment utilized a simple one-interval, yes/no simple detection task.

Table 12: experiment design for aims 1 and 2 (sound detection experiment)

Groups	Silence	-4 SL	0 SL	4 SL	10 SL	16 SL
Normal Hearing	DV	DV	DV	DV	DV	DV
Hearing loss	DV	DV	DV	DV	DV	DV

Soft and loud blocks that were previously separated in the pilot study were combined in the same blocks for the following reasons: 1) The 10 SL tone level did not elicit different pupillary response magnitude if presented with lower-level tones (in the soft block) or with the higher-level tones (in the loud block) as expected. This finding suggests that the pupillary response magnitude can reflect the loudness perception of the sound relative to the context of its presentation; 2) Loudness scaling studies commonly restrict tone level randomization, so no consecutive tones are more than 30 dB from each other. This suggests no concern for perceptual influence on the loudness of tones preceding another tone that is less than 30 dB apart. For this reason, it was found appropriate to include all sound levels ranging from -4 dB SL to +16 dB SL in the same block.

Each sensation level used in this experiment was chosen to best address one main difference between the conflicting loudness models that were tested in this study, as shown in Figure 22 below. The 0 SL condition examined the subtle difference between the two models. The 4 and 10 SL conditions were the low sensation level conditions that examined the loudness right above threshold and might distinguish one loudness growth from another. The 16 SL condition was chosen as the higher sensation level condition that was expected to manifest the abnormal growth of loudness in one model (Moore & Glasberg, 2004) and not the other (Florentine et al.). The 16 SL condition was chosen instead of 15 SL in order to keep the sound level difference between 4, 10, and 16 SL constant.



	Rapid Growth	Softness Imperception
Loudness at elevated threshold **(primary assumption)**	Same as loudness at normal threshold	Higher than loudness at normal threshold in *some* listeners
Growth of loudness near threshold	Same as normal growth up to 4-10 dB SL followed by rapid growth (recruitment) (larger slope than normal)	Same (not significantly greater) as normal growth at any SL (at least up to 15 SL)
Loudness at mid levels	Grows more rapidly than normal	Grows more rapidly only in *some* listeners

Figure 22: Main differences between loudness models for listeners with cochlear hearing loss dictated signal sensation levels chosen in the experiment.

The -4 SL condition was included to further verify that sounds in the 0 SL condition were presented at the participant's threshold. It would be expected that the percentage of "yes" responses at the -4 SL condition would be significantly smaller (less than 30%) than the percentage of "yes" responses at the 0 SL condition.

5.2.2 Subjective Rating Experiment (for Specific Aim 3)

A repeated measure between-subject design was used to address this aim. Again, the independent variable was the hearing condition (normal hearing versus hearing loss). The dependent variable was the difference in loudness rating between two tone levels within the range of 28 dB SL in each participant. The difference in rating of two tones was looked at (as opposed to the rating for each tone) in order to quantify the loudness dynamic range between each pair of tones. The pairs of tones examined for the differences in rating were: 4 and 16 SL, 4 and 20 SL, 4 and 24 SL, and between 4 and 28 SL. More detailed information on defining the dependent variable is included in the following sections.

Table 13: experimental design for Aim 3.

	-4 SL	0 SL	4 SL	8 SL	12 SL	16 SL	20 SL	24 SL	28 SL
NH	DV	DV	DV	DV	DV	DV	DV	DV	DV
HL	DV	DV	DV	DV	DV	DV	DV	DV	DV

Participants always completed experiment I (sound detection experiment) first. Completing experiment II (subjective rating experiment) first prior to experiment I might make participants unconsciously think about judging the loudness of the tones while presented in the sound detection experiment, which would add an additional undesirable source of cognitive activity within the intended simple sound detection task.

5.3 STIMULI

To examine the perception of loudness of low-level sounds, it was decided to use the simplest form of acoustic signals: pure tones. All other sounds can be broken down into combinations of sine waves. Additionally, pure tones were used as stimuli in the experiment to replicate the signal that has previously been used to test the SI model (Florentine et al., 2005; Moore, 2004).

A 4000 Hz tone was chosen for several reasons. First, the pilot data showed a more robust pupillary response to low-level sounds at 4000 Hz compared to 500 Hz. Second, testing at a high frequency sound increases the chance of recruiting listeners with hearing loss. Third, 4000 Hz was the frequency most tested among the high frequency tones in the previous studies that examined the concept of softness imperceptions (Florentine et al., 2005; Moore, 2004); and therefore, the current study findings will be appropriately comparable to the previous studies when using 4000 Hz tones.

A set of 11 audio files were generated and served as the acoustic stimuli for both experiments. They are variations of ten levels of 500 ms (including 20 ms rise/fall ramps shaped with a raised-cosine function) pure tones and one 500 ms of silence audio file.

All acoustic stimuli were generated using a digital audio editing program, Adobe Audition CC2017. Pure tone levels were generated in such a way that all tone levels were relative to the level of one of these tones, which was be the condition “0 SL” in this experiment. This tone was set at the participant’s hearing threshold level. This way, once the 0 SL tone was set at the participant’s threshold, each time the experiment was running, the presentation levels of all the other tones (example: conditions -4 SL, 4 SL, 10 SL, and 16 SL) maintained the targeted sensation level. Figure 21 shows an illustration of how the levels of the tones are relative to each other.

The 4000 Hz tone was digitally generated with mono channel, a sampling rate of 44.1 KHz, 16-bit depth, and with full amplitude (full voltage). The tone then was digitally attenuated by the amount of 30 dB and saved to serve as the 0SL condition. The same full amplitude tone was also attenuated by different amounts to serve as the -4, 4, 8, 10, 12, 16, 20, 24, and 28 SL conditions in the experiments. Table 14 shows the SPL level of tones used in experiment I, along with the digital attenuation needed to achieve the relative differences between them.

Table 14: Digital manipulations to a full amplitude 4000 Hz tone to achieve tones with intensity levels that are relative to each other.

Pure tone digital amplitude	***SLM reading (dB SPL)***	***Intensity difference from the 0 SL condition tone (dB SPL)***	***Use of tone in the experiment***
Full amplitude	86.3-86.5	NA	Not used in experiment
-33.9 dB	60.9-60.9	-4	Served as -4 SL condition

Table 14 (continued).

-30 dB	64.9	0	Depending on each participant's threshold, this tone was attenuated to be set his/her threshold Served as 0 SL condition
-26.2 dB	68.8	4	Served as 4 SL condition
-20.4 dB	74.9	10	Served as 10 SL condition
-14.6 dB	81	16	Served as 16 SL condition

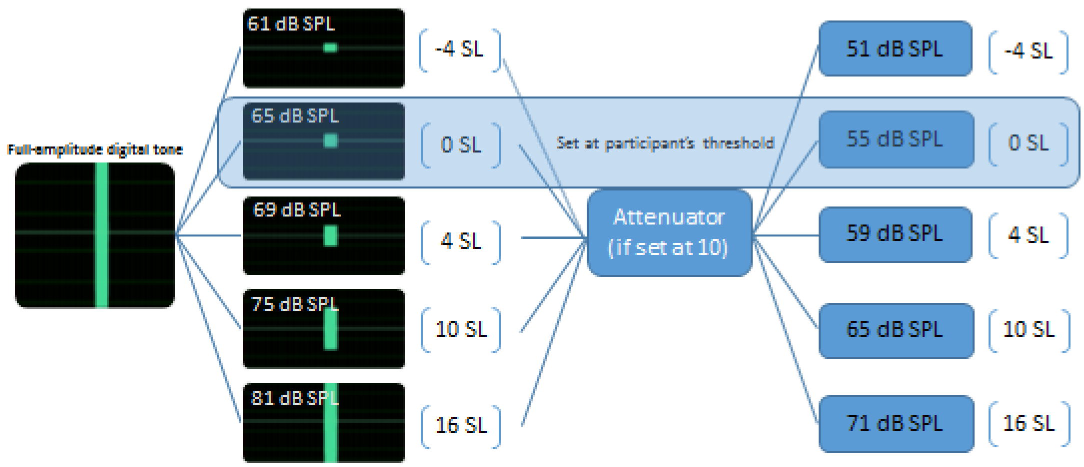


Figure 23: Generating acoustic stimuli with different sensation levels (left panel). Right panel is an example of tones set at a participant's threshold of 55 dB SPL.

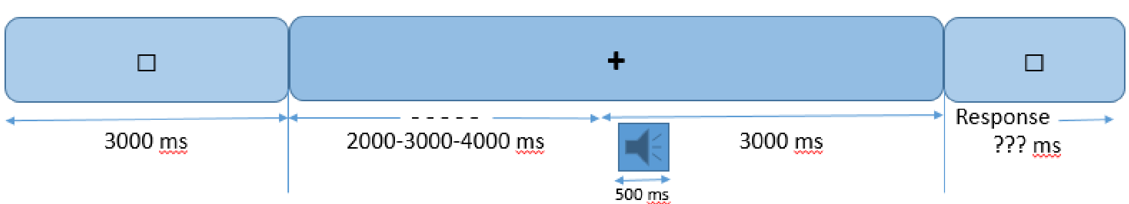


Figure 24: Design of the trials within the sound detection block.

5.4 PROCEDURES

All procedures including completing the consent form, screening, and experimental procedures, were conducted in the Auditory Processing Lab of the Department of Communication Sciences and Disorders at the University of Pittsburgh. The entire research session flowchart is illustrated in Figure 25. All screening and experimental tasks were completed in one 2-hour session. Participants were allowed to ask for breaks between blocks and between experiments.

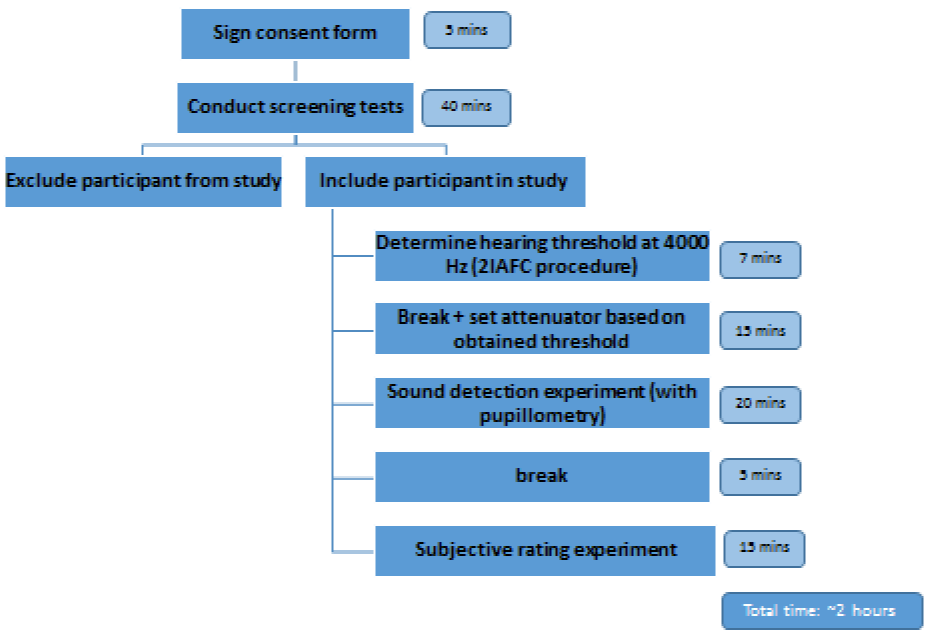


Figure 25: Flow chart of research session.

Consenting: An IRB-approved consent form was provided to each participant to read and sign prior to the beginning of any procedure. Participants were given ample time and encouraged to ask questions or ask for explanation or clarification regarding the form or the study details.

5.4.1 Screening procedures

Participants were asked a series of case history questions to ensure eligibility for the study (Appendix A). The questions included demographic, medical, and audiologic questions. An otoscopic examination was performed to ensure that the participant's ear canals were free from occluding wax. A pure tone audiometric hearing test following the ASHA guidelines (Association, 2005) was conducted at octave frequencies from 250 Hz-8000 Hz in both ears tested separately (ANSI, 2004). The MADSEN Astera[2] audiometer from Otometrics was used with ER-3 insert earphones. Uncomfortable loudness level of 4000 Hz tone for the right ear was obtained using the same audiometer. In each run, a 4000 Hz pulsed pure tone was first presented at 60 dB HL for approximately one-second duration. Tones were presented in an ascending manner in 5 dB steps until the participant indicated that his uncomfortable loudness level had been reached. The instructions and method were consisted with Cox et al. (1997), (see instruction in Appendix B). Prior to the test, participants were provided with the Loudness Chart of the Cox Contour Test (Cox, et al., 1997) shown in Appendix C. Participants who demonstrated a UCL that was 2 standard deviations outside the normal range according to normative data by Pascoe (1988) (Appendix D), were excluded from the study. Word recognition testing was conducted for the right ear only using the same audiometer. Speech was presented at 40 dB SL above pure tone average (PTA), or 20 dB SL above the 2000 Hz audiometric threshold, whichever was greater. Participants were presented with an NU.6 25-word list and had to score within the 95% confidence interval of the normative data to be included in the study.

A vision screening was conducted using a Snellen chart (Bailey & Lovie, 1980). All participants were required to have a visual acuity of 20/20 or better when tested under the binocular condition without eye glasses with a viewing distance of 31.5 inches = 80 cm, which was equal to that of the viewing of the computer screen in the experiment. Presence of an adequate light reflex pupillary reaction and near vision accommodation reaction was assessed by the experimenter. Before recording those reflexes, the experimenter flashed an otoscope beam light on participants' eyes and observed the pupil contractions. Then, the experimenter asked participants to follow the tip of a pen positioned 1 meter away from the eye moving slowly toward then slowly away from the eye. Once both reflexes were determined to be adequately present, participants were seated in the experimental setting (the booth) and their pupil diameters were recorded in partial darkness while they gazed at a fixation cross on the screen. The participants were instructed to do nothing but look at the cross in the screen while the experimenter adjusted the eye camera. At least 15 seconds later, the experimenter turned the sound booth light on. Then, in medium lighting, the experimenter repeated the near vision accommodation eliciting task, but with the camera recording the pupil this time. The recorded pupil diameter in partial darkness and the pupil diameter range of constriction during the light reflex (amplitude of light reflex) and accommodation reflex (near vision reflex) were examined later and compared with normative ranges. These data were also used as descriptive data for the study sample.

After completing the screening procedures and determining eligibility for the study, the enrolled participants completed the following procedures:

5.4.2 Descriptive procedures

Absolute hearing threshold for 500 ms-4000 Hz pure tone in the right ear was measured using an adaptive two-intervals-alternative forced choice procedure (2I-AFC). This hearing threshold was obtained using an ER-2 insert earphone. All experimental stimuli were presented through ER-2 insert earphones. Each trial contained two observation intervals that were displayed in a computer screen in front of the participant as two boxes, marked as 1 and 2. The intervals were separated by 500 ms and were marked by consecutive yellow lights as shown in Figure 26 (left panel). The stimulus was presented with equal a priori probability in either the first or the second interval. Participants were required to click on the interval (box) which contained the tone using a mouse. A 250-ms green light indicated the correct answer 100 ms after the participant's response as shown in Figure 26 (right panel). The next trial began after a 500-ms delay. A single threshold measurement was based on three interleaved adaptive tracks. In each trial, one track was selected at random. An illustration of the 2I-AF trials is shown in Figure 25. For each track, the signal level started at 50 dB SPL. It decreased after two consecutive correct responses, and increased following one incorrect response. This procedure converged on the signal level yielding 79.4% correct responses (Levitt, 1971). The step size was initially 5 dB and decreased to 2 dB after the second reversal. Reversals occurred when successive signal levels changed direction from decreasing to increasing or vice versa. Each track ended after eight reversals. The threshold for one track was calculated as the average of the signal levels at the fourth and fifth reversals. One threshold measurement was taken as the average threshold across three tracks. The thresholds obtained from each track had to meet two criteria to be accepted, otherwise, the track was re-administered. A track could be re-administered a maximum of two times, otherwise, the participant was excluded from the study due to the high chance of inability to follow test instructions. Each track threshold had to be within 15 dB from the participant's own 4000 Hz audiometric hearing threshold. The difference between the highest and the lowest track thresholds could not exceed 10 dB difference. This adaptive hearing threshold procedure was previously written in Python language and run through the Spyder platform from the same computer that ran the experimental stimuli. Appendix B shows the verbal instructions provided to participants for the 2I_AFC hearing threshold procedure.

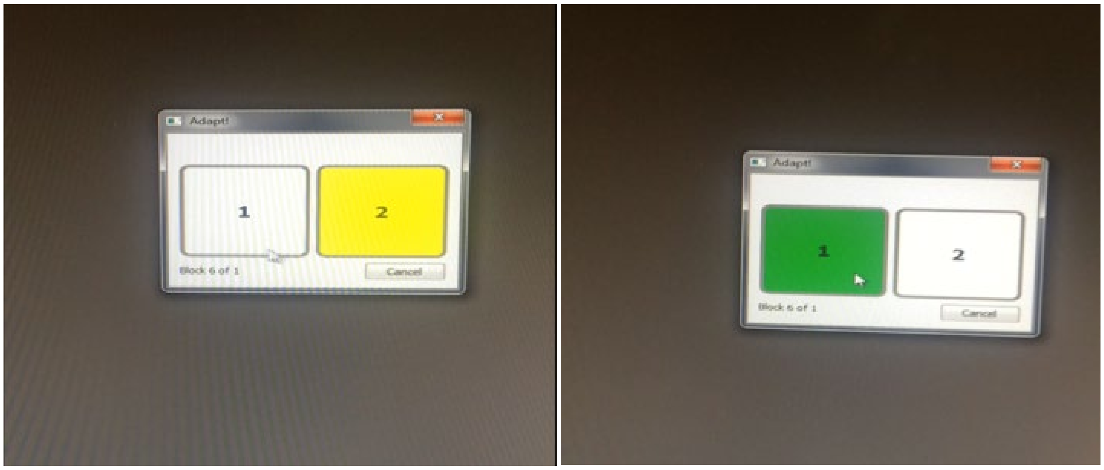


Figure 26: Illustration of two intervals in screen in front of participant during 2I-AFC procedure.

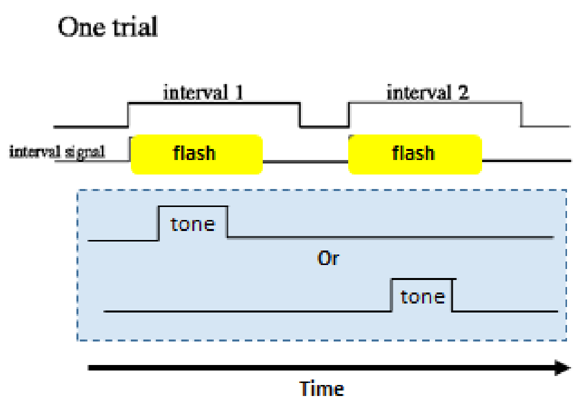


Figure 27: Design of trial in 2I-AFC procedure.

5.4.3 Experimental procedures

All participants completed two experimental tasks: the sound detection experiment (with pupillometry), and the subjective rating experiment (without pupillometry).

5.4.3.1 Experiment I: sound detection experiment

All experimental procedures took place in a 1.5x2.2m double-walled sound-treated booth that meets specifications for maximum permissible ambient noise levels (ANSI, 2003). The booth was dimly lit with most luminance coming from the booth window and the stimuli screen. Luminance of the screen and the testing sound booth were kept constant throughout the entirety of the sound detection experiment. The experiment setup is shown in Figure 28. Participants were seated at a table in front of the computer monitor (size of the display: 13.5x10.5", display resolution: 1024×768) at a distance of approximately 32.2 inches (82 cm) (angle eye = 2.73 degrees) with a four-key response keypad in front of them. The experiment instructions and the fixation images (for the purpose of pupil dilation measurement) were displayed on this monitor

For experimental data collection, a computer running the experimental control software, SuperLab (Cedrus, Phoenix, Arizona) controlled the presentation of acoustic and visual stimuli, the timing operations, and the keypad response data acquisition. Acoustic stimuli were played out via an external Cakewalk soundcard attached to the experimental computer, fed to an attenuator (that is set at a different level for each participant based on their 4000 Hz threshold), and provided to the participant's right ear via ER-2 insert earphone. During the experiment, participants were fit with bilateral insert earphones, but stimuli were presented to the right ear only. A one-minute duration of the 4000 Hz tone that was used as the 0SL condition in the experiment (without applying any attenuation), was the calibration tone. A Larson-Davis 824 sound level meter was used to measure intensity of the calibration tone from the right ER-2 insert earphones at 65 dB SPL when the attenuator was set at zero.

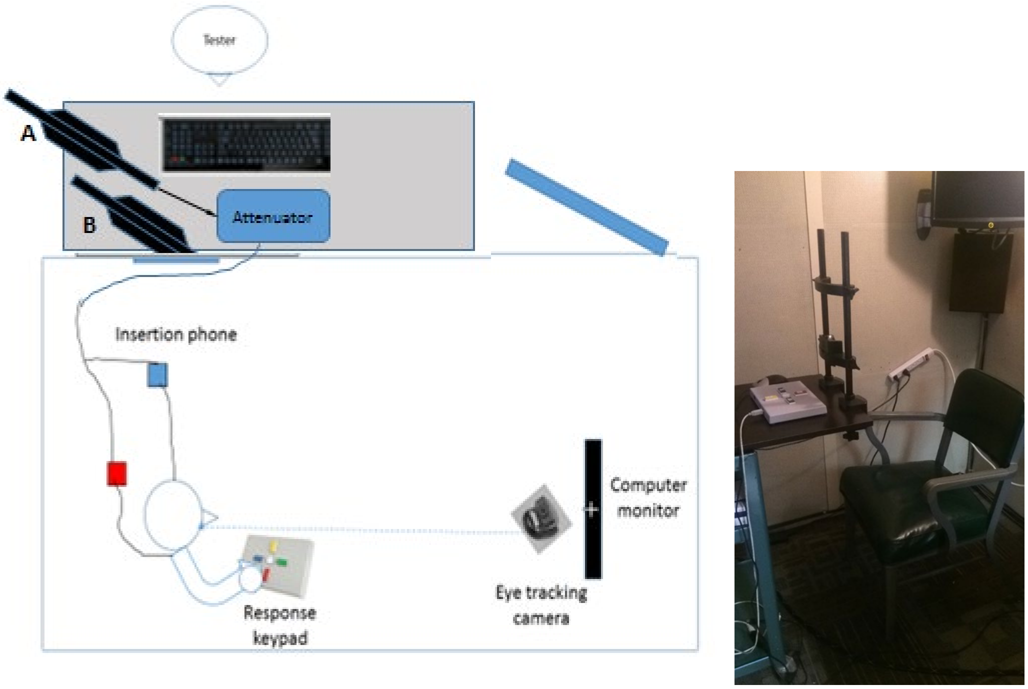


Figure 28: Experiment setup. Computer A is running the experimental stimuli. Computer B is controlling pupillometry.

Practice trials

Prior to the experiment, participants performed a sequence of practice trials to ensure they were familiarized with the response keypad and the stimuli, and that they understood the task and directions. A list of practice trials along with the specific instructions that were displayed on the screen is as follow:

1. Practice keypad: In order to familiarize the participants with the response keypad, they were presented with a picture of the keypad on the screen (Figure 29) and were instructed as follow: "*In order to get familiar with the experiment task, you will go through several practice trials. First, you need to feel comfortable with the response keypad. You will be using the response keypad without looking at it throughout the experiment. So please take a minute to memorize the layout of the keypad*".

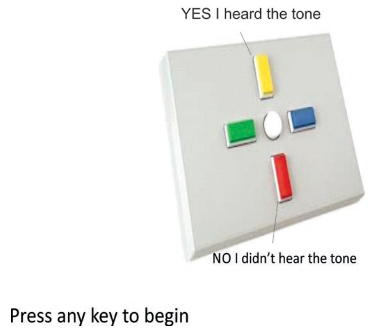


Figure 29: The response keypad used in the experiment.

Next, they were instructed as follows: "*Please follow the commands by pressing the appropriate key on the keypad WITHOUT looking at it. You will receive a feedback for each item telling you if you are correct or incorrect*".

Participants were provided with a feedback indicating correct and incorrect responses by presenting happy or sad faces on the screen.

2. <u>Practice sound detection task</u>: In order to familiarize the participants with the sound detection task in the study, they were instructed as follows:

"*We want you to be familiar with the task in this experiment. You will see a square on the screen. When the square changes to a cross, a tone might be presented during the cross period. Once the square comes back on the screen, press "Yes" if you heard the tone, press "No" if you did not hear the tone. Please always look at the image (square or cross) on the screen during the task. Please keep your eyes open during running the experiment and try not to blink especially when you see the cross. Press any key to begin the practice*".

Five experimental stimuli were used in the practice trials (silence, 10 SL, 12 SL, 16 SL). Practice stimuli were chosen to be high SL stimuli to ensure that participants can easily detect tones in the practice trials and get familiar with the task. Participants were provided with feedback by presenting a happy or sad face on the screen based on their correct and incorrect response.

Along with the written instructions on the screen, participants were also provided with verbal instructions before starting the practice trials and again before starting the main experiment (see appendix B for the script of instructions). If the participant showed accurate responses during the practice trials, he/she proceeded to the experimental trails. Otherwise, the participant was given more verbal instructions and repeated the practice trials. Participants had to master the practice procedure by the third time, otherwise, the participant was excluded from the study.

Experimental trials

Participants completed a total of 72 trials (5 SL levels × 12 + 12 silence trials) divided into two blocks. Each block took about 7 to 10 minutes to be completed. Participants were encouraged to take a break in between blocks and were allowed to take as long as they wanted. The order of the stimuli within each block was random.

At the beginning of the block, participants were reminded about the experiment task with written instructions on the screen as follows: "*This is the beginning of the main experiment. As you did in the practice, when the square changes to a cross, you might hear a tone, but this time, the tone could be very soft sometimes (the softest sound you can hear). So please pay attention. When you see the square back, press "Yes" if you heard the tone. Press "No" if you did not hear the tone. Your response will not be registered until the square is back. So please wait for the square to tell us. (+) listen for the tone (□) press "Yes" or "No" and wait. Press any key to start*".

Participants were required to look at a fixation square at the center of the screen. Three seconds later, a fixation cross appeared. Participants were instructed to listen carefully during the presentation of the cross as a tone might be presented. Tone onset was preceded by a randomly selected 2, 3, or 4 sec pre-stimulus interval. Random pre-stimulus interval was inserted to avoid stimulus anticipation and response habituation. Three seconds post tone onset, a fixation square re-appeared signaling the participant to respond if he or she heard the tone or not by pressing the yes/no button that is placed on the table in front of participant. A fixed 3 sec post tone interval (without initiation of a response) provided sufficient time for the pupillary response to occur without overlapping with the following trial, or with the motor response effect on pupil dilation. The initial 3000 ms square fixation period of each trial was included for the following reasons:

1) The initial square fixation period separated it from the cross fixation period that participants were required to attend and detect the signal. Reducing this period minimized the fatigue that might be suffered by participants if they were required to continuously attend to signal detection.

2) Recorded pupillary data during the initial square fixation period contained any motor impact from button activation on pupillary dilation and therefore, the motor-induced response was separated from the task-induced response. It also separated it from the cross fixation period from which the pupil diameter baseline was obtained (200 ms pre-tone onset).

3) In order to reduce the amount of blinking by participants during pupillometry recording, they were instructed to keep their eyes open during the sound detection experiment and to refrain from blinking within each block. Within these trials, if participants must blink, they were instructed to do so during the square fixation period, if possible. This way, any distortion of the data due to blinking will be minimal during the critical pupillometry recordings.

Pupillometry

Throughout the sound detection experiment, participants' pupil diameters were monitored using an ASL Eye-Trac 6 system (Applied Science Laboratories, Bedford, MA), which consists of a video camera and an infrared light source pointed at the participant's right eye. The ASL Eye-Trac 6 system was used with a head and chin rest that has been shown to be very effective in minimizing head movements, keeping the participant in focal range of the video camera, and ensuring a consistent distance from the screen (Raney, Campbell, & Bovee, 2014). A chin-rest was mounted at the end of the table, and the height was kept at the same position between participants as much as possible and was only minimally adjusted when participant's chin could not comfortably be positioned at that height. The ASL Eye-Trac 6 system employs a Pupil-Centre Corneal Reflection (PCCR) method (Mason, 1969) to calculate pupil size and to track the eye diameter and eye gaze location at 60 Hz (i.e., every 16.7 ms). The spatial resolution of the pupillometer was 0.1 mm. As the ASL program was run by a separate computer, a PCI-DIO 24 board manufactured by Measurement Computing was used to connect it to the experimental computer which ran SuperLab software. The pupil tracking by the ASL system was able to synchronize with the stimuli presentation, having markers of the onset of each fixation sign and tone in the pupil dilation traces.

The luminance of the visual field was controlled and kept constant to avoid the floor and ceiling of the range of pupil size, which is affected relatively more strongly by the light reflex than by cognitive load (Beatty, 1982a; Beatty & Lucero-Wagoner, 2000a; Zekveld & Kramer, 2014). The monitor had a grey background and the fixation images were thick, black square or plus signs with equal dimensions of 1.5x1.5 inches. Using grey screen with black fixation signs minimized the change of luminance with changing visual stimuli (Lemercier et al., 2014).

Prior to the experiment, the eye tracking camera was adjusted to provide accurate pupil diameter reading for a 4 mm diameter of a model eye. Once the camera's zoom and focus functions were adjusted for calibration to obtain a reading from the model eye, they were not adjusted afterward between participants. In case a participant's eye image was not clear, the camera was moved slightly forward or backwards to obtain a clear image of the participant's pupil.

5.4.3.2 Experiment II: subjective rating experiment

In the second experiment, participants provided subjective ratings related to the loudness of the signals presented. It would have been optimal to obtain the loudness subjective ratings of stimuli simultaneously with the pupil data; however, pupillary function is very sensitive to cognitive functions that would occur if the participant were asked to not only detect the sound and report its presence, but also to judge its loudness and make another decision. Therefore, subjective ratings of loudness were obtained in a separate experiment.

The setting for this experiment was the exact setting of the sound detection experiment except that participants did not have to use the chin rest, the luminance could be adjusted higher, and responses were not given through the keypad. Sound levels ranging from -4 SL up to +28 SL (including the tones used in the sound detection experiment), ascending in 4 dB steps, were used as the stimuli in this experiment. The 4 SL condition sound was just above the participant's threshold, which allowed for the comparison between the two models of loudness at threshold examined in this study. The pure tone and the procedures for generating the audio files for this experiment were the same as those used to generate audio files for the sound detection experiment.

Based on the review of psychoacoustic subjective measures of perceived loudness in section 2.5, these measures can be categorized by two methods: loudness matching and loudness scaling. Loudness scaling is when a listener relates the physical magnitude of sounds to their subjective loudness. Loudness scaling techniques have been primarily used to obtain loudness growth information. The most common methods that have been used to derive loudness scaling are the following: magnitude estimation (ME), magnitude production (MP), cross modality matching (CMM), and categorical scaling (CS).

Loudness scaling in general is an easier psychoacoustic task compared to loudness matching which requires more training. Specifically, magnitude estimation of loudness is the most used technique to measure the growth of loudness within the dynamic range (Hellman & Zwislocki, 1963). It has been used with a variety of acoustic type and duration stimuli (Epstein & Florentine, 2006a; Poulton, 1989; Stevens & Hall, 1966). Other psychoacoustic measures have been examined using magnitude estimation, such as temporal integration. Magnitude estimation has been shown to provide loudness data that are consistent with other data obtained using MP and CMM from both listeners with normal and impaired hearing (Hellman, 1999; Hellman & Meiselman, 1988, 1993). In a standard ME task, listeners are asked to provide a numerical value that matches the loudness of a test tone.

A relative ME measure was utilized in the current study. In this study, participants were asked to rate the loudness of a test tone in relation to the loudness of a standard tone whose loudness was called 100. The standard tone was a tone that was set at 20 dB SL. The test tone was randomly selected from 10 tones with levels ranged from -4 SL up to 28 SL in 4 dB steps.

The relative ME technique was chosen to address SA 3 because it allows for anchoring the listeners perception at a certain point, but at the same time not restricting their responses. This anchor makes it possible to make individual loudness function slopes more meaningful, and allows for direct collection of outcome data. Although it has been argued that the use of a standard produces a bias (Hellman & Zwislocki, 1963), loudness functions collected with and without the use of a standard reference have shown to be in sufficient agreement with each other (Hellman & Zwislocki, 1963). Because the loudness functions obtained in the present study were compared between subjects, the bias (if existed) became less relevant.

In the subjective rating experiment, each participant completed a total of 90 trials (9 tone levels x 10 repetitions) where tones were randomized for each trial. Before starting the experiment, participants were provided with verbal instruction (see appendix B) and went through 10 practice trials. Experimenter made sure participants recoded their ratings accurately in the response sheet and that they didn't have any questions before starting the experimental trials. Data from practice trials were not included in the analysis. In each trial, a pair of 500 ms, 4000 Hz tones presented consecutively (with 500 ms inter-stimulus interval) to the participant's right ear (see Figure 30). The first tone was always the standard tone of 20 dB SL. The second tone was the test tone that changed every trial. The test tone was randomly selected every time from a list of tones of different levels (-4, 0, 4, 8, 12, 16, 20, 24, and 28 dB SL), so each tone condition was presented 10 times. On the screen, participants fixated on numbers 1 and 2 presented simultaneously with the onset of the first and second tone, respectively. After each trial, participants could take as long as needed to record their loudness rating for the second tone and once they are ready for the next pair of tones, they were instructed to press any button in the keypad.

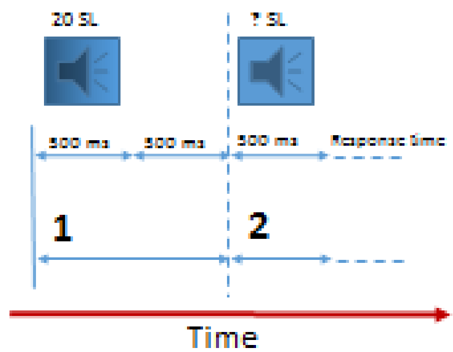


Figure 30: Design of trials in subjective rating experiment.

5.5 DATA REDUCTION / PROCESSING / STATISTICAL ANALYSIS

5.5.1 Pre-experiment data

Pre-experiment baseline pupil diameter

It could be argued that a difference between the groups in pupil dilation response in the study is inflated by the difference in their baseline pupil diameter. This difference in pupil size could be related to the difference in age between the study groups or the difference in an unknown impact of hearing loss on baseline pupil size. It is generally accepted that pupil size tends to decrease with increasing age (Piquado, Isaacowitz, & Wingfield, 2010; Van Gerven, Paas, Van Merriënboer, & Schmidt, 2004) and therefore have a more restricted range of dilation. Findings of such studies that examined age related difference in pupil reactivity were drawn from older adults above the age of 67 (Piquado et al., 2010; Van Gerven et al., 2004). Given that the current study's cut-off age was 69, this age related reduction in pupil reactivity becomes less concerning. Although the study groups included participants of similar age range, the HL group was significantly older, $t(48.196) = -3.957$, $p = <.001$. In order to determine if the baseline pupil diameter was different in the different hearing groups, the baseline pupil diameter data obtained from participants, prior to starting any experimental tasks, was subjected to an independent-samples t-test. There were three outlier values in the NH group data, as assessed by inspection of a boxplot. It was decided to not exclude the outliers, but instead adjust them to values that are less extreme (i.e., the next smallest value). Baseline diameter was normally distributed in both groups, as assessed by Shapiro-Wilk's test ($p > .05$). There was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .108$). The normal hearing group had larger baseline pupil diameter ($M = 5.464$, $SD = .842$) than the hearing loss group ($M = 5.043$, $SD = 1.192$). However, there was no statistically significant difference in baseline pupil diameter between the hearing groups, $M = 0.421$, $SE = 0.330$, $t(36) = 1.274$, $p = .211$.

The pre-experiment pupil size baseline was obtained from the pupil data recorded prior to the onset of pupil light reflex. As discussed earlier, this recording was obtained at least 5 minutes after completing the audiometric screening tasks and before starting any experimental tasks. Participants were seated in an almost dark booth (only luminance coming from the booth window and the grey-backgrounded computer screen) and were instructed to fixate on a black cross on the screen and stare at it while their pupil was being recorded. The average pupil diameter for the duration of 500 ms, selected -700 ms to -200 ms relative to onset of constriction. Onset of pupil constriction was determined for each participant individually by visual inspection of the response curves. The pre-experiment baseline data were calculated from 21 participants (out of 29) in the NH group and 17 participants (out of 29) in the HL group. The age distribution among these specific participants (Figure 31) were representative of the age distribution of the whole sample size (see Figure 42 for comparison). Due to the sudden change in the surrounding luminance, the camera's recording function was often interrupted. As a result, the data recorded from the rest of the participants in this task included a lot of missing and artifactual data either at the baseline time or at the dip of the reflex and hence, baseline diameter and/or reflex amplitude data were not able to be calculated for those participants.

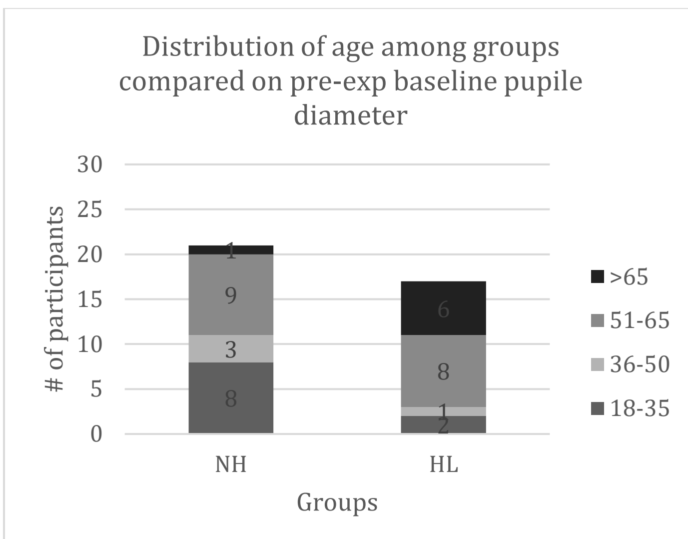


Figure 31: Age distribution among participants compared on pre-experiment pupil baseline diameter.

Pre-experiment pupillary light reflex and accommodation reflex

The amplitude of light reflex is the difference between the average baseline pupil diameter calculated above, and the average of 200 ms post maximum constriction (dip in the graph). The onset of maximum constriction was determined for each participant individually by visual inspection of the response curve.

The range of constriction for each participant was considered as an indication of the range of the pupil's reactivity. This index was previously used as an indicator for pupil dilation reactivity (Piquado et al., 2010).

The amplitude of pupillary light reflex data were calculated from 13 (out of 29) participants in the NH group and 7 (out of 29) participants in the HL group and was subjected to an independent-samples t-test to determine if there were differences in the possible range of pupil constriction between the two groups. Again, the distribution of age among the groups (Figure 32) was still similar to that of the whole study sample (see Figure 42 for comparison). There were no outliers in the pupil light reflex amplitude data, as assessed by inspection of a boxplot. The amplitude of light reflex data in both groups were normally distributed, as assessed by Shapiro-Wilk's test ($p > .05$). There was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = .280$). There was no statistically significant difference in the amplitude of light reflex between the NH and the HL group, $M = -0.267$, $SE = 0.285$, $t(18) = -0.937$, $p = .361$.

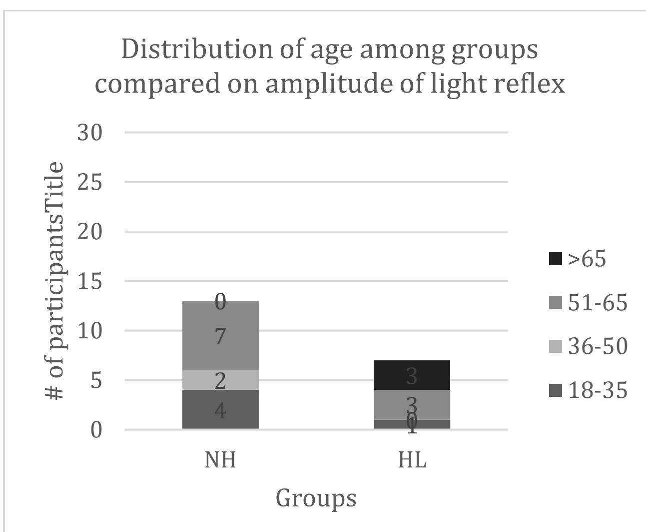


Figure 32: Age distribution among participants compared on amplitude of light reflex.

It should be noted, however, that age in these participants was negatively correlated with pre-experiment baseline pupil diameter, $r(36)=-0.38$, $p=.016$. Age statistically explained 14% of the variability in pre-experiment baseline pupil diameter. When the correlation between age and the magnitude of the pupillary light reflex were examined, they were not significantly correlated, $r(18)=-=.41$, $p=.075$. This indicates that older participants in this sample tended to have smaller pupil diameter compared to younger participants, but their pupil range of reactivity (defined as their reactivity to light) did not differ from younger participants in this sample (Figure 33). Age was significantly negatively correlated with near vision (accommodation) reflex amplitude, $r(9)=-.77$, $p=.006$.

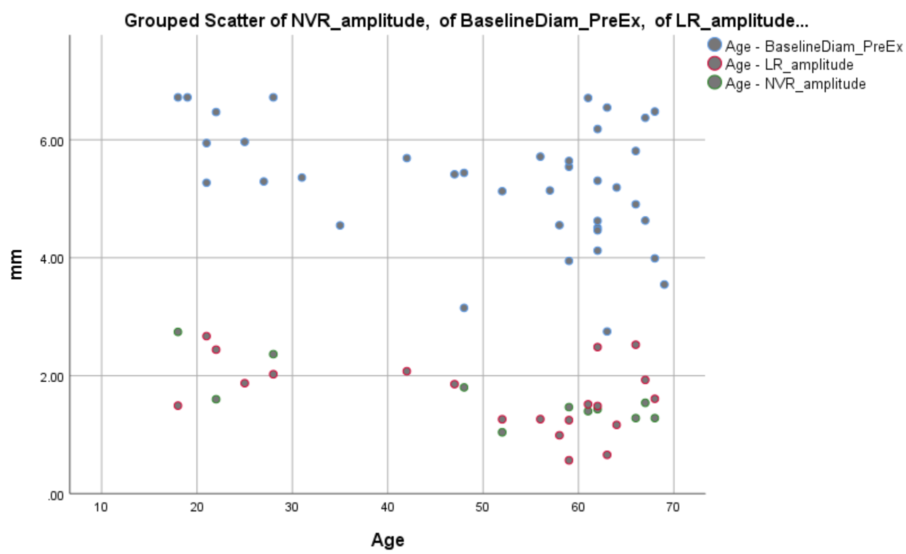


Figure 33: Scatterplot of age and pre-experiment baseline pupil diameter (blue dots), age and amplitude of light reflex (red dots), and age and amplitude of accommodation (near vision) reflex (green dots).

The amplitude of accommodation reflex is the difference between the average 20 smallest pupil diameter values and the 20 largest diameter values recorded during the task. Again this calculated range for each participant was considered as an indication of the range of the pupil's reactivity. Pupil diameter recordings from the NV reflex procedure were extremely corrupted and therefore, NV reflex amplitude was able to be calculated from only 7 participants in the NH group and 4 participants in the HL group. Amplitude of NV reflex was not significantly different between these two small sample sized groups, *M* = .323, *SE* = 0.314, $t(9) = 1.031$, $p = .329$.

2I-AFC hearing thresholds

As expected, the SPL hearing thresholds for 4K Hz tones obtained from participants were highly correlated with their audiometric threshold in dB HL for that frequency, $r(56) = 0.960$, $p < .001$ (=.000) (Figure 34).

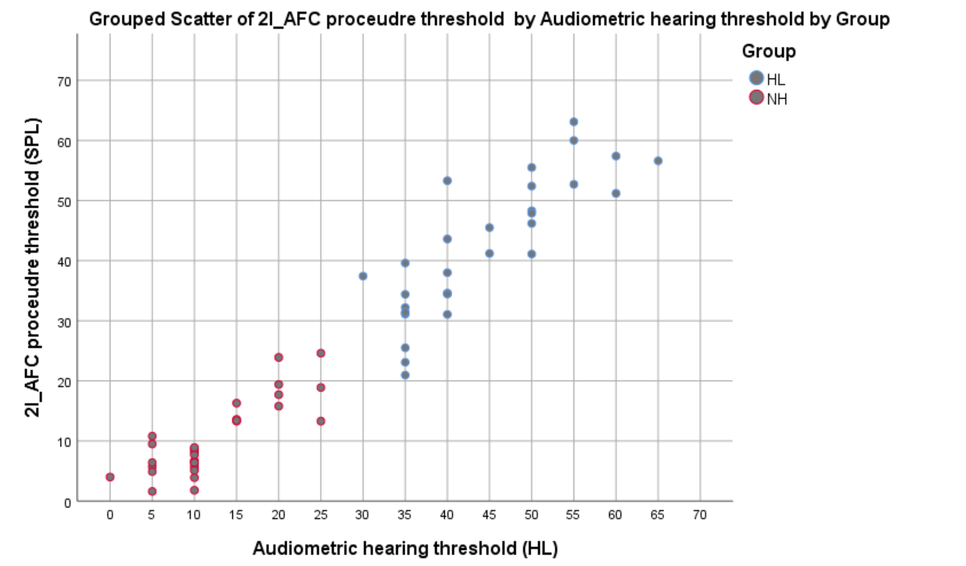


Figure 34: Correlation of audiometric hearing threshold (in dB HL) with 2I_AFC hearing threshold (in dB SPL).

5.5.2 Experiment I: sound detection experiment

5.5.2.1 Behavioral data

To examine the extent of accurate presentation of 0SL condition tones at participant's threshold, the percentage of "yes" responses across all tone conditions were examined. In order to ensure that tones presented at 0 SL condition were accurately presented at threshold to the nearest approximation, pupillary data obtained from participants who responded "yes" 100% of the times (or 0% of the times) across all those near threshold conditions (-4, 0, and 4 SL) were excluded from the analysis. This process resulted in the exclusion of 8 participants from the NH group and 2 participants from the HL group. As long as the participant showed an increase in percent from 0 SL to 4 SL condition or from -4 SL to 0 SL condition, even if response at 0SL was 100% or 0%, this participant was still included in the analysis, since this change in response indicated an approximation to a response that is near enough to threshold. Figure 36 shows the averaged participants' "yes" responses across tone conditions. Although the average percent of "yes" response at 0 SL tone condition in the study groups (NH: M=45, SD= 33.780, HL: M= 32.034, SD= 34.735) did not match the percent of detection proposed by the threshold search procedure (i.e., 2I_AFC) of 80% (Levitt, 1971), this was not surprising given the different sound detection procedure (a non-AFC, non-adaptive) utilized in the experiment that included a larger inter-stimulus interval, which predicts a decreased performance in detection.

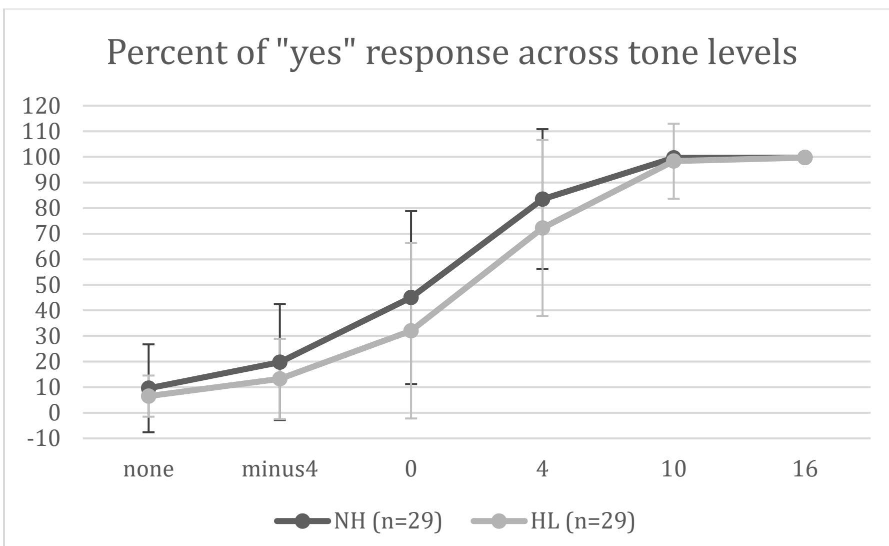


Figure 35: Mean percent of "yes" responses across tone levels for both study groups. Bars indicate 1 standard deviation.

Participants' sound detection sensitivity across tone levels was examined. Each participant's sensitivity index (d prime, d') was calculated at each tone level. D prime is a measure of sensitivity, whereas proportion correct is affected by both sensitivity and bias. D prime was calculated from each participant's false alarm (ratio of "yes" responses when no tone was presented) and hit rates (ratio of "yes" responses when tone was presented) across tone levels. Figure 36 shows the average d' across tone levels for each group. In signal detection theory, larger absolute values of d' mean that a participant is more sensitive to the difference between the signal present and signal absent distributions, while values near zero indicate no discrimination, i.e., a chance performance. Although participants in the study did not show a typical percent of "yes" response at threshold, they, showed the d' (i.e., around 1) that is consistent with what is typically used to determine threshold (Micheyl, Schrater, & Oxenham, 2013), (NH: M=1.2, SD= 1.162, HL: M=.888, SD= 1.209). By observing similar detectability index (d') around threshold between the two groups, *M* = .334, $t(56)= 1.073$, $p= .229$, different pupil responses to sound detection were more likely to be unconfounded by differences in task performance.

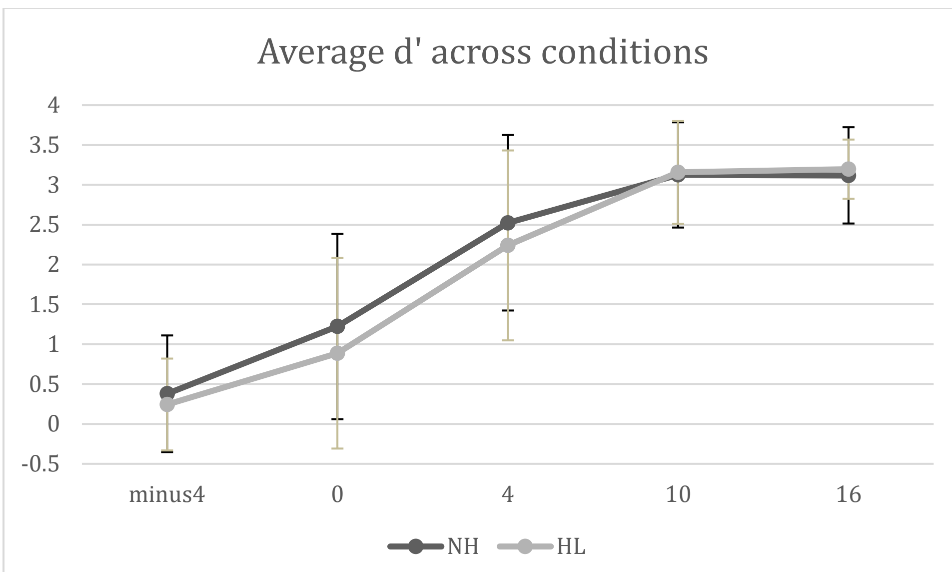


Figure 36: Mean d' scores across tone levels for both study groups. Bars indicate 1 standard deviation.

5.5.2.2 Pupillometry data

For each trial, the collected pupil data were divided into four segments. The initial square segment was the pupil data collected at the beginning (3000 ms) of each trial. The middle pre-tone onset segment was the pupil data collected with the cross onset until the tone onset (data from the last 200 ms of this segment was averaged and served as the baseline diameter to be subtracted from the pupil response after). This segment was 2000, 3000, or 4000 ms long. The post tone onset segment was the pupil data collected with the tone onset until the second square onset (the beginning of the response period). This segment was 3000 ms long. Button response segment was the pupil data collected with the onset of the second square until the participant's initiation of the response (button press). R-codes (a programming language) that were specifically written for this study data were used to extract the pupil data recorded during the experimental trials and to separate these three segments. Pre-tone and post-tone segments comprised the trial section that was subjected to pre-analysis processing described below. The post tone-onset segment comprised the data that was subjected to statistical analysis to answer the research question (SA 1 and 2). Figure 37 shows the distinction between the four segments, segments included in data processing, and segments included in data statistical analysis.

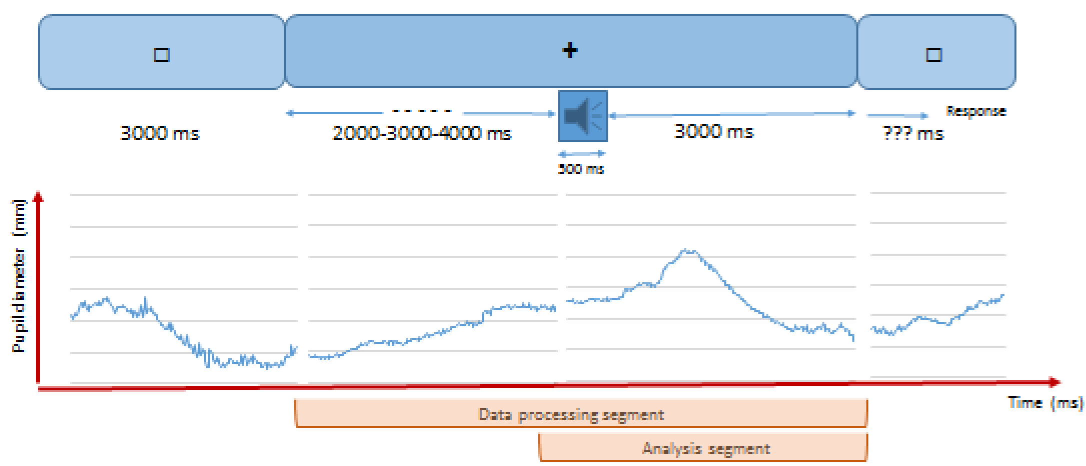


Figure 37: Hypothetical recorded pupil diameter data in each trial as it was segmented for processing and analysis.

The output of the eye tracker data files included the major axis length (width) of the pupil and the minor axis length (height) at each time sample (i.e., every 16.67 ms, 60Hz sampling rate). The minor axis becomes small as the eye rotates away from the camera, however, the major axis does not, which indicates that the width is a more accurate pupil size measurement during eye movements (Kuchinsky et al., 2013).

The pupil data processing was performed using CHAP software (Hershman, Henik, & Cohen, 2018, manuscript submitted for publication). CHAP is a soon-to be open source software written in MATLAB to process and analyze pupillometry data. In each trial, outliers were identified as values 2.5 standard deviation away from the trial mean. Outlier values were converted to missing data. Blinks were identified using a novel noise-based algorithm (Hershman, Henik, & Cohen, 2018), which accuracy is in agreement with other common blink detection methods e.g., (Mathôt, 2013). Trials comprised of over 50% missing data (occur during blinks, head movements, difficulty of the eye-tracking system to register the pupil measurement) were rejected from the analysis (Kuchinsky et al., 2013; Siegle, Steinhauer, Stenger, Konecky, & Carter, 2003b; Silk et al., 2009). Missing data in accepted trials were replaced using linear interpolation of values measured just before and after each identified blink (Siegle, Steinhauer, Carter, Ramel, & Thase, 2003a; Silk et al., 2009; Steinhauer et al., 2004). Data cleaning resulted in the elimination of total of M (SD)= 6.1 (12.05) trials per participant in the NH group and M (SD)= 6.68 (11.63) trials per participant in the NH group. A portion of these rejected trials (~30%) were trials where participants responded “No”, that were not included in all types of analysis. For each tone level condition, participants with at least one valid trial were included in the analysis.

The data were time-aligned, adjusted to baseline, and averaged across all the accepted traces for a given condition. Baseline pupil diameter was defined as the mean pupil diameter recorded 200 ms before tone onset, which is consistent with studies that used a pure tone as a stimulus (e.g.,Schlemmer et al., 2005; Silk et al., 2009). The CHAP output pupil data curves were then smoothed using a five-point moving average filter.

As mentioned earlier, it could be argued that a difference between the groups in pupil dilation response in the study is inflated by a difference in their baseline pupil diameter due to any unknown impact of hearing loss on pupil size within a listening task in general that is not related to stimulus. Therefore, the difference between groups in pupil diameter within the listening experiment was first examined. Among the study participants, the average diameter of 1 second prior to tone onset for one of the conditions (16SL condition was arbitrary chosen) was not significantly larger in the NH group (*M*= 4.812, *SD*= .949) compared to the HL group (*M*= 4.356, *SD*= .793), *M*= .456, *SE*= .229, *t*(56)= 1.985, *p*= .052. The correlation between participants' hearing thresholds and their baseline pupil diameter pre tone-onset was examined and found to be not correlated, *r*(56)= -.16, *p*= .232 (Figure 38).

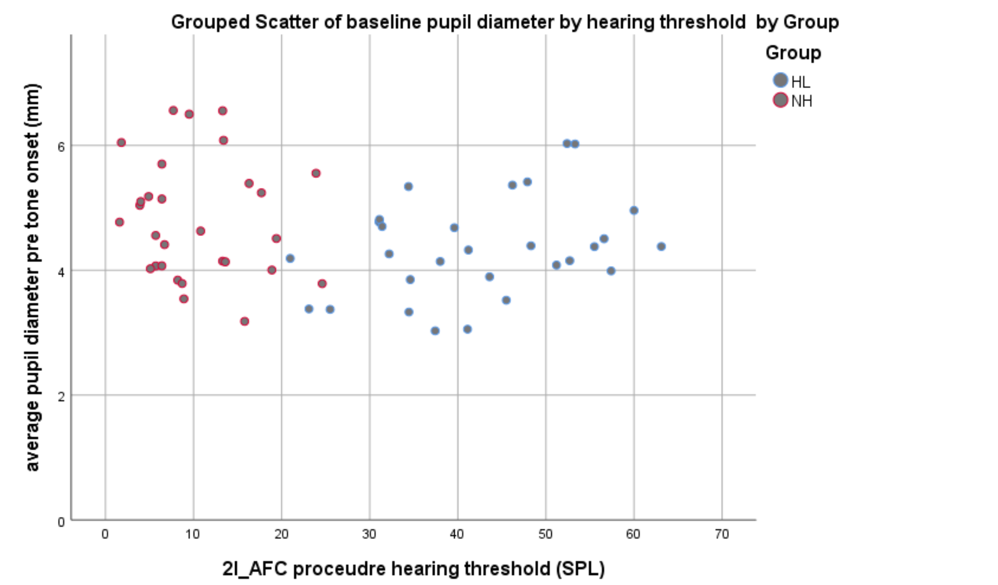


Figure 38: Scatterplot shows the relationship between participants' hearing threshold obtained from the 2I_AFC procedure in dB SPL and their 1 second average baseline pupil diameter pre-tone onset.

Additionally, Age of participants was not related to peak pupil dilation at 0SL condition, $r_s(52)$ = -.233, *p*= .097 (See Figure 39), or to peak dilation at 16 SL condition, $r_s(58)$ = -.147, *p*= .270 (See Figure 40).

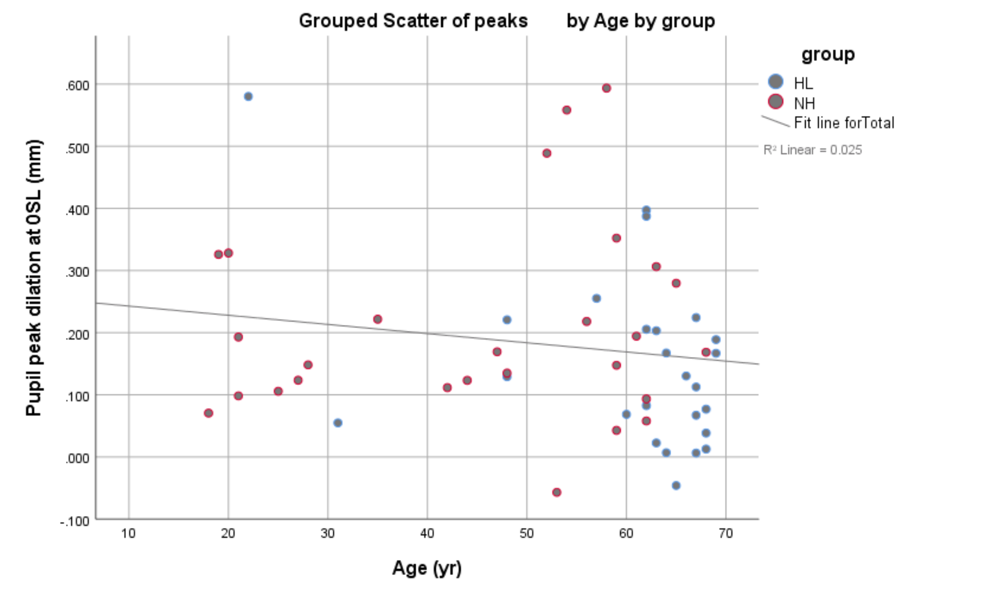


Figure 39: Correlation between age of participants and pupil peak dilation at 0 SL

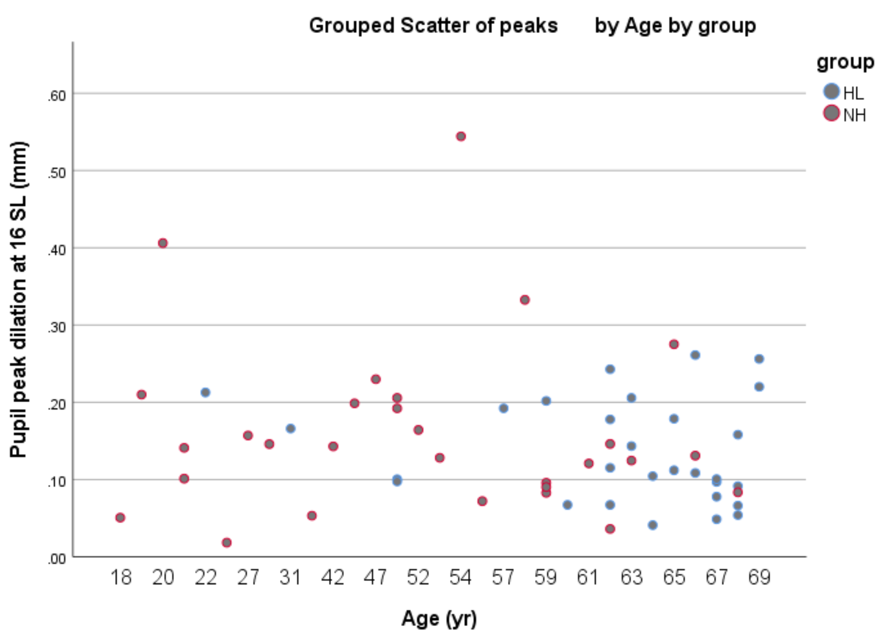


Figure 40: Correlation between age of participants and pupil peak dilation at 16 SL.

In pupillary recordings, there are three measures that are typically extracted from each interval of interest: mean pupil dilation, peak dilation, and latency to peak dilation (Beatty & Lucero-Wagoner, 2000b). Pupil dilation data from this experiment were analyzed using two different methods. The first was by looking at the difference in pupil dilation curves via statistical tests at each time point along the waveforms using Guthrie and Buchwald's method (Guthrie & Buchwald, 1991). This technique has been successfully applied in previous publications on pupil dilation data sets (Siegle, Ichikawa, & Steinhauer, 2008; Siegle et al., 2003a; Siegle et al., 2003b). The second was the more traditional approach by submitting DVs (mean, peak, and latency of peak) to linear mixed models with restricted maximum likelihood estimators (REML).

Guthrie and Buchwald (1999) approach

To test for differences between HL and NH group in the pupil dilation trajectory from 0 to 3000 ms at each tone level, individual Student's t tests were performed at each time point. This gave several time intervals of p-values significant at .1 and .05. To determine whether these intervals were statistically significant or due to autocorrelation, the correction proposed by Guthrie and Buchwald (1991) was used. This requires using Monte Carlo simulations based on the original data to generate the distribution of the length of significant time intervals under the assumption that there is no difference in response between groups.

To do this, first the series of pupil responses in each tone condition was determined to be (in an autocorrelation structure of order 1) AR(1) by examining plots of the autocorrelation function and the partial autocorrelation function (Shumway & Stoffer, 2011). Then the autocorrelation of the responses was estimated (Guthrie & Buchwald, 1991) and found to range from 0.91-0.97, depending on the tone condition. One-thousand data sets were generated with the same number of time points, autocorrelation, and sample size per group as the data to be analyzed. To generate the distribution of significant interval lengths, Student's t tests were performed at all time points and the length of the time interval of significant responses was recorded for each run. Based on this, interval lengths longer than the 95^{th} percentile of this distribution were considered significant.

The same method was used when testing for tone level differences within each group, but instead of using a Student's t test, a one-way ANOVA and corresponding F test were used. Pairwise tone level condition differences were tested for using Tukey's correction for multiple comparisons. This analysis was done with R version 3.2.1.

Linear Mixed Models

For each participant, pupil data were averaged for each condition. Pupil dilation mean, peak, and latency of that peak were derived within a time window starting at 500 ms post tone onset and ending at 3000 ms (with the end of the trial before the participant's response). Based on a visual inspection of the data, this range captured the entire rise and fall of the pupil response across conditions. Mean dilation, peak dilation, and latency to peak dilation were retrieved from the averaged curve. Mean dilation was the average of all values in the time window of interest. Peak dilation was identified as the average of the three maximum values within time window (Beatty & Lucero-Wagoner, 2000b). Peak dilation latency was the average of the tone that those three peaks occurred. All pupil dilation measures in this experiment will be reported in millimeters.

A mixed model was used to analyze the mean, peak, and latency data. Specifically, a restricted maximum likelihood estimators (REML) method was chosen. This analysis was used because it has several advantages over analyses of variance. First, it tolerates missing values and does not exclude participants with missing values from the analyses. This was an important factor due to the existence of missing data in the current study at different tone levels (mainly at -4, 0, and 4 SL). Second, mixed-effect modeling can be used to account for random effects of participants in analyses with repeated measures to improve the accuracy of inferences (Baayen, Davidson, & Bates, 2008).

The fixed effects were hearing condition group, tone levels, and a group by tone level interaction with subject as the random effect. A Q-Q plot was used to determine if the dependent variables were normally distributed. After model fit, heteroskedasticity were assessed using a scatter plot of residuals versus fitted values. Some evidence of heteroskedasticity was detected, but not enough to justify using any transformations. An exchangeable correlation structure was used, as it was the only type that the sample size would support. Linear contrasts based on the model were used to test for differences in groups and tone levels, and Tukey's method was used to adjust for multiple comparisons. Chi-squared values which returned p-values < .05 were considered statistically significant. Analysis was performed in STATA 15.1.

5.5.3 Experiment II: subjective rating experiment

Subjective ratings across all tone levels within the loudness function (up to 28SL) were compared between the NH and HL groups. The geometric mean of each participant's 10 responses at each tone level was calculated (Epstein & Florentine, 2006a; Gescheider & Hughson, 1991; Hellman & Zwislocki, 1963; Walker, 2002). A linear mixed model with REML with these means as the response was used to analyze the data. Fixed effects were group membership, tone level, and a group by tone interaction; subject was the random effect. Correlation was examined visually and found to be AR(2) using Bayesian Information Criteria. Linear contrasts based on the model were used to test for differences in groups and tone levels, and Tukey's method was used to adjust for multiple comparisons. Chi-squared values which returned p-values < .05 were considered statistically significant.

In order to more specifically investigate the recent proposed difference in "loudness" dynamic range in listeners with hearing loss (as opposed to only decreased "intensity" dynamic range), the distance in perception between a tone near threshold and a tone enough above threshold that is expected to exhibit an abnormally larger loudness in listeners with hearing loss (regardless of which model the listener's loudness function fits) was examined. The 4 SL tone condition was just above the participant's threshold, which allows for the comparison between the two models of loudness at threshold examined in this study. By comparing the difference in *loudness* dynamic range between 4 SL (near threshold level) and 16 SL, 20 SL, 24 SL, and 28 SL (supra threshold levels where faster growth of loudness is expected to manifest), finding could suggest (and locate) the manifestation (or not) of abnormal loudness growth within the loudness function in the hearing loss group. The difference in tone ratings was calculated for each participant. Then a Student's t test was used to analyze differences between the NH and the HL groups. P-value < .05 was considered statistically significant. Analysis was performed using STATA 15.1.

6.0 RESULTS

6.1 PARTICIPANTS

A hundred and fourteen potential participants attended the research session and signed the consent form. Seventy participants (32 with hearing loss and 38 with normal hearing) satisfied all of the study criteria and participated in the research activities. Of those 70 participants, one of participants in the NH group was not able to complete the first experiment (sound detection experiment) due to abnormally small pupil size that prevented pupil diameter recording, but completed the subjective rating experiment. It was decided not to include this participant's subjective rating data in the analysis to maintain the consistency in data obtained from both experiments.

Participants of sound detection experiment

Based on participants' behavioral responses in the sound detection experiment (experiment I), data from 8 participants in the NH group and 2 participants in the HL group had to be excluded from the analysis. The rational for exclusion was explained in section 5.5.2.1. One participant in the HL group found to have a conductive hearing loss (15 dB or greater air-bone gap at more than one frequency), and therefore, this participant's data were excluded from analysis. The whole exclusion process resulted in a sample size of 58 participants (29 in each group) included in the analysis of data from the sound detection experiment. Past studies of pupil dilation that compared responses between two groups have drawn conclusions from approximately 20-30 participants in each group (Ohlenforst et al., 2017; Siegle, Granholm, Ingram, & Matt, 2001) and less than 20 participants in each group (Siegle et al., 2003a; Van Gerven et al., 2004).

Age of participants ranged from 18 to 68 (M= 46.21, SD= 16.78, median= 52) in the NH group, and from 22 to 69 (M= 60.93, SD= 10.95, median= 64) in the HL group. Figure 42 shows the distribution of age among the participants whose data were included in analysis (sound detection data). Figure 41 shows the average pure tone air conduction audiometric hearing threshold for both groups. Gender was equally distributed in both groups. Each group included 15 females and 14 males. None of the participants in the normal hearing group reported a diagnosis of a psychological disorder. However, 4 participants in the hearing loss group reported either mild depression or mild anxiety, but only 2 of them were under medication for it (Celexa or Duloxotine). Due to the difficulty of recruiting in the hearing loss groups, these participants were included when their pupil size and amplitude of light and accommodation reflexes were judged by the PI to be non-significant/normal. Reported etiology of hearing loss in the HL group varied between congenital (n=1), noise induced (n=4), presbycusis (n=6), a combination of noise induced and presbycusis (n=1), and unknown etiology (n=9). Eight participants in the HL group reported no difficulty hearing, so they were not aware of their hearing loss. Reported self-perceived duration of hearing loss ranged from 1 to 40 years (M=11.6, SD= 9.2).

All participants were consented and completed a one visit, 2-hour research session.

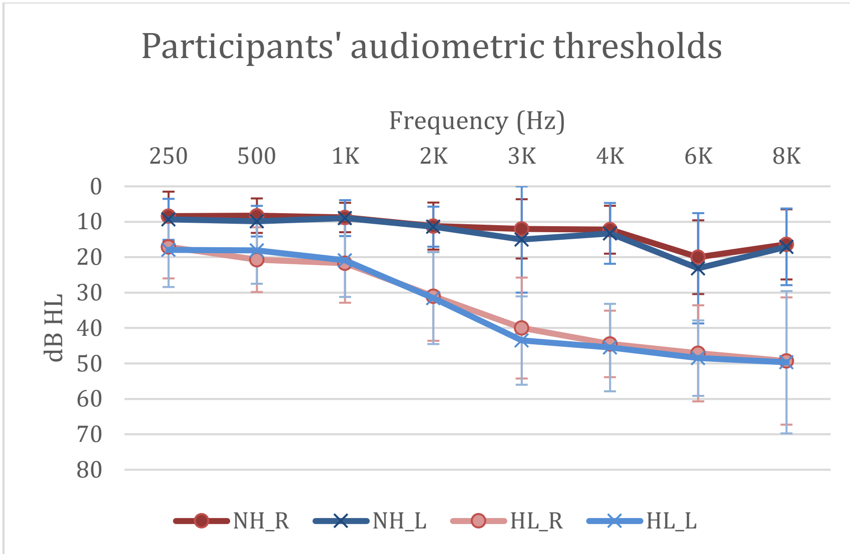


Figure 41: Participants' average audiometric thresholds (bars indicate 1 standard deviation).

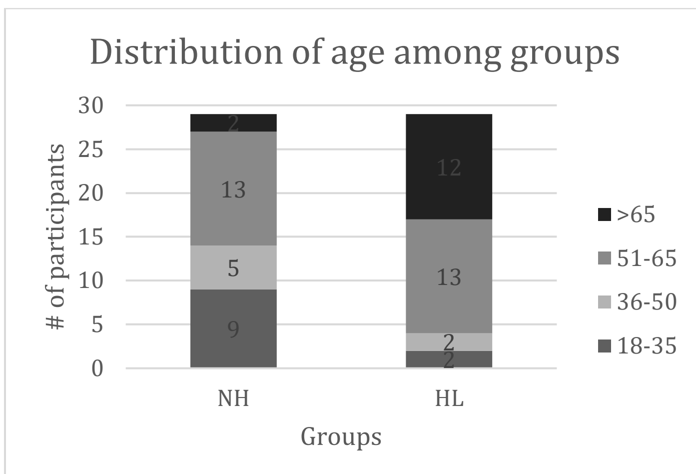


Figure 42: Distribution of age among study groups.

Participants in the subjective rating experiment

All participants in the sound detection experiment participated in the subjective rating experiment. Of those 58 participants, three did not provide complete responses in the loudness rating sheet, so their loudness ratings could not be matched with ordered stimulus conditions and hence could not be analyzed.

Thus, a sample size of 55 participants (27 in the NH group and 28 in the hearing loss group) included in the analysis of the subjective rating experiment.

6.2 EXPERIMENT I: SOUND DETECTION EXPERIMENT

6.2.1 Specific Aim 1 and 2

Is there a significant difference in the pupillary dilation response to tones presented **at threshold** (0 dB SL) between listeners with hearing loss and listeners with normal hearing?

Is there a significant difference in the pattern of change of pupillary dilation response with increasing sound level within the range of **low sensation levels** (4 dB SL, 10 dB SL, and 16 dB SL) between listeners with hearing loss and listeners with normal hearing?

6.2.1.1 Guthrie and Buchwald (1999) approach

Averaged pupil dilation waveforms were contrasted between groups at each tone level condition. Figure 43 shows the averaged waveforms for both groups at each tone level condition.

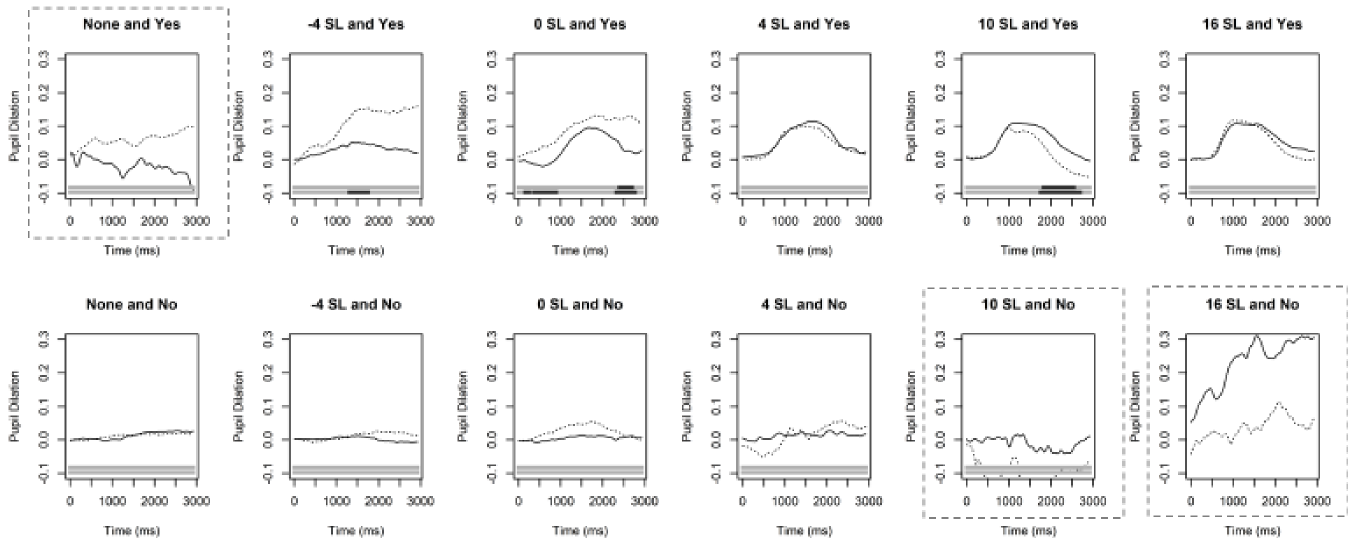


Figure 43: Average pupil waveforms and the results if significance tests. Waveforms from normal hearing listeners (dotted line) and listeners with hearing loss (solid line) at each tone level (columns) when participants responded "yes, I heard it" (upper panel) and when participants responded "no, I did not hear it" (lower panel). Regions of significant differences in the individual t-tests are highlighted along the x-axis by black bars. Top bar indicates significance at 0.05 p-value, and bottom bar indicates significance at 0.1 p-value. Graphs in dotted line boxes were obtained from minimum number of participants and trials but were presented for data completion.

There were no significant differences in the pupil dilation responses between the groups when participants did not hear the tones (responded "no, I did not hear it"). Table 15 shows the number of participants who contributed to the grand mean waveform at each tone level condition for both responses (yes and no), as well as the average number of trials included for each participant.

Table 15: Sample size for responses at each tone level in each group (NH/HL) and average number of trials included in the analysis (M) and standard deviation (SD).

		None	-4 SL	0 SL	4 SL	10 SL	16 SL
Yes	# of participants	24 (11/13)	37 (19/18)	52 (27/25)	58 (29/29)	58 (29/29)	58 (29/29)
	# of trials/participant	M=(1.1/0.8) SD=(2.1/1.0)	M=(2.2/1.5) SD=(2.6/1.7)	M=(5.0/3.4) SD=(3.9/4.0)	M=(9.4/7.8) SD=(3.4/3.8)	M=(11.9/11.5) SD=(4.8/2.8)	M=(11.7/11.7) SD=(5.2/3.5)
No	# of participants	58 (29/29)	58 (29/29)	50 (25/25)	28 (12/16)	5 (2/3)	2 (1/1)
	# of trials/participant	M=(10.6/10.9) SD=(5.4/3.2)	M=(9.5/9.9) SD=(5.8/3.4)	M=(6.7/8.1) SD=(6.4/5.1)	M=(2.6/3.6) SD=(6.6/5.0)	M=(0.1/0.2) SD=(0.3/0.8)	M=(0/0) SD=(0.2/0.2)

When participants perceived the tone (responded “yes”), the following differences were found. At -4 SL tone condition, there was a significant window of greater pupil dilation in the NH group from 1309 to 1819 ms, D=-.099, $t(35)= 1.85$, $p= .073$, $d= -.585$. This observed difference appeared to occur in the rising phase of the pupil response. At 0 SL tone condition, there was a significant window of greater pupil dilation in the NH group from 153 to 969 ms, D=-.048, $t(50)= 1.90$, $p= .063$, $d= -.529$, and from 2329 to 2839 ms, D= -.096, $t(50)= 2.08$, $p= .043$, $d= -.562$. This observed difference appeared to occur in both rising phase and the phase that followed maximum dilation. At 10 SL tone condition, there was a significant window of greater pupil dilation in the HL group from 1751 to 2771 ms, D=.058, $t(56)= 2.08$, $p= .042$, $d= .541$. This observed difference appeared to occur in the phase that follows maximum dilation. There were no significant differences between the groups in the pupil dilation response to other tone conditions (4 SL and 16 SL) and silence condition.

Averaged pupil dilation waveforms were also contrasted between tone level conditions within each group. Figure 44 shows the averaged waveforms for all tone levels in each group.

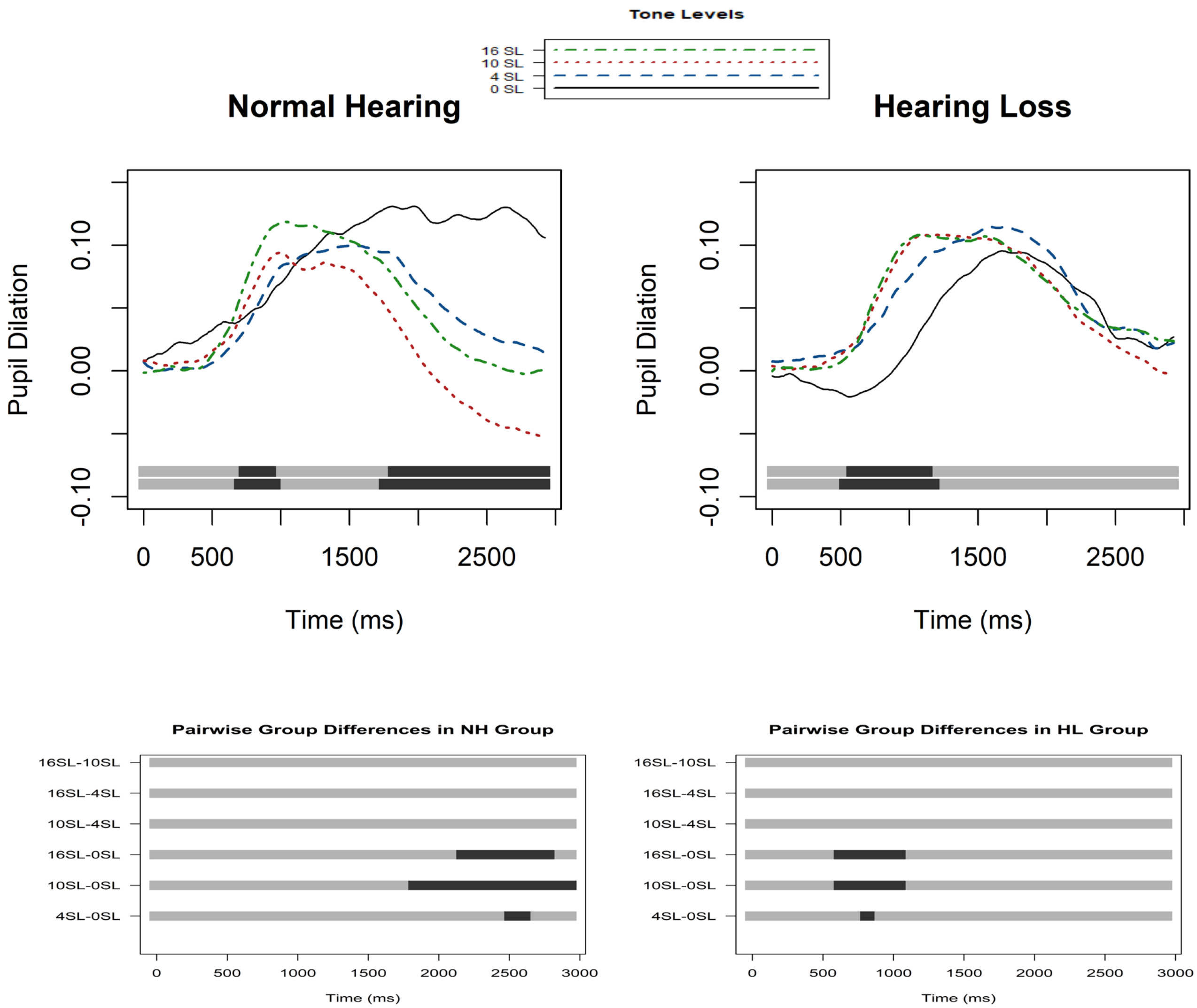


Figure 44: Average pupil dilation response to different tone levels within each group. Responses to all tone level conditions (Black= 0SL, Blue= 4 SL, Red= 10 SL, Green= 16 SL) from normal hearing listeners (left panel) and listeners with hearing loss (right panel) when participants responded "yes, I heard it". Regions of significant differences (main effect of tone level) in the individual ANOVA tests are highlighted along the x-axis by black bars. Top bar indicates significance at 0.05 p-value, and bottom bar indicates significance at 0.1 p-value. Lower graph: Highlighted regions of significant differences in pairwise comparison within a group.

In the normal hearing group, there was a window of significant main effect of tone level extending from 697 to 1020 ms post tone onset. The pupil dilation increased with increasing tone level, $F(3,110)= 5.4$, $p= .032$. However, pairwise comparison with Tukey's correction revealed no window of significant difference between any individual tone levels within that region. There also was another window of significant main effect of tone level condition in a later phase of the response, with pupil dilation decreasing with increasing tone level, from 1751 to 2924 ms, $F(3,110)= 9.0$, $p= .010$. Pairwise comparison with Tukey's correction revealed that there was a window of significantly smaller pupil dilation in 4 SL tone condition (D= .105) from 2516 to 2737 ms, 10 SL condition (D= .147) from 1836 to 2924 ms, and 16 SL tone condition (D= .117) from 2176 to 2856 ms than that in 0 SL condition, $t(54)= 2.12$, $p= .039$, $d= .643$; $t(54)= 3.01$, $p= .004$, $d= .868$; and $t(54)= 2.16$, $p= .018$, $d= .695$, respectively. No time window of significant difference existed between any other tone level conditions.

In the hearing loss group, opposite main effect and pairwise comparison findings existed in the later phase of the response in the NH group was found in the early phase of the response in the HL group. There was a window of significant main effect of tone level extending from 527 to 1241 ms post tone onset. The pupil dilation increased with increasing tone level, $F(3,108)= 10.5$, $p= .016$. Pairwise comparison with Tukey's correction revealed that there was a window of significantly larger peak dilation in 4 SL tone condition (D= -.060) from 816 to 901 ms, 10 SL condition (D= -.075) from 646 to 1122 ms, and 16 SL tone condition (D= -.078) from 629 to 1122 ms than that in 0 SL condition, $t(52)= 2.15$, $p= .036$, $d= -.598$; $t(52)= 2.71$, $p= .009$, $d= -.730$; and $t(52)= 2.70$, $p= .009$, $d= -.752$, respectively. No time window of significant difference existed between any other tone level conditions.

6.2.1.2 Linear Mixed Models

As a comparison to Guthrie and Buchwald's (1991) approach on analyzing the pupillary data, a traditional approach on analyzing the mean, peak and latency of peak dilation was performed. A mixed model was used to analyze the data. The fixed effects were hearing group with two levels (NH and HL), tone level condition with 4 levels (0 SL, 4 SL, 10 SL, and 16 SL), and a group by tone level interaction with subject as the random effect. Figure 45 shows the averaged mean, peak, and peak latency of pupil dilation across tone level conditions in the NH group and HL group.

Peak pupil dilation

There was no main effect of group, $\chi^2(1)$= 1.30, *p*= 0.254, but there was a main effect of tone level on peak pupil dilation, $\chi^2(3)$= 9.99, *p*= 0.0186, with a decrease in peak dilation with increasing tone level overall. Peak pupil dilation in the NH group significantly decreased with increasing tone level $\chi^2(3)$= 11.86, *p*= 0.0079, whereas the peak pupil dilation did not differ with increasing tone level in the HL group. $\chi^2(3)$= 1.22, *p*= 0.747. Pairwise comparison within the NH group showed that peak pupil dilation at 0 SL tone condition was significantly larger than that at 4 SL, 10 SL, and 16 SL, *D*=-.052, *z*= -2.60, *Delta-method SE*= .02, *p*= .009, *d*= -.416, *D*=-.065, *z*= -3.26, *Delta-method SE*= .02, *p*= .001, *d*= -.509; and *D*=-.045, *z*= -2.28, *Delta-method SE*= .02, *p*= .023, *d*= -.326, respectively No significant difference existed between any other conditions. There was no significant interaction effect of group by tone level, $\chi^2(3)$= 2.73, *p*= .434. Table 16 shows the descriptive statistics values for the peak pupil dilation in both groups.

Table 16: Descriptive statistics of peak pupil dilation in both groups.

Group	*NH*		*HL*	
	Mean	Standard Error	Mean	Standard Error
0 SL	0.2073	0.0299	0.1503	0.0286
4 SL	0.1550	0.0171	0.1411	0.0181
10 SL	0.1416	0.0163	0.1323	0.0121
16 SL	0.1613	0.0210	0.1367	0.0123

Mean pupil dilation

There was a significant interaction effect of group x tone level condition in mean pupil dilation $\chi^2(3)$= 12.89, *p*= .0049. Overall, there was no significant main effect of tone level, $\chi^2(3)$= 6.34, *p*= .096, or main effect of group, $\chi^2(1)$= 0.0, *p*= .983. However, there was a main effect of tone level on mean pupil dilation in the NH group $\chi^2(3)$= 18.42, *p*= .0004, but not in the HL group, $\chi^2(3)$= 1.27, *p*= .736. Pairwise comparison within the NH group showed that mean pupil dilation at 0 SL was significantly larger than that at 4 SL, 10 SL, and 16 SL, *D*=-.044, *z*=-2.45, *Delta-method SE*= .018, *p*= .014, *d*= -.421; *D*=-.077, *z*= -4.28, *Delta-method SE*= .018, *p*< .0001, *d*= -.696; and *D*=-.045, *z*= -2.48, *Delta-method SE*= .018, *p*= .013, *d*= -.376, respectively. No significant differences existed between any other conditions. When examining group differences at each tone level, the NH groups showed significantly larger mean pupil dilation than the HL group at 0 SL tone level condition, $\chi^2(1)$= 5.43, *p*= .0197, *D*= -.515, *d*= -.057. No significant difference existed between the groups at any other conditions. Table 17 shows the descriptive statistics values for the mean pupil dilation in both groups.

Table 17: Descriptive statistics of mean pupil dilation in both groups.

Group	*NH*		*HL*	
	Mean	Standard Error	Mean	Standard Error
0 SL	0.1030	0.0233	0.0460	0.0182
4 SL	0.0596	0.0159	0.0676	0.0147
10 SL	0.0263	0.0166	0.0626	0.0096
16 SL	0.0589	0.0183	0.0681	0.0090

Pupil dilation peak latency

There was no main effect of group $\chi^2(1)= 0.47$, $p= .4915$, but there was a significant main effect of tone level on peak latency, $\chi^2(3)= 32.15$, $p< .0001$, with sooner dilation peak occurring with increasing tone level overall. This main effect of tone level existed in both groups, NH: $\chi^2(3)= 20.09$, $p= .0002$, HL: $\chi^2(3)= 13.05$, $p= .0045$. There was no significant interaction effect of group by tone level on peak latency, $\chi^2(3)= .80$, $p= .850$. Pairwise comparison within the NH group revealed that peak latency at 0 SL tone condition was significantly delayed than that at 4 SL, 10 SL, and 16 SL, *D*= -365.106, *z*= -2.69, *Delta-method SE*= 135.586, *p*= .007, *d*= -.582; *D*= -492.117, *z*= -3.63, *Delta-method SE*= 135.586, *p*< .0001, *d*= -.741; and *D*= -560.117, *z*= -4.13, *Delta-method SE*= 135.586, *p*< .0001, *d*= -.883, respectively. No significant differences existed between any other conditions. The same differences were also found within the HL group. Peak latency at 0 SL tone condition was significantly delayed than that at 4 SL, 10 SL, and 16 SL, *D*= -316.069, *z*= -2.28, *Delta-method SE*= 138.558, *p*= .023, *d*= -.572; *D*= -467.311, *z*= -3.37, *Delta-method SE*= 138.558, *p*= .001, *d*= -.867; and *D*= -402.632, *z*= -2.91, *Delta-method SE*= 138.558, *p*= .004, *d*= -.699, respectively. No significant differences existed between any other conditions. Table 18 shows the descriptive statistics values for the pupil dilation peak latency in both groups.

Table 18: Descriptive statistics of pupil dilation peak latency in both groups.

Group	*NH*		*HL*	
	Mean	Standard Error	Mean	Standard Error
0 SL	1875.66	128.475	1879.74	118.762
4 SL	1514.75	100.706	1560.09	88.315
10 SL	1387.74	107.364	1408.85	69.663
16 SL	1319.74	87.454	1473.5	92.341

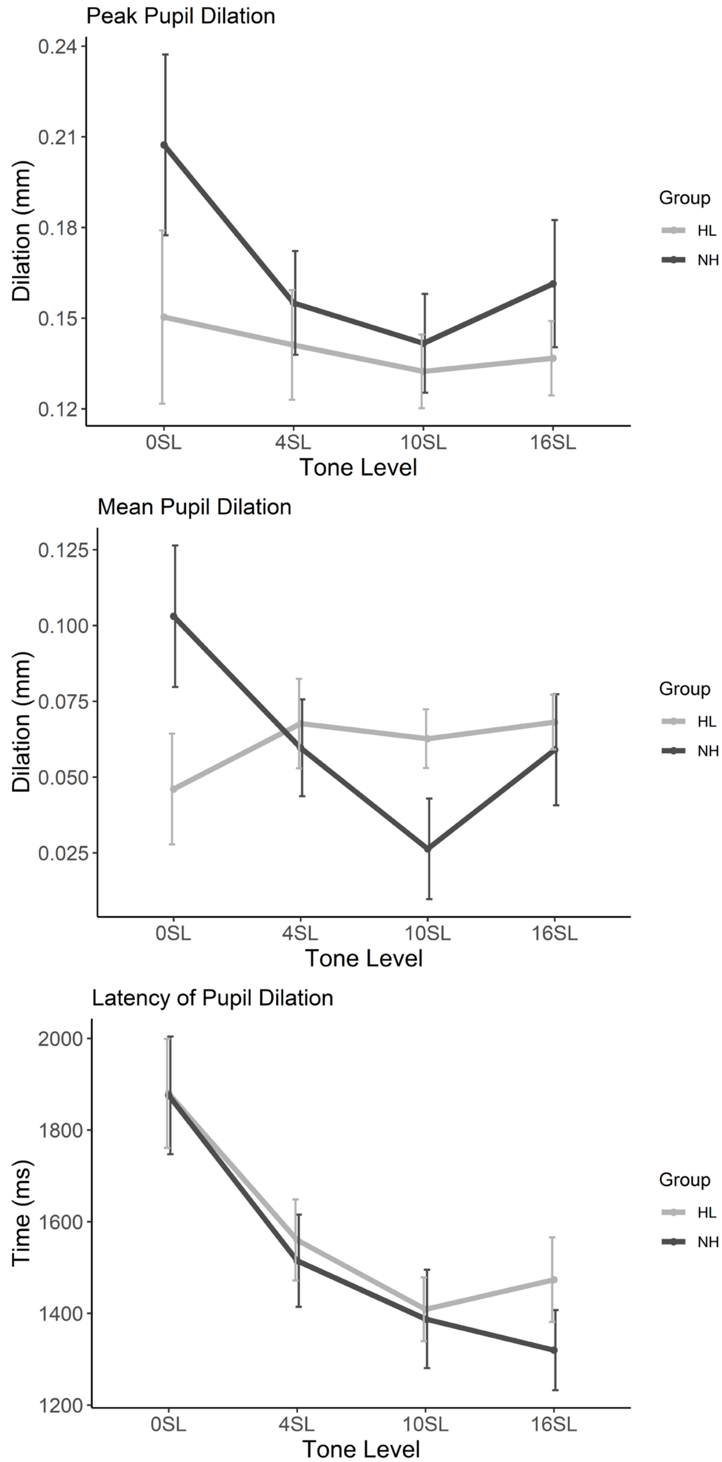


Figure 45: Averaged mean, peak, and peak latency of pupil dilation across tone level conditions in the NH group and HL group.

6.3 EXPERIMENT II: SUBJECTIVE RATING EXPERIMENT

6.3.1 Specific Aim 3

Is the dynamic range in terms of perceived loudness as measured by the difference in subjective loudness ratings of two tones presented at different SLs (difference between 4 SL and 16, 20, 24, and 28 SL, respectively) significantly different in listeners with normal hearing and listeners with hearing loss?

Before comparing the distance between the ratings of different tones, the whole loudness function obtained was contrasted between the groups.

Figure 46 shows the loudness ratings obtained across tone levels from each group. There was a significant main effect of tone level $\chi^2(8) = 3032.41$, $p < 0.0001$ on loudness rating. Loudness rating significantly increased with increasing tone level. Main effect of tone level was significant in both NH and HL groups, $\chi^2(8) = 1535.26$, $p < 0.0001$, $\chi^2(8) = 1517.63$, $p < 0.0001$, respectively. In both groups, pairwise comparison revealed a significant increase in loudness rating with each increase in tone level within the loudness function except for -4 to 0 SL, and 0 to 4 SL. In the normal hearing group, there was no significant increase in loudness rating from -4 to 4 SL, and from 4 to 8 SL. All other significant increases in rating were consistent in both groups of listeners. Statistics values for these pairwise comparison are shown in Table 19.

Table 19: Statistics values for the pairwise comparison across tone levels within both groups. Shaded cells indicate differences in ratings that were not statistically significant.

Group	***NH***				***HL***			
	Contrast	Delta-method Std. Err.	Unadjusted z	P>|z|	Contrast	Delta-method Std. Err.	Unadjusted z	P>|z|
-4 vs 0 SL	1.572	4.300	0.37	0.715	4.589	4.219	1.09	0.277
-4 vs 4 SL	8.2268	4.300	1.91	0.056	10.198	4.2196	2.42	0.016
-4 vs 8 SL	15.33107	4.300	3.57	<0.0001	20.763	4.2196	4.92	<0.0001
-4 vs 12 SL	30.435	4.300	7.08	<0.0001	33.824	4.2196	8.02	<0.0001
-4 vs 16 SL	48.995	4.300	11.39	<0.0001	56.146	4.2196	13.31	<0.0001
-4 vs 20 SL	84.612	4.300	19.68	<0.0001	90.062	4.219	21.34	<0.0001
-4 vs 24 SL	95.156	4.300	22.13	<0.0001	102.666	4.683	21.92	<0.0001
-4 vs 28 SL	105.011	4.300	24.42	<0.0001	120.014	4.683	25.62	<0.0001
0 vs 4 SL	6.6546	4.300	1.55	0.122	5.608	4.219	1.33	0.184
0 vs 8 SL	13.758	4.300	3.20	0.001	16.174	4.219	3.83	<0.0001
0 vs 12 SL	28.863	4.300	6.71	<0.0001	29.234	4.219	6.93	<0.0001
0 vs 16 SL	47.422	4.300	11.03	<0.0001	51.557	4.219	12.22	<0.0001
0 vs 20 SL	83.040	4.300	19.31	<0.0001	85.472	4.219	20.26	<0.0001
0 vs 24 SL	93.584	4.300	21.76	<0.0001	98.076	4.683	20.94	<0.0001
0 vs 28 SL	103.439	4.300	24.06	<0.0001	115.424	4.683	24.64	<0.0001
4 vs 8 SL	7.1042	4.300	1.65	0.099	10.565	4.219	2.50	0.012
4 vs 12 SL	-22.209	4.300	-5.16	<0.0001	-23.625	4.219	-5.60	<0.0001
4 vs 16 SL	-40.768	4.300	-9.48	<0.0001	-45.948	4.219	-10.89	<0.0001
4 vs 20 SL	-76.385	4.300	-17.76	<0.0001	-79.864	4.219	-18.93	<0.0001
4 vs 24 SL	-86.929	4.300	-20.22	<0.0001	-92.468	4.683	-19.74	<0.0001

Table 19 (continued).

4 vs 28 SL	-96.785	4.300	-22.51	<0.0001	-109.816	4.683	-23.45	<0.0001
8 vs 12 SL	-15.104	4.300	-3.51	<0.0001	-13.060	4.219	-3.10	0.002
8 vs 16 SL	-33.664	4.300	-7.83	<0.0001	-35.382	4.219	-8.39	<0.0001
8 vs 20 SL	-69.281	4.300	-16.11	<0.0001	-69.298	4.219	-16.42	<0.0001
8 vs 24 SL	-79.825	4.300	-18.56	<0.0001	-81.902	4.683	-17.49	<0.0001
8 vs 28 SL	-89.680	4.300	-20.86	<0.0001	-99.250	4.683	-21.19	<0.0001
12 vs 16 SL	18.559	4.300	4.32	<0.0001	22.322	4.219	5.29	0.026
12 vs 20 SL	54.176	4.300	12.60	<0.0001	56.238	4.219	13.33	<0.0001
12 vs 24 SL	64.720	4.300	15.05	<0.0001	68.842	4.683	14.70	<0.0001
12 vs 28 SL	74.575	4.300	17.34	<0.0001	86.1901	4.683	18.40	<0.0001
16 vs 20 SL	35.617	4.300	8.28	<0.0001	33.915	4.219	8.04	<0.0001
16 vs 24 SL	46.161	4.300	10.74	<0.0001	46.519	4.683	9.93	<0.0001
16 vs 28 SL	56.016	4.300	13.03	<0.0001	63.867	4.68	13.64	<0.0001
20 vs 24 SL	10.544	4.300	2.45	0.014	12.604	4.683	2.69	0.007
20 vs 28 SL	20.399	4.300	4.74	<0.0001	29.951	4.683	6.40	<0.0001
24 vs 28 SL	9.855	4.300	2.29	0.022	17.347	5.030	3.45	0.001

There was no significant main effect of group, $\chi^2(1)= 2.36$, $p= .1248$. However, the hearing loss group rated the loudness of 28 SL tones significantly louder than the normal hearing group, $\chi^2(1)= 6.88$, p= .0087. No significant differences existed between the groups at any other tone level conditions. There was also no significant interaction effect between group and tone level, $\chi^2(8)= 7.24$, $p= .5107$.

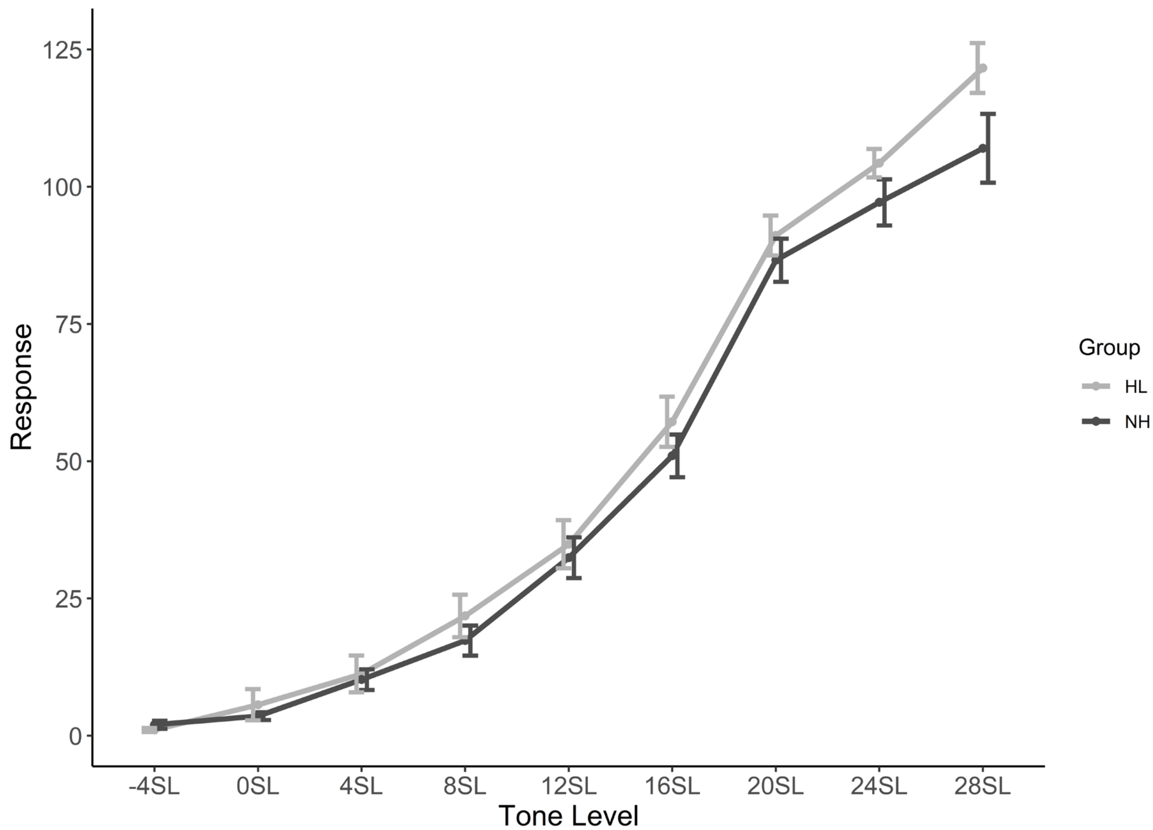


Figure 46: loudness functions obtained by magnitude estimation of subjective loudness rating from participants in both groups.

Figure 47 shows the difference in loudness rating of 4 SL and 16 SL, 4 SL and 20 SL, 4 SL and 24 SL, and 4 SL and 28 SL. The dynamic range of loudness between these pairs of tones did not differ between the groups, 4-16 difference: D= 5.16, $t(51)$= -.85, p= .3968, d= .238; 4-20 difference: D= 3.47, $t(51)$= .40, p= .5905, d= .129; 4-24 difference: D= 5.48, $t(43)$= .54, p= .4763, d= .178; and 4-28 difference: D= 13.0, $t(43)$= 1.27, p= .2100, d= .336.

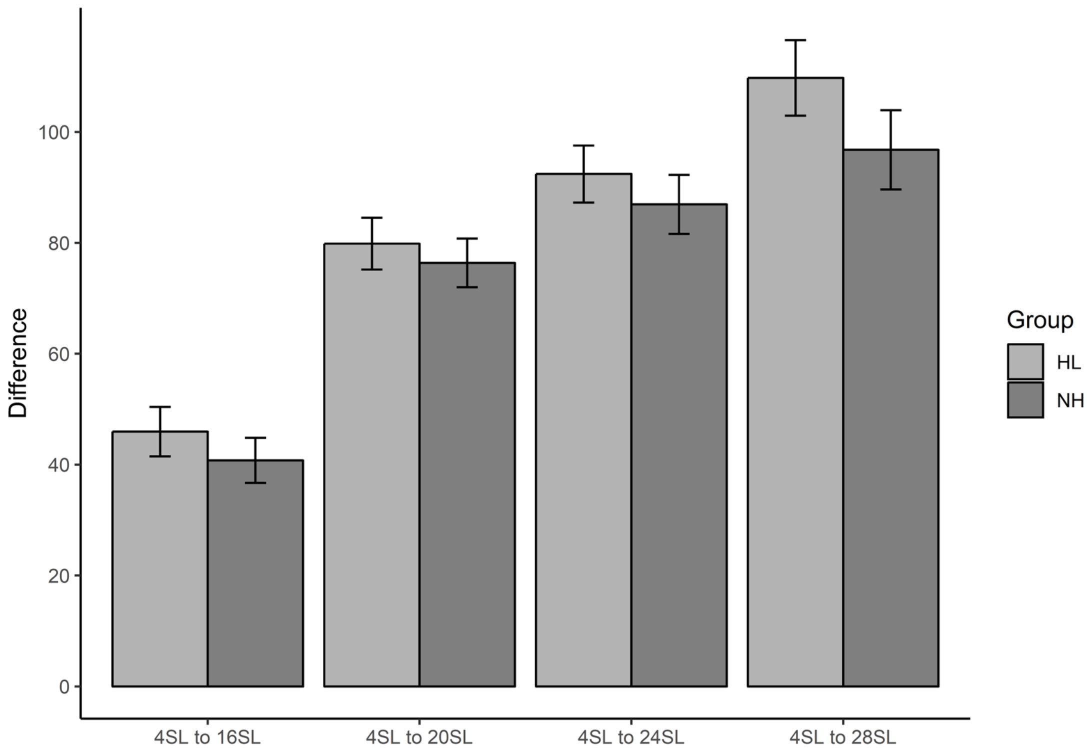


Figure 47: Difference in loudness ratings between 4 SL tone and higher SL tones compared between the groups.

7.0 DISCUSSION

This study sought to determine how listeners with hearing loss perceive the loudness of low sensation level sounds (sounds at and near threshold) compared to normal hearing listeners. Specifically, do individual with hearing loss perceive sounds at threshold as loud as normally hearing listeners? Do loudness of sounds near elevated threshold grows abnormally fast with increasing sound level?

7.1 EXPERIMENT I: SOUND DETECTION EXPERIMENT

7.1.1 Specific Aim 1

The first aim focused on differences in pupil dilation across groups at threshold (0 SL condition). The results of Guthrie and Buchwald's (1991) method of time-by-time comparison of pupil dilation post tone onset between the groups revealed that normal hearing listeners showed a pupil dilation response that was significantly different from listeners with hearing loss in both rising phase and overall behavior. The magnitude of maximum dilation in normal hearing listeners was similar to the magnitude of maximum dilation in listeners with hearing loss. Dilation started sooner for normal hearing listeners and once dilation reached maximum, it was sustained longer and did not return back to baseline as quickly, compared to listeners with hearing loss. The results of the more standard approach of analyzing pupillary data by examining the mean, peak, and latency of peak difference, complimented the results of Guthrie and Buchwald's approach. The normal hearing group showed significantly larger mean pupil dilation across the whole time of the response, but not larger peak dilation or delayed peak dilation than listeners with hearing loss. These results line up with the difference found in the overall behaviors of the response, and not in the magnitude of maximum dilation (peak), or the timing where that maximum dilation occurred (peak latency). It should be noted that latency of peak dilation was shifted at threshold (0 SL) in both groups compared to latency of peak dilation at higher tone level conditions. This delayed peak dilation at threshold reflects a generally longer time needed for cognitive processing and an increased processing load (Van Der Meer et al., 2010) that was not shown when detecting tones above threshold; hence, peak dilation in response to tones above threshold all occurred at similar latencies.

In the neuropsychology, cognitive psychology, and psychophysiology literature, pupil dilation has previously been established as a physiological marker of decision-making uncertainty that gradually increases during the decision process (de Gee, Knapen, & Donner, 2014; Lempert, Chen, & Fleming, 2015; Nassar et al., 2012). It has also been established as an indicator of cognitive effort (Beatty, 1982b) which can be related to task uncertainty. Perceived changes like those in detection at near-threshold tasks also can be reflected in changes in pupil dilation (Nassar et al., 2012). In this study, the task of sound detection at threshold (at 0 SL) is a difficult task that requires decision-making and involves uncertainty, all of which is not the case in the task of sound detection at other tone level conditions (i.e., above threshold sound detection at 4 SL, 10 SL, and 16 SL conditions).

Listeners with normal hearing showed longer (more sustained) pupillary dilation response to sounds presented at their threshold compared to the response in listeners with hearing loss at their elevated threshold. This finding is consistent with an increase in cognitive processing load duration in normally hearing listeners, as well as increased decision uncertainty when detecting tones at threshold. Previous evidence suggests that pupils dilate during decision formation and peak just after making the decision (de Gee et al., 2014; Urai, Braun, & Donner, 2017).

The earlier initiation of pupillary response in normal hearing listeners is less understood. It could suggest more uncertainty in the decision-making involved in the detection task, compared to listeners with hearing loss whose response rising phase matched with the response to easier detections of 4 SL, 10 SL, and 16 SL tones.

The more defined pupillary response in listeners with hearing loss that initiated and ended in a similar (although temporally shifted) manner to responses obtained in easier detection tasks with less uncertainty (i.e., at 4, 10, and 16 SL), might reflect an easier detection task for sounds at their elevated thresholds compared to normal hearing listeners detecting sounds at their normal threshold. This easier detection task in listeners with hearing loss could be explained due to louder perception of that tone at threshold. This interpretation of louder perception of tones at threshold in listeners with hearing loss is plausible given the equal presentation level of tones (0 SL), and equal performance (i.e., detection sensitivity) in the task in both groups, but different cognitive processing involved within the task, as indicated by the pupil response. The difference in the auditory evoked pupillary response indicated different cognitive processing load duration (and not magnitude) to accomplish the same performance. Although it is not completely clear, it has been suggested that the magnitude of peak pupil dilation reflects the maximum processing load carried out by the listening task, and that the mean dilation is more related to the total processing load in the time course of cognitive activity (Granholm, Morris, Asarnow, Chock, & Jeste, 2000; Siegle et al., 2001; Siegle et al., 2003b).

This finding is interesting given what is known about aging and hearing loss as factors related to less release from effort (Zekveld, Kramer, & Festen, 2011). A more sustained pupil dilation in normal hearing listeners indicated that the task involved a reduced duration of effortful processing for listeners with hearing loss, and that could be related to louder perception of tones at threshold, making the detection task less effortful. Given similar baseline pupil sizes between the groups (HL group did not have a significantly larger baseline pupil diameter and therefore, was expected to have similar pupil range of reactivity), the difference in their pupil responses found in this study is otherwise unexplainable.

This finding of larger loudness at threshold in listeners with hearing loss supports the SI loudness model by Buus and Florentine (2002). The decreased cognitive processing load duration apparent in the pupil dilation response in listeners with hearing loss also agrees with the finding of faster reaction time to detect tones at threshold in listeners with hearing loss found by Florentine et al. (2005), suggesting a more certain detection compared to delayed response time shown by normal hearing listeners suggesting a softer perception for tones at threshold and hence an uncertain detection. In that study, similar interpretations were made about louder perception at threshold in these listeners. This finding also agrees with clinical observation during pure tone audiometric testing. Once tone presentation approaches threshold, clinicians observe faster response to tones at elevated threshold than tones at a normal threshold.

It should be noted that the observed difference between the groups in pupil dilation response at threshold (0 SL condition) was also present in the initial phase of the response to tones presented just below threshold (-4 SL condition). Normal hearing listeners showed a larger and more sustained dilation than listeners with hearing loss, indicating similar processing demand and uncertainty difference between groups in this difficult detection task.

7.1.2 Specific Aim 2

The second aim focused on differences in pupil dilation across groups with increasing tone level near threshold (0 SL, 4 SL, 10 SL, and 16 SL). Once stimuli were presented above threshold, task difficulty was not expected to be a factor in different cognitive processing load, since the stimuli are now always easily detectable, and the pupillary response appeared to be insensitive to reflect different loudness perception. Increased subjective ratings with increased tone level obtained from the subjective rating experiment in this study (see section 7.2.1 for discussion) confirms the change in perceived loudness with increasing tone level that was not reflected in the pupil dilation response. All pupil dilation responses started at similar time right after 500 ms post tone onset, but contrary to what was hypothesized, did not change in amplitude, latency, or duration (mean dilation) with increasing level in both groups. This suggests that, unless the perceived change in loudness is drastic, possibly resulting in a different level of arousal or startle, the pupil response is expected to stay unchanged with increasing tone level. Pupil dilation response to startle sounds (95, 100, and 105 dB SPL) had previously shown to peak with higher amplitude than peaks obtained in this study (Wiemer, 2014).

Guthrie and Buchwald's (1991) approach in comparing the pupil dilation response between the groups showed that response at 10 SL condition peaked at a similar time in both groups, but was sustained longer in listeners with hearing loss. This difference was not shown when compared to the mean overall dilation between the groups and could not be explained within the scope of this study. Because the same pattern of increased sustained dilation in listeners with hearing loss at 10 SL tone condition was not found at 16 SL tone condition, it could not be interpreted as possible increased loudness growth at those levels in listeners with hearing loss, which would have been consistent with the rapid growth model of loudness by Moore et al., (2004).

While it is difficult to pinpoint cognitive functions underlying pupillary dilation response, it can be concluded that the pupil dilation response could be an indirect index of perceived loudness when it affects the cognitive processing involved in the detection of sounds, and not an index of perceived loudness. Given that tones above threshold in this study were all near threshold tones, they elicited pupillary responses similar in latency, magnitude, and duration. This pattern of pupil dilation with increasing tone level could not be interpreted to neither support nor refute the rapid growth loudness model by Moore et al (2004) that suggest abnormal growth of loudness starting above 4-10 SL in listeners with hearing loss. Should pupil dilation can directly index perceived loudness of tones near threshold, one of the hypothesized growth pattern of pupil dilation would have supported one model of loudness perception versus the other. Figure 48 shows the hypothesized growth of peak pupil dilation with increasing tone level as described by each model if peak dilation directly reflected perceived loudness.

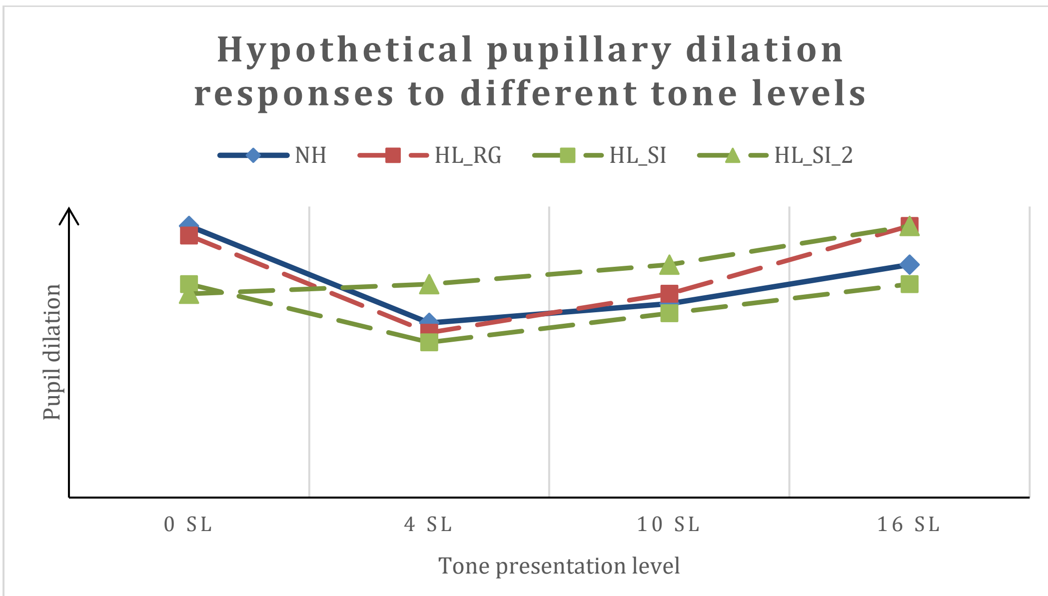


Figure 48: Hypothetical mean pupillary dilation response across the experimental sound levels obtained from NH listeners (solid blue), listeners with HL if their responses do not support SI (dashed red), listeners with HL if their responses support SI (dashed green and dashed-dotted green).

7.1.3 Discussion of results as a contribution to pupillometry literature

In the current study, the detection of sound elicited a pupillary response. To the best of the author's knowledge, this is the first study to record pupillary responses to at threshold sound detection. In general, the presence of pupil dilation response during the sound detection task was only present when tones were perceived (responded yes) and not when tones were present but not perceived (responded no). This is consistent with the results of the two studies that recorded pupil response during visual stimuli detection by Hakerem and Sutton (1966); Privitera et al. (2010).

Average pupil dilation measures obtained in this study were in agreement with measures obtained in other pupillometry studies that collected an auditory-evoked pupil dilation response. Similar to previous studies, peak dilation was measured between 1-2 seconds post-tone onset (Liao, et al, 2011).

In contrast to the results by Dybäck and Wallgren (2016), where participants were passive listeners (not required to report whether they heard the tone or not), all tones perceived in this study elicited a pupil dilation response. In the study by Dyback and Wallgren (2016), although tones were presented above threshold (ranged from 30 to 70 dB SPL presented to listeners with normal hearing), they did not elicit a pupil dilation response possibly because tones were not significant to the listeners (Privitera et al., 2010).

7.2 EXPERIMENT II: SUBJECTIVE RATING EXPERIMENT

7.2.1 Specific Aim 3

Listeners in both groups in the study subjectively rated the loudness of tones at 0 dB SL similarly. The loudness rating did not show a difference in perception shown by the pupil dilation data. This is not surprising given the shortcoming of subjective ratings as a measure. Asking participants to rate the loudness of a stimuli within a numerical scale is also restricting them to choosing the same lowest value in their individual scale every time they experience the lowest sensation, regardless of whether that lowest sensation is actually different levels. This will manifest similar ratings even when the lowest sensations are reached at different levels.

The loudness function obtained from listeners with HL in this study did not exhibit a different (steeper) slope from that obtained from NH listeners. This finding suggests that listeners with HL in this study did not exhibit abnormal growth of loudness above their elevated threshold (rapid growth of loudness did not start at any SL up to 28 SL). Although the loudness function obtained from listeners with HL did not start at a larger loudness at threshold, their similar loudness growth pattern above threshold to the NH group supports the SI model of loudness.

It is worth noting that loudness scaling methods, including magnitude estimation (ME), have been accepted as the gold standard for many advantages, but are still questioned in their ability to uncover the true sensory representation without response bias. Therefore, it is possible that ME loudness data in the present study were not able to produce the true loudness function for the participants. It has been argued that in all subjective loudness scaling functions, by varying the stimulus range, stimulus spacing, and instruction parameters, response biases are introduced that distort the form of loudness function (Schneider & Parker, 1990).

It is not uncommon for loudness functions obtained by magnitude estimation measures to show a lack of loudness recruitment in listeners with hearing loss. Knight and Margolis (1984) found that loudness growth functions in listeners with normal hearing and those with presbycusis (symmetrical hearing loss) obtained by ME measures were not significantly different in steepness; however, steeper functions were obtained from listeners with asymmetrical hearing loss. Actually, it has not always been believed that listeners with presbycusis show signs of loudness recruitment (Békésy & Wever, 1960; Pestalozza & Shore, 1955). This either suggests that ME measure is not optimal to assess loudness growth functions and to detect loudness growth differences, or that listeners with hearing loss do not necessarily show steeper growth of loudness above threshold compared to normal hearing listeners. Most studies that were able to demonstrate a steeper than normal growth of loudness in listeners with hearing loss either utilized loudness matching techniques, where they asked listeners with asymmetrical hearing loss to match the loudness of suprathreshold tones presented to the normal ear to the loudness of another tone presented to the impaired ear (e.g., Moore, 2004), or obtained loudness growth functions from listeners with more severe degree of hearing losses than those in this study (e.g., Hellman & Meiselman, 1993).

In order to examine the loudness growth within the loudness function in a different way, the difference of loudness ratings at two points within the dynamic range of intensity was compared between the groups to assess the possible change in the dynamic range in terms of loudness. The dynamic range of loudness between 4 SL and higher SLs remained similar in both groups with increasing level. Equal distances between ratings of 4 SL and higher SLs (16, 20, 24, and 28 SL) in both groups indicated equal amounts of loudness growth from 4 SL to those higher sensation level tones. This finding also supports the SI loudness model that suggests no change in loudness growth pattern with hearing loss.

8.0 RESEARCH AND CLINICAL IMPLICATIONS

This study enhances our understanding of loudness perception near threshold in listeners with hearing loss, and provides support for the softness imperceptions loudness model. The results of this study are considered preliminary and provide a foundation for replication in a larger sample.

This is the first study that objectively measured the perceived loudness of near threshold auditory stimuli and the first to compare this measure between listeners with NH and HL. The pupillary responses obtained in this study serve as an expansion to the pupillometry literature as they are in agreement with the only other available data at threshold obtained in response to visual stimuli (Hakerem & Sutton, 1966).

This study also demonstrated the value of pupillometry as an objective measure of threshold. Once stimuli were at/near threshold, the properties of signal detection theory of sensitivity index (the component of decision making at threshold) showed prominent in the pupillary response (increased latency). Taking the finding of this study along with others that reported responses to passive tone detection (Dybäck & Wallgren, 2016) and those reported responses to near threshold visual stimuli (Hakerem & Sutton, 1966) could motivate research toward utilizing pupillary responses as an objective measure for hearing threshold.

Findings of the current study demonstrate the importance of considering not only sensory factors, but also the cognitive factors of hearing loss when developing hearing technologies such as hearing aids. The difference in processing load and processing duration for detecting sounds at threshold in listeners with hearing loss feeds into the emerging field of cognitive hearing (Arlinger, Lunner, Lyxell, & Kathleen Pichora-Fuller, 2009). More research is needed to further clarify the underlying causes of this difference beside what was claimed in this study.

9.0 LIMITATIONS AND FUTURE DIRECTIONS

There are several limitations of the present study that need to be mentioned. Inferences made from the results of the current study about the different amount of perceived loudness of tones at threshold are intuitively appropriate (attractive), but they exclude the possibility of an inherent unknown difference in decision-making and metacognitive process that might occur with, or be caused by hearing loss. Despite the fact that the two groups in the study did not have different pre-experiment and in-experiment baseline pupil diameter, they might be different in how they cognitively process in decision making tasks. This possible difference between the groups could possibly be natural difference in decision making between younger and older listeners. The NH group in this study were younger than the HL groups. Although one would expect a longer processing duration for decision making in the older group (HL group), the longer sustained processing was observed in the NH group. The number of trials averaged per participant was less than optimal for such physiologic data. When reaching near threshold levels, and as would be expected, a big number of the trials presented at those levels were not perceived (participant responded "No") and therefore, were excluded from the analysis of responses to "perceived" tones. Several participants in the study contributed data at near threshold level conditions (-4, 0 SL) with a single trial only. Not having any averaging for those subjects might have increased noise in the overall averaged data. A study design that includes at least triple the number of trials per condition would have been optimal to account for excluded trials due to missed stimuli, as well as noisy recordings and excessive blinks.

Future research is needed to investigate factors within the hearing loss that might correlate with the manifestation of larger loudness at threshold. Whether it is amount of threshold elevation (degree of hearing loss), physiological cause of hearing loss, amount of increased internal noise due to hearing loss, this research could go to many directions. Including listeners with wider range of hearing loss severity and those with possible inner hair cell loss might confirm the relationship of the SI model and outer hair cell loss. This also could make it possible to covariate the degree of hearing loss in the analysis of the response to sound.

10.0 CONCLUSION

Pupil dilation is not a direct index for perceived loudness. It is an indirect index of loudness if the difference in perceived loudness results in a difference in the cognitive processing load required for the task, or possibly a difference in arousal level (like sounds that are non-startle inducing and those that are startle inducing). Based on the findings of the current study, loudness of tones at threshold might be perceived louder (if not at least perceptually different) with presence of hearing loss. Loudness growth above threshold in listeners with hearing loss seems to be similar to that in normal hearing listeners. Reconsideration for the definition of loudness recruitment and its underlying assumption of equal loudness at threshold should be further investigated. Acknowledging the existence of the SI model of loudness perception in the literature is appropriate.

APPENDIX A

CASE HISTORY FORM

Participant Number: ______________ Date: ___________________

Sex: Male Female

Dominant hand: Right Left

Age: ______________ Date of birth: ___________________________

Native/First Language: ____

1. Do you have hearing loss? Yes No
   If yes, which ear: Right Left Both

   When did it start? ____

   Cause of hearing loss: ____

2. Have you ever used hearing aids or any amplification devices? Yes No
   If yes, explain: ____

   When did you start using the device? ____

   For how long? ____

3. Do you feel one ear is better than the other? Yes No
   If yes, which ear is better? Right Left

4. Are you in good general health? Yes No
   If no, please explain your medical condition: ____________________________________

5. Have you had any ear surgery? Yes No
   Explain: ____

6. Have you had any recent ear infections, drainage, or pain in your ears? Yes No

If yes, please indicate the date: ____

7. Do you wear eye glasses? Yes No
8. Do you wear contact lenses? Yes No
9. Have you had any eye surgery? Yes No
   Explain: ____

10. Have you ever had any condition that affects your brain such as: a stroke, seizure, hemorrhage, brain tumor, or other type of neurological condition?
    Yes No

11. Have you ever been diagnosed with any psychological disorder such as: anxiety disorder, schizophrenia, severe depression? Yes No
    If yes, please explain: ____

12. Are you currently using any anticholinergic drugs (listed below) in regular basis?
    Yes No

□ Parkinson's medications

□ Diphenhydramine (Benadryl)

□ Trihexyphenidyl (Artane)

□ Benztropine mesylate (Cogentin)

□ Biperiden (Kineton)

□ Antipsychotics

□ Clomipramine (Anafranil)

□ Chlorpromazine (Thorazine)

□ Clozapine (Clozaril)

□ Fluphenazine (Prolixin)

□ Loxapine (Loxitane)

□ Olanzapine (Zyprexa)

□ Perphenazine (Trilafon)

□ Pimozide (Orap)

□ Quetiapine (Seroquel)

□ Thioridazine (Mellaril)

□ Thiothixene (Navane)

□ Trifluoperazine (Stelazine)

APPENDIX B

VERBAL INSTRUCTIONS PROVIDED THROUGHOUT THE SESSION FOR VARIOUS TESTING

Test	Read instructions
Uncomfortable loudness level (UCL)	*"The purpose of this test is to find your uncomfortable loudness level. You will hear tones that increase in volume. I want you to make a judgment about how loud the sounds are using the loudness categories in the sheet. Pretend you are listening to the radio at that volume. How loud would it be? Tell me when the sound reaches a level that is uncomfortably loud (a category of 7). Keep in mind that this level is louder than you would ever want on your radio no matter what mood you were in. When the loudness of the tone reaches that level, say "seven" and I will be able to hear you. When judging the loudness of a tone, it is Ok to skip a category, or to repeat a category. Do you have any questions?"*
	"You will see two boxes in the screen [show with your fingers where the boxes will be located]*, box one and box two. Each time, they will light up one after the other. When each box lights up, it means that there could've been a tone. It is*

	always going to be one tone, either when box one lit up or when box two lit up. Your job is to click on the box where you think you heard the tone. It has nothing to do with which ear you heard the tone, they are always going to be in your right ear. It has to do with when did you hear the tone, this time or this time? [Show times with hand gesture]. *The box that you click will flash in green if your choice was correct and in red if your choice was wrong. You can take your time to decide where you heard the tone, but as soon as you click the box, the next tone will be played, so you really have to keep listening. Sometimes the tone might be too soft that you can barely hear it, just do your best guess. Sometimes you will not hear a tone but you have to click a box to go to the next tone. Do you have any question?"*
Practice trial in sound detection experiment	*"I will insert these earphones in your ears. You will go through two sets of experiments today. However, we will start with a practice. You will use this keypad to give responses. Just follow the instructions that will come up on the screen. They will teach you how to use the keypad and then you will go through some practice trials of the experiment. Basically the instructions will tell you that you will always be looking either at a cross or a square. It will start with a square then will change to a cross. When you see the cross, it means listen carefully, there might be a tone. When you see the block, it means now give me an answer, did you hear the tone or not by pressing the appropriate button in the keypad. You have to wait for the square to give your answer. In the main experiment we want you to keep your head in the same location and your eyes*

	wide open and try not to blink. So start practicing that in the practice trials. Do you have any questions? Once you are done with the practice, I will come and make sure you don't have any questions before we start the main experiment".
	"In this experiment, you will write down your response in the appropriate row in this sheet of paper. You will be looking at the screen but you do not need to use the head/chin rest, and you do not need to keep your eyes open because we will not be recording your pupil. You will go through practice trials first, just follow the instructions in the screen. You will hear some tones and make loudness judgments by assigning a number. Each time, you will hear two consecutive tones, tone one then tone two. You will see in the screen, 1 then 2, corresponding to when you hear those tones. if I tell you that the first tone is always going to be the same tone and the loudness of that tone is a 100, how would rate the loudness of the second tone in reference to the first one? I want you to assign a number to the second tone. This number could be ANY number of ANY scale. Take your time to decide on the number, write it down in the appropriate row in the sheet, then you can click any button in the keypad to go to the next pair of toes. [Give verbal examples with visual cues for how trials will sound and look like]. *"If tone 2 is louder, that would be a number that is bigger than a 100. If it is softer, that would be a number that is smaller than a 100. How much bigger or smaller? That is really subjective and up to your judgement. If you think tone 2 is exactly twice as loud then that would be?* [Prompt participant to say 200], *but if you think that tone 2 was exactly half as*

	loud, then that would be? [Prompt participant to say 50]. *There is no right or wrong. Just try to be precise as much as you can". Are you ready?"*

APPENDIX C

THE LOUDNESS CATEGORIES FROM THE CONTOUR TEST (COX ET AL., 1997)

Loudness Categories

7. Uncomfortably loud
6. Loud, but o.k.
5. Comfortable, but slightly loud
4. Comfortable
3. Comfortable, but slightly soft
2. Soft
1. Very soft

APPENDIX D

UNCOMFORTABLE LOUDNESS LEVEL NORMATIVE DATA

Mean, standard deviation (SD), and number of data points (N), for LDL at four frequencies (in dB HL) measured by Pascoe (1988). Also provided are minimum, maximum, range at each frequency and grand mean, SD, and N. Table obtained from Keller (2006).

	500 Hz			1000 Hz			2000 Hz			4000 Hz					
HL	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Grand Mean	Grand SD	Grand N
0	95.4	9.9	14	101.7	6.8	6	--	--	--	--	--	--	97.3	8.4	20
5	97.7	8.7	22	101.3	6.4	12	--	--	--	--	--	--	99.0	8.1	34
10	99.1	7.3	39	100.3	6.9	15	110.0	--	1	90.0	--	1	99.5	7.3	56
15	97.5	6.8	57	99.1	6.5	32	100.0	--	1	100.0	--	1	98.1	6.7	91
20	95.0	7.8	45	100.0	7.1	26	95.0	2.9	2	91.7	2.9	3	96.6	7.7	76
25	100.4	8.9	40	102.4	8.5	40	100.0	9.6	5	100.0	--	2	101.3	8.6	87
30	102.5	8.4	54	100.8	8.1	40	100.5	9.1	11	106.7	12.6	3	101.8	8.4	108
35	97.9	8.7	57	103.3	7.4	51	103.3	7.3	24	110.0	--	1	101.1	8.4	133
40	100.2	9.7	45	102.1	9.5	63	104.6	9.0	52	105.0	7.1	15	102.6	9.3	175
45	106.6	9.5	32	104.2	8.9	63	105.6	8.9	63	106.3	11.4	19	105.4	9.3	177
50	105.3	9.8	33	108.2	9.2	47	107.7	9.1	84	107.8	10.2	34	107.5	9.4	198
55	106.7	8.5	27	108.4	7.8	44	107.6	8.3	82	107.8	8.9	66	107.7	8.4	219
60	108.2	6.1	17	108.3	6.5	26	110.6	7.1	78	110.3	8.4	67	110.0	7.5	188
65	108.3	7.5	6	112.9	6.8	22	115.2	8.0	30	113.9	9.3	58	113.8	8.5	116
70	103.8	4.8	4	114.3	6.7	7	115.5	8.4	22	115.4	7.7	61	114.8	8.0	94
75	116.3	6.4	8	122.5	10.6	2	116.9	7.2	24	116.6	5.7	57	116.8	6.2	91
80	115.0	5.8	4	120.0	10.0	3	119.7	4.6	16	120.5	6.0	33	119.8	5.9	56
85	125.0	--	1	117.5	4.2	6	120.7	6.7	7	120.2	6.5	22	120.0	6.1	36
90	130.0	--	1	127.5	3.5	2	120.0		1	123.3	5.6	26	123.7	5.6	30
95	120.0	--	1	--	--	--	--	--	--	130.5	7.2	10	129.6	7.6	11
100	--	--	--	--	--	--	--	--	--	126.6	6.3	16	126.6	6.3	16
105	--	--	--	--	--	--	135.0	7.1	2	132.5	4.2	6	133.1	4.6	8
110	--	--	--	135.0	--	1	136.7	2.9	3	131.7	7.6	3	134.3	5.4	7
115	--	--	--	--	--	--	--	--	--	136.7	5.8	3	136.7	5.8	3
120	--	--	--	--	--	--	--	--	--	140.0	--	1	140.0	--	1
GRAND	106.5	7.9	507	109.5	7.4	507	111.8	7.3	503	114.9	7.0	479			2031
Minimum	95			99			95			90					
Maximum	130			135			137			140					
Range	35			36			42			50					

Pascoe (1988) pooled four frequencies mean MCL and UCL (dB HL) as a function of frequency. Below graphs obtained from original article Pascoe (1988).

Figure 4 displays the pooled four-frequency mean MCL and mean UCL in reference to threshold (HTL). The range of plus and minus one Standard Deviation for the above mentioned means is also shown. In general, the slopes of mean MCLs and mean UCLs seem to be different for Hearing Levels within the 0 to 45 dB range than for the 50 to 120 dB range. Mean UCLs stay close to 100 dB HL for thresholds within the 0 to 40 dB HL range. For greater Hearing Levels, the mean UCL increases 5 dB for every 10 dB increase in HL.

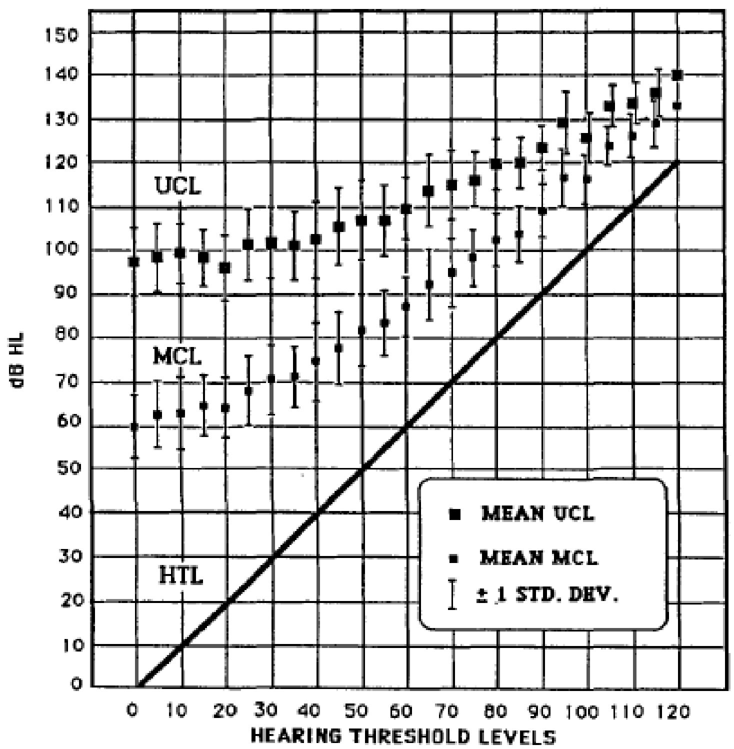


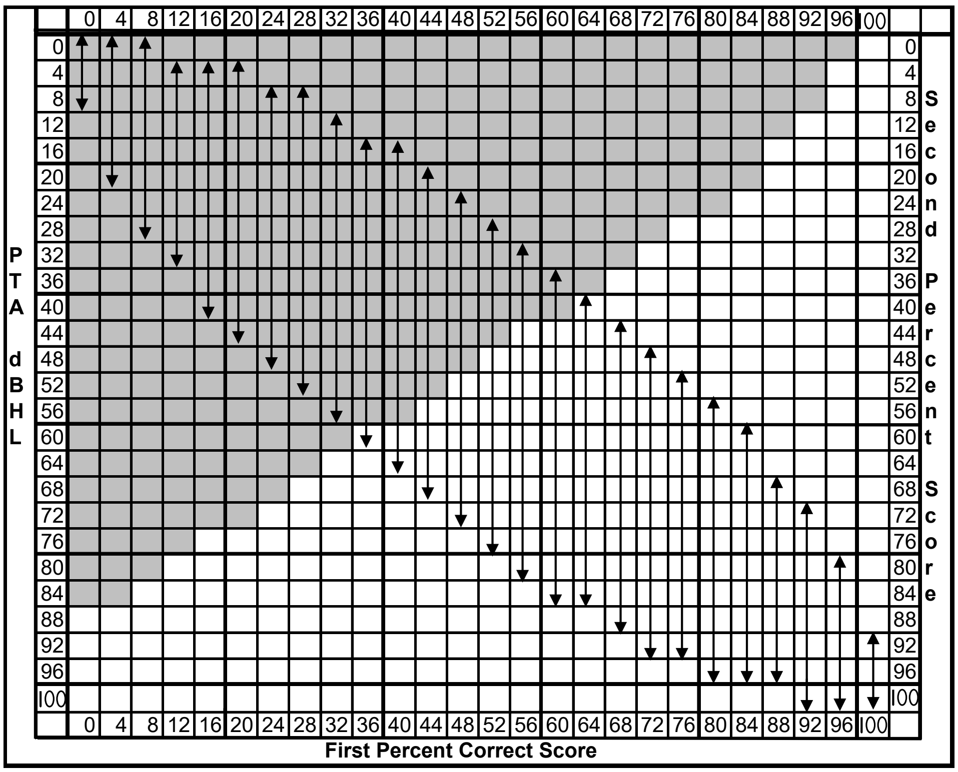
Fig. 4

MEAN COMFORT AND DISCOMFORT LEVELS
RE HEARING THRESHOLD LEVELS

APPENDIX E

WORD RECOGNITION SCORE NORMATIVE DATA

The 95% confidence limit (CL) for PBmax for 25-item NU-6 word lists (Dubno et al., 1995).



95% Confidence Limit for PBmax on NU6 25-word list.Plot score according to PTA on left ordinate and percent correct score on the abscissa. If it falls in the shaded area, it is considered disproportionately low. (Adapted from Dubno et al.,1995)

95% Critical differences for 25-word list. Plot first and second score according to the abscissa and right ordinate. If it falls within the arrow, the two scores are not significantly different (Adapted from Thornton & Raffin, 1978)

BIBLIOGRAPHY

Al-Salim, S. C., Kopun, J. G., Neely, S. T., Jesteadt, W., Stiegemann, B., & Gorga, M. P. (2010). Reliability of categorical loudness scaling and its relation to threshold. *Ear and hearing, 31*(4), 567.

ANSI. (1994). American National Standard Acoustical Terminology. *ANSI Sl*, 1-1994.

ANSI, A. (2004). S3. 6 Specification for Audiometers. *Society of America, New York, NY, USA*.

Antikainen, J., & Niemi, P. (1983a). Neuroticism and the pupillary response to a brief exposure to noise. *Biological psychology, 17*(2), 131-135.

Antikainen, J., & Niemi, P. (1983b). Neuroticism and the pupillary response to a brief exposure to noise. *Biological Psychology, 17*(2-3), 131-135. doi:10.1016/0301-0511(83)90013-3

Arlinger, S., Lunner, T., Lyxell, B., & Kathleen Pichora-Fuller, M. (2009). The emergence of cognitive hearing science. *Scandinavian journal of psychology, 50*(5), 371-384.

Association, A. S.-L.-H. (2005). Guidelines for manual pure-tone threshold audiometry.

Ayasse, N. D., Lash, A., & Wingfield, A. (2017). Effort not speed characterizes comprehension of spoken sentences by older adults with mild hearing impairment. *Frontiers in Aging Neuroscience, 8*, 329.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and language, 59*(4), 390-412.

Bailey, I. L., & Lovie, J. (1980). The design and use of a new near-vision chart. *American Journal of Optometry and Physiological Optics, 57*(6), 378-387.

Bala, A., & Takahashi, T. (2000). Pupillary dilation response as an indicator of auditory discrimination in the barn owl. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 186*(5), 425-434.

Barry, R. J. (2009). Habituation of the orienting reflex and the development of preliminary process theory. *Neurobiology of learning and memory, 92*(2), 235-242.

Beatty, J. (1982a). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol Bull, 91*(2), 276-292.

Beatty, J. (1982b). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin, 91*(2), 276.

Beatty, J., & Lucero-Wagoner, B. (2000a). The Pupillary System. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of Psychophysiology* (2nd ed., pp. 142-162). Cambridge: Cambridge University Press.

Beatty, J., & Lucero-Wagoner, B. (2000b). The pupillary system. *Handbook of psychophysiology, 2*, 142-162.

Beatty, J., & Wagoner, B. L. (1977). Activation and signal detection: A pupillometric analysis In. Los Angeles: University of California, Los Angeles: Human Neurophysiology laboratory.

Beck, D. L. (2015). BrainHearing: Maximizing hearing and listening. *Hearing Review, 3*(20).

Békésy, V., & Wever, E. (1960). *Experiments in hearing* (Vol. 8): McGraw-Hill New York.

Bilecen, D., Seifritz, E., Scheffler, K., Henning, J., & Schulte, A.-C. (2002). Amplitopicity of the human auditory cortex: an fMRI study. *Neuroimage, 17*(2), 710-718.

Blamey, P. J., & Martin, L. F. (2009). Loudness and satisfaction ratings for hearing aid users. *Journal of the American Academy of Audiology, 20*(4), 272-282.

Boettcher, F. A., & Salvi, R. J. (1993). Functional changes in the ventral cochlear nucleus following acute acoustic overstimulation. *The Journal of the Acoustical Society of America, 94*(4), 2123-2134.

Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology, 45*(4), 602-607.

Bradshaw, J. (1967). Pupil size as a measure of arousal during information processing. *Nature, 216*(5114), 515-516.

Brand, T., & Hohmann, V. (2001a). Effect of Hearing Loss, Centre Frequency, and Bandwidth on the Shape of Loudness Functions in Categorical Loudness Scaling. *Audiology, 40*(2), 92-103.

Brand, T., & Hohmann, V. (2001b). Effect of Hearing Loss, Centre Frequency, and Bandwidth on the Shape of Loudness Functions in Categorical Loudness Scaling: Efecto de la hipoacusia, la frecuencia central y el ancho de banda, en la configuración de la funciones de sonoridad en una escala categóries de sonoridad. *Audiology, 40*(2), 92-103.

Brand, T., & Hohmann, V. (2002). An adaptive procedure for categorical loudness scaling. *The Journal of the Acoustical Society of America, 112*(4), 1597-1604.

Buus, S., & Florentine, M. (2001). *Modifications to the power function for loudness*: Citeseer.

Buus, S., & Florentine, M. (2002). Growth of loudness in listeners with cochlear hearing losses: recruitment reconsidered. *J Assoc Res Otolaryngol, 3*(2), 120-139.

Buus, S., Florentine, M., & Poulsen, T. (1997). Temporal integration of loudness, loudness discrimination, and the form of the loudness function. *The Journal of the Acoustical Society of America, 101*(2), 669-680.

Buus, S., Greenbaum, H., & Scharf, B. (1982). Measurements of equal loudness and reaction times. *The Journal of the Acoustical Society of America, 72*(S1), S94-S94.

Buus, S., Müsch, H., & Florentine, M. (1998). On loudness at threshold. *The Journal of the Acoustical Society of America, 104*(1), 399-410. doi:doi:http://dx.doi.org/10.1121/1.423295

Byrne, D., Dillon, H., Ching, T., Katsch, R., & Keidser, G. (2001). NAL-NL1 procedure for fitting nonlinear hearing aids: Characteristics and comparisons with other procedures. *Journal of American Academy of Audiology, 12*(1), 37-51.

Cai, S., Ma, W.-L. D., & Young, E. D. (2009). Encoding intensity in ventral cochlear nucleus following acoustic trauma: implications for loudness recruitment. *Journal of the Association for Research in Otolaryngology, 10*(1), 5-22.

Carney, L. H. (1994). Spatiotemporal encoding of sound level: Models for normal encoding and recruitment of loudness. *Hearing research, 76*(1), 31-44.

Chalupper, J., & Fastl, H. (2002). Dynamic loudness model (DLM) for normal and hearing-impaired listeners. *Acta Acustica united with Acustica, 88*(3), 378-386.

Chen, Z., Hu, G., Glasberg, B., & Moore, B. (2011). A new method of calculating auditory excitation patterns and loudness for steady sounds. *Hearing research, 282*(1), 204-215.

Cheng, T. C., & Wareing, M. J. (2012). Three-year ear, nose, and throat cross-sectional analysis of audiometric protocols for magnetic resonance imaging screening of acoustic tumors. *Otolaryngology--Head and Neck Surgery, 146*(3), 438-447.

Chocholle, R. (1940). Variation des temps de réaction auditifs en fonction de l'intensité à diverses fréquences. *L'année psychologique, 41*(1), 65-124.

Cornelisse, L. E., Seewald, R. C., & Jamieson, D. G. (1995). The input/output formula: A theoretical approach to the fitting of personal amplification devices. *The Journal of the Acoustical Society of America, 97*(3), 1854-1864.

Cox, R. M. (1995). Using loudness data for hearing aid selection: the IHAFF approach. *The Hearing Journal, 48*(2), 10-39.

Cox, R. M., Alexander, G. C., Taylor, I. M., & Gray, G. A. (1997). The contour test of loudness perception. *Ear and hearing, 18*(5), 388-400.

Davidson, S. A., Wall, L. G., & Goodman, C. M. (1990). Preliminary studies on the use of an ABR amplitude projection procedure for hearing aid selection. *Ear and hearing, 11*(5), 332-339.

Davis, H. (1964). Enhancement of evoked cortical potentials in humans related to a task requiring a decision. *Science, 145*(3628), 182-183.

De Cheveigne, A. (2005). Pitch perception models. In *Pitch* (pp. 169-233): Springer.

de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proceedings of the National Academy of Sciences, 111*(5), E618-E625.

Deckers, L., & Hricik, D. (1984). Orienting and humor responses: A synthesis. *Motivation and emotion, 8*(3), 183-204.

Dubno, J. R., Lee, F.-S., Klein, A. J., Matthews, L. J., & Lam, C. F. (1995). Confidence limits for maximum word-recognition scores. *Journal of Speech, Language, and Hearing Research, 38*(2), 490-502.

Dybäck, M., & Wallgren, J. (2016). Pupil dilation as an indicator for auditory signal detection: Towards an objective hearing test based on eye tracking. In.

Einhauser, W., Koch, C., & Carter, O. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in human neuroscience, 4*, 18.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of the National Academy of Sciences, 105*(5), 1704-1709.

Ekman, G., Berglund, B., & Berglund, U. (1966). Loudness as a function of the duration of auditory stimulation. *Scandinavian journal of psychology, 7*(1), 201-208.

Elberling, C. (1999). Loudness scaling revisited. *Journal-American Academy Of Audiology, 10*, 248-260.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature neuroscience, 16*(8), 1146-1153.

Ellis, M. R., & Wynne, M. K. (1999). Measurements of loudness growth in 1/2-octave bands for children and adults with normal hearing. *American Journal of Audiology, 8*(1), 40-46.

Epstein, M., & Florentine, M. (2005a). Inferring basilar-membrane motion from tone-burst otoacoustic emissions and psychoacoustic measurements. *J Acoust Soc Am, 117*(1), 263-274.

Epstein, M., & Florentine, M. (2005b). A test of the equal-loudness-ratio hypothesis using cross-modality matching functions. *J Acoust Soc Am, 118*(2), 907-913.

Epstein, M., & Florentine, M. (2006a). Loudness of brief tones measured by magnitude estimation and loudness matching. *J Acoust Soc Am, 119*(4), 1943-1945.

Epstein, M., & Florentine, M. (2006b). Reaction time to 1- and 4-kHz tones as a function of sensation level in listeners with normal hearing. *Ear Hear, 27*(4), 424-429. doi:10.1097/01.aud.0000224743.50604.6c

Epstein, M., & Florentine, M. (2006c). Reaction time to 1-and 4-khz tones as a function of sensation level in listeners with normal hearing. *Ear and hearing, 27*(4), 424-429.

Epstein, M., & Marozeau, J. (2010). Loudness and intensity coding. *The oxford handbook of auditory science: Hearing, 3*.

Epstein, M., Marozeau, J., & Florentine, M. (2006). Basilar-membrane activity and loudness. *Proceedings of Fechner Day, 22*(1), 181-186.

Evans, E. (1975). The sharpening of cochlear frequency selectivity in the normal and abnormal cochlea. *Audiology, 14*(5-6), 419-442.

Evans, E. (1981). The dynamic range problem: place and time coding at the level of cochlear nerve and nucleus. In *Neuronal mechanisms of hearing* (pp. 69-85): Springer.

Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: facts and models* (Vol. 22). New York;Berlin;: Springer.

Fechner, G. T. (1948). Elements of psychophysics, 1860.

Fletcher, H. (1937). Relation between loudness and masking. *The Journal of the Acoustical Society of America, 9*(1).

Fletcher, H. (1940). Auditory patterns. *Reviews of modern physics, 12*(1), 47.

Fletcher, H., & Munson, W. (1933). Loudness, Its Definition, Measurement and Calculation*. *Bell System Technical Journal, 12*(4), 377-430.

Fletcher, H., & Munson, W. (1937). Relation between loudness and masking. *The Journal of the Acoustical Society of America, 9*(1), 78-78.

Florentine, M. (2011). *Loudness*: Springer.

Florentine, M., Buus, S., & Poulsen, T. (1996). Temporal integration of loudness as a function of level. *The Journal of the Acoustical Society of America, 99*(3), 1633-1644. doi:doi:http://dx.doi.org/10.1121/1.415236

Florentine, M., Buus, S., & Rosenberg, M. (2005). Reaction-time data support the existence of softness imperception in cochlear hearing loss. In *Auditory Signal Processing* (pp. 30-39): Springer.

Florentine, M., & Epstein, M. (2006). To honor stevens and repeal his law (for the auditory ststem). *Proceedings of Fechner Day, 22*(1), 37-42.

Florentine, M., & Zwicker, E. (1979). A model of loudness summation applied to noise-induced hearing loss. *Hearing research, 1*(2), 121-132.

Fowler, E. P. (1937). Measuring the sensation of loudness: a new approach to the physiology of hearing and the functional and differential diagnostic tests. *Archives of Otolaryngology, 26*(5), 514-521.

Fridberger, A., Zheng, J., & Nuttall, A. (2002). Alterations of basilar membrane response phase and velocity after acoustic overstimulation. *Hearing research, 167*(1), 214-222.

Gabay, S., Pertzov, Y., & Henik, A. (2011). Orienting of attention, pupil size, and the norepinephrine system. *Attention, Perception, & Psychophysics, 73*(1), 123-129.

Gabriel, B., Kollmeier, B., & Mellert, V. (1997). Influence of individual listener, measurement room and choice of test-tone levels on the shape of equal-loudness level contours. *Acta Acustica united with Acustica, 83*(4), 670-683.

Gescheider, G. A., & Hughson, B. A. (1991). Stimulus context and absolute magnitude estimation: A study of individual differences. *Perception & Psychophysics, 50*(1), 45-57.

Glasberg, B., & Moore, B. (1986). Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments. *The Journal of the Acoustical Society of America, 79*(4), 1020-1033. doi:doi:http://dx.doi.org/10.1121/1.393374

Glasberg, B., & Moore, B. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing research, 47*(1-2), 103-138.

Glasberg, B., & Moore, B. (2002). A model of loudness applicable to time-varying sounds. *Journal of the Audio Engineering Society, 50*(5), 331-342.

Goldinger, S. D., & Papesh, M. H. (2012). Pupil dilation reflects the creation and retrieval of memories. *Current Directions in Psychological Science, 21*(2), 90-95.

Goldwater, B. (1972). Psychological significance of pupillary movements. *Psychological bulletin, 77*(5), 340.

Granholm, E., Morris, S., Asarnow, R. F., Chock, D., & Jeste, D. V. (2000). Accelerated age-related decline in processing resources in schizophrenia: evidence from pupillary responses recorded during the span of apprehension task. *Journal of the International Neuropsychological Society, 6*(1), 30-43.

Gröschel, M., Ryll, J., Götze, R., Ernst, A., & Basta, D. (2014). Acute and long-term effects of noise exposure on the neuronal spontaneous activity in cochlear nucleus and inferior colliculus brain slices. *BioMed research international, 2014*.

Guthrie, D., & Buchwald, J. S. (1991). Significance testing of difference potentials. *Psychophysiology, 28*(2), 240-244.

Hakerem, G., & Sutton, S. (1966). Pupillary response at visual threshold. *Nature*.

Hall, D. A., Haggard, M. P., Summerfield, A. Q., Akeroyd, M. A., Palmer, A. R., & Bowtell, R. W. (2001). Functional magnetic resonance imaging measurements of sound-level encoding in the absence of background scanner noise. *The Journal of the Acoustical Society of America, 109*(4), 1559-1570.

Harrison, R. (1981). Rate-versus-intensity functions and related AP responses in normal and pathological guinea pig and human cochleas. *The Journal of the Acoustical Society of America, 70*(4), 1036-1044.

Harrison, R., & Evans, E. (1979). Some aspects of temporal coding by single cochlear fibres from regions of cochlear hair cell degeneration in the guinea pig. *Archives of oto-rhino-laryngology, 224*(1-2), 71-78.

Hart, H. C., Palmer, A. R., & Hall, D. A. (2002). Heschl's gyrus is more sensitive to tone level than non-primary auditory cortex. *Hearing research, 171*(1), 177-190.

Heinz, M. G., Issa, J. B., & Young, E. D. (2005). Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment. *Journal of the Association for Research in Otolaryngology, 6*(2), 91-105.

Heinz, M. G., & Young, E. D. (2004). Response growth with sound level in auditory-nerve fibers after noise-induced hearing loss. *Journal of neurophysiology, 91*(2), 784-795.

Hellman, R. P. (1976). Growth of loudness at 1000 and 3000 Hz. *The Journal of the Acoustical Society of America, 60*(3), 672-679.

Hellman, R. P. (1993). Can magnitude scaling reveal the growth of loudness in cochlear impairment. *Sensory research: Multimodal perspectives*, 1-18.

Hellman, R. P. (1999). Cross-modality matching: A tool for measuring loudness in sensorineural impairment. *Ear and hearing, 20*(3), 193.

Hellman, R. P., & Meiselman, C. H. (1988). Prediction of individual loudness exponents from cross-modality matching. *Journal of Speech, Language, and Hearing Research, 31*(4), 605-615.

Hellman, R. P., & Meiselman, C. H. (1990). Loudness relations for individuals and groups in normal and impaired hearing. *The Journal of the Acoustical Society of America, 88*(6), 2596-2606.

Hellman, R. P., & Meiselman, C. H. (1993). Rate of loudness growth for pure tones in normal and impaired hearing. *The Journal of the Acoustical Society of America, 93*(2), 966-975.

Hellman, R. P., & Zwislocki, J. (1961). Some factors affecting the estimation of loudness. *The Journal of the Acoustical Society of America, 33*(5), 687-694.

Hellman, R. P., & Zwislocki, J. (1963). Monaural loudness function at 1000 cps and interaural summation. *The Journal of the Acoustical Society of America, 35*(6), 856-865.

Hershman, R., Henik, A., & Cohen, N. (2018). A novel blink detection method based on pupillometry noise. *Behavior research methods, 50*(1), 107-114.

Humes, L. E., & Ahlstrom, J. B. (1984). Relation between reaction time and loudness. *Journal of Speech, Language, and Hearing Research, 27*(2), 306-310.

Hunter, L. L., Ries, D. T., Schlauch, R. S., Levine, S. C., & Ward, W. D. (1999). Safety and clinical performance of acoustic reflex tests. *Ear and Hearing, 20*(6), 506-513.

Jackson, I., & Sirois, S. (2009). Infant cognition: going full factorial with pupil dilation. *Developmental science, 12*(4), 670-679.

Javel, E., & Mott, J. B. (1988). Physiological and psychophysical correlates of temporal processes in hearing. *Hearing research, 34*(3), 275-294.

Jeng, P. S.-F. (1992). Loudness Predictions Using a Physiologically Based Auditory Model.

Johnson, D. (1980). The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *The Journal of the Acoustical Society of America, 68*(4), 1115-1122.

Johnson, J., & Cox, R. (2013). *Is normal loudness the appropriate goal for hearing aid fittings?* Paper presented at the Presented at the annual conference of the American Auditory Society, Scottsdale, AZ.

Joris, P. X. (2009). Recruitment of Neurons and Loudness: Commentary on “Encoding intensity in ventral cochlear nucleus following acoustic trauma: implications for loudness recruitment” by Cai et al. J. Assoc. Res. Otolaryngol. DOI: 10.1007/s10162-008-0142-y. *JARO: Journal of the Association for Research in Otolaryngology, 10*(1), 1.

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

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science, 157*(3785), 218-219.

Keidser, G., Dillon, H., Carter, L., & O’Brien, A. (2012). NAL-NL2 empirical adjustments. *Trends in amplification*, 1084713812468511.

Keidser, G., Seymour, J., Dillon, H., Grant, F., & Byrne, D. (1999). An efficient, adaptive method of measuring loudness growth functions. *Scandinavian audiology, 28*(1), 3-14.

Khahnemann, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Attention, Perception, & Psychophysics, 2*(3), 101-105.

Kiang, N., Moxon, E., & Levine, R. (1970). *Auditory-Nerve Activity in Cats with Normal and Abnormal Cochleas.* Paper presented at the Ciba Foundation Symposium-Sensorineural Hearing Loss.

Killion, M. C., & Fikret-Pasa, S. (1993). The 3 types of sensorineural hearing loss: Loudness and intelligibility considerations. *Hearing journal, 46*, 31-31.

Knight, K. K., & Margolis, R. H. (1984). Magnitude estimation of loudness II: Loudness perception in presbycusic listeners. *Journal of Speech, Language, and Hearing Research, 27*(1), 28-32.

Koelewijn, T., de Kluiver, H., Shinn-Cunningham, B. G., Zekveld, A. A., & Kramer, S. E. (2015). The pupil response reveals increased listening effort when it is difficult to focus attention. *Hearing research, 323*, 81-90.

Kohfeld, D. L., Santee, J. L., & Wallace, N. D. (1981). Loudness and reaction time: I. *Perception & Psychophysics, 29*(6), 535-549.

Kuchinsky, S. E., Ahlstrom, J. B., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2014). Speech-perception training for older adults with hearing loss impacts word recognition and effort. *Psychophysiology, 51*(10), 1046-1057.

Kuchinsky, S. E., Ahlstrom, J. B., Vaden, J. K. I., Cute, S. L., Humes, L. E., Dubno, J. R., & Eckert, M. A. (2013). Pupil size varies with word listening and response selection difficulty in older adults with hearing loss. *Psychophysiology, 50*(1), 23-34. doi:10.1111/j.1469-8986.2012.01477.x

Langers, D. R., van Dijk, P., Schoenmaker, E. S., & Backes, W. H. (2007). fMRI activation in relation to sound intensity and loudness. *Neuroimage, 35*(2), 709-718.

Launer, S. (1995). Loudness perception in listeners with sensorineural hearing impairment. *Unpublished Ph. D. thesis, Oldenburg University, Oldenburg, Germany*.

Launer, S., Hohmann, V., & Kollmeier, B. (1995). Modeling loudness growth and loudness summation in hearing-impaired listeners. *The Journal of the Acoustical Society of America, 97*(5), 3277-3277.

Launer, S., Holube, I., Hohmann, V., & Kollmeier, B. (1996). Categorical loudness scaling in hearing-impaired listeners-Can loudness growth bei predicted from the audiogram? *Audiological Acoustics, 35*, 156-163.

Lavín, C., San Martín, R., & Jubal, E. R. (2013). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience, 7*.

Lemercier, A., Guillot, G., Courcoux, P., Garrel, C., Baccino, T., & Schlich, P. (2014). Pupillometry of taste: Methodological guide–from acquisition to data processing-and toolbox for MATLAB. *The Quantitative Methods for Psychology, 10*(2), 179-195.

Lempert, K. M., Chen, Y. L., & Fleming, S. M. (2015). Relating pupil dilation and metacognitive confidence during auditory decision-making. *PLoS One, 10*(5), e0126588.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical society of America, 49*(2B), 467-477.

Liao, H.-I., Kidani, S., Yoneya, M., Kashino, M., & Furukawa, S. (2016a). Correspondences among pupillary dilation response, subjective salience of sounds, and loudness. *Psychonomic bulletin & review, 23*(2), 412-425.

Liao, H., Yoneya, M., Kidani, S., Kashino, M., & Furukawa, S. (2016b). Human pupillary dilation response to deviant auditory stimuli: effects of stimulus properties and voluntary attention. *Frontiers in neuroscience, 10*.

Liberman, M. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *The Journal of the Acoustical Society of America, 63*(2), 442-455.

Liberman, M. (1982). Single-neuron labeling in the cat auditory nerve. *Science, 216*(4551), 1239-1241.

Liberman, M., Dodds, L., & Learson, D. (1986). Structure-function correlation in noise-damaged ears: A light and electron-microscopic study. In *Basic and applied aspects of noise-induced hearing loss* (pp. 163-177): Springer.

Liberman, M., & Kiang, N. (1984). Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate-and phase-level functions. *Hearing research, 16*(1), 75-90.

Lin, F. R., Niparko, J. K., & Ferrucci, L. (2011). Hearing loss prevalence in the United States. *Archives of internal medicine, 171*(20), 1851-1853.

Logue, A. (1976). Individual differences in magnitude estimation of loudness. *Perception & Psychophysics, 19*(3), 279-280.

Marozeau, J., Epstein, M., Florentine, M., & Daley, B. (2006). A test of the Binaural Equal-Loudness-Ratio hypothesis for tonesa). *The Journal of the Acoustical Society of America, 120*(6), 3870-3877.

Marozeau, J., & Florentine, M. (2007). Loudness growth in individual listeners with hearing losses: a review. *J Acoust Soc Am, 122*(3), EL81. doi:10.1121/1.2761924

Marshall, L., & Brandt, J. F. (1980). The relationship between loudness and reaction time in normal hearing listeners. *Acta oto-laryngologica, 90*(1-6), 244-249.

Mathôt, S. (2013). A simple way to reconstruct pupil size during eye blinks. *Retrieved from doi, 10*, m9.

May, B. J., Little, N., & Saylor, S. (2009). Loudness perception in the domestic cat: reaction time estimates of equal loudness contours and recruitment effects. *Journal of the Association for Research in Otolaryngology, 10*(2), 295-308.

Meddis, R., Lecluyse, W., Clark, N. R., Jürgens, T., Tan, C. M., Panda, M. R., & Brown, G. J. (2013). A computer model of the auditory periphery and its application to the study of hearing. In *Basic Aspects of Hearing* (pp. 11-20): Springer.

Merriam-Webster. (2004). *Merriam-Webster's collegiate dictionary*: Merriam-Webster.

Micheyl, C., Schrater, P. R., & Oxenham, A. J. (2013). Auditory frequency and intensity discrimination explained using a cortical population rate code. *PLoS Comput Biol, 9*(11), e1003336.

Middleton, J., Kiritani, T., Pedersen, C., Turner, J., Shepherd, G., & Tzounopoulos, T. (2011). Mice with behavioral evidence of tinnitus exhibit dorsal cochlear nucleus hyperactivity because of decreased GABAergic inhibition. *Proceedings of the National Academy of Sciences, 108*(18), 7601-7606.

Modh, D., Katarkar, A., Alam, N., Jain, A., & Shah, P. (2014). Relation of distortion product otoacoustic emission and tinnitus in normal hearing patients: a pilot study. *Noise and Health, 16*(69), 69.

Moore, B. (1995). *Perceptual consequences of cochlear damage*: Oxford University Press.

Moore, B. (1997). An introduction to the psychology of hearing (Academic, San Diego). *and, 313*, 159-167.

Moore, B. (2000). Dead regions in the cochlea: Diagnosis, perceptual consequences, and implications for the fitting of hearing aids. *Trends in amplification, 5*.

Moore, B. (2001). Chapter Thirteen Loudness, Pitch and Timbre.

Moore, B. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology & Neurotology, 24*(2), 243-254.

Moore, B. (2004). Testing the concept of softness imperception: loudness near threshold for hearing-impaired ears. *The Journal of the Acoustical Society of America, 115*(6), 3103-3111.

Moore, B. (2012). *An Introduction to the Psychology of Hearing* (6th ed.). Bingley: Emerald.

Moore, B., & Glasberg, B. (1983). Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America, 74*(3), 750-753.

Moore, B., & Glasberg, B. (1996). A revision of Zwicker's loudness model. *Acta Acustica united with Acustica, 82*(2), 335-345.

Moore, B., & Glasberg, B. (1997). A model of loudness perception applied to cochlear hearing loss. *Auditory Neuroscience, 3*(3), 289-311.

Moore, B., & Glasberg, B. (2004). A revised model of loudness perception applied to cochlear hearing loss. *Hear Res, 188*(1-2), 70-88. doi:10.1016/s0378-5955(03)00347-2

Moore, B., Glasberg, B., & Baer, T. (1997). A model for the prediction of thresholds, loudness, and partial loudness. *Journal of the Audio Engineering Society, 45*(4), 224-240.

Moore, B., Glasberg, B., Hess, R., & Birchall, J. (1985). Effects of flanking noise bands on the rate of growth of loudness of tones in normal and recruiting ears. *The Journal of the Acoustical Society of America, 77*(4), 1505-1513.

Moore, B., Vickers, D., Plack, C., & Oxenham, A. (1999). Inter-relationship between different psychoacoustic measures assumed to be related to the cochlear active mechanism. *The Journal of the Acoustical Society of America, 106*(5), 2761-2778.

Moore, B. C. (2007). *Cochlear hearing loss: physiological, psychological and technical issues*: John Wiley & Sons.

Morita, T., Naito, Y., Nagamine, T., Fujiki, N., Shibasaki, H., & Ito, J. (2003). Enhanced activation of the auditory cortex in patients with inner-ear hearing impairment: a magnetoencephalographic study. *Clinical neurophysiology, 114*(5), 851-859.

Mueller, H. G., & Powers, T. A. (2001). *Consideration of auditory acclimatization in the prescriptive fitting of hearing aids.* Paper presented at the Seminars in Hearing.

Mulders, W., Ding, D., Salvi, R., & Robertson, D. (2011). Relationship between auditory thresholds, central spontaneous activity, and hair cell loss after acoustic trauma. *Journal of Comparative Neurology, 519*(13), 2637-2647.

Mulla, R., Palmer, C., Wassel, N., & Lao, A. (2013). *Relationship between measured and perceived loudness perception*. Poster presented at AudiologyNOW, American Academy of Audiology annual meeting. Anaheim, CA.

Naber, M., Frässle, S., Rutishauser, U., & Einhäuser, W. (2013). Pupil size signals novelty and predicts later retrieval success for declarative memories of natural scenes. *Journal of vision, 13*(2), 11-11.

Narayan, S. S., Temchin, A. N., Recio, A., & Ruggero, M. A. (1998). Frequency tuning of basilar membrane and auditory nerve fibers in the same cochleae. *Science, 282*(5395), 1882-1884.

Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature neuroscience, 15*(7), 1040.

Neely, S. T., Rodriguez, J., Liu, Y.-W., Jesteadt, W., & Gorga, M. P. A computational model of loudness density.

Netser, S., Ohayon, S., & Gutfreund, Y. (2010). Multiple manifestations of microstimulation in the optic tectum: eye movements, pupil dilations, and sensory priming. *Journal of neurophysiology, 104*(1), 108-118.

Nunnally, J. C., Knott, P. D., Duchnowski, A., & Parker, R. (1967). Pupillary response as a general measure of activation. *Attention, Perception, & Psychophysics, 2*(4), 149-155.

Ohlenforst, B., Zekveld, A. A., Lunner, T., Wendt, D., Naylor, G., Wang, Y., . . . Kramer, S. E. (2017). Impact of stimulus-related factors and hearing impairment on listening effort as indicated by pupil dilation. *Hearing research, 351*, 68-79.

Oxenham, A. J., & Plack, C. J. (1997). A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing. *The Journal of the Acoustical Society of America, 101*(6), 3666-3675.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International journal of human-computer studies, 59*(1), 185-198.

Pascoe, D. P. (1988). Clinical measurements of the auditory dynamic range and their relation to formulas for hearing aid gain. *Hearing aid fitting: Theoretical and practical views*, 129-152.

Patterson, R. D., & Holdsworth, J. (1996). A functional model of neural activity patterns and auditory images. *Advances in speech, hearing and language processing, 3*(Part B), 547-563.

Pestalozza, G., & Shore, I. (1955). Clinical evaluation of presbycusis on the basis of different tests of auditory function. *The Laryngoscope, 65*(12), 1136-1163.

Pfingst, B. E., Hienz, R., Kimm, J., & Miller, J. (1975). Reaction− time procedure for measurement of hearing. I. Suprathreshold functions. *The Journal of the Acoustical Society of America, 57*(2), 421-430.

Phillips, D. (1987). Stimulus intensity and loudness recruitment: neural correlates. *The Journal of the Acoustical Society of America, 82*(1), 1-12.

Pickles, J. O. (2012). *An introduction to the physiology of hearing* (Vol. Fourthition.). Bingley, United Kingdom: Emerald Group Publishing, Limited.

Piquado, T., Isaacowitz, D., & Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology, 47*(3), 560-569.

Plack, C. J. (2004). auditory perception. In M. W. E. (Ed.) (Ed.), *Psychology: An international perspective* (Vol.

, pp. 1-20). london: Psychology Press.

Poulton, E. C. (1989). *Bias in quantifying judgements*: Taylor & Francis.

Pratt, H., & Sohmer, H. (1977). Correlations between psychophysical magnitude estimates and simultaneously obtained auditory nerve, brain stem and cortical responses to click stimuli in man. *Electroencephalography and clinical neurophysiology, 43*(6), 802-812.

Preuschoff, K., Marius't Hart, B., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in neuroscience, 5*.

Privitera, C. M., Renninger, L. W., Carney, T., Klein, S., & Aguilar, M. (2010). Pupil dilation during visual target detection. *Journal of Vision, 10*(10), 3-3.

Rajkumar, S., Muttan, S., Jaya, V., & Vignesh, S. (2013). Comparative analysis of different prescriptive formulae used in the evaluation of real ear insertion gain for digital hearing aids. *Universal Journal of Biomedical Engineering, 1*(2), 32-41.

Raney, G. E., Campbell, S. J., & Bovee, J. C. (2014). Using eye movements to evaluate the cognitive processes involved in text comprehension. *Journal of visualized experiments: JoVE*(83).

Rasetshwane, D. M., Neely, S. T., Kopun, J. G., & Gorga, M. P. (2013). Relation of distortion-product otoacoustic emission input-output functions to loudness. *The Journal of the Acoustical Society of America, 134*(1), 369-383.

Relkin, E. M., & Doucet, J. R. (1997). Is loudness simply proportional to the auditory nerve spike count? *The Journal of the Acoustical Society of America, 101*(5), 2735-2740.

Rennies, J., Verhey, J. L., & Fastl, H. (2010). Comparison of loudness models for time-varying sounds. *Acta Acustica united with Acustica, 96*(2), 383-396.

Rhode, W. S. (1971). Observations of the vibration of the basilar membrane in squirrel monkeys using the Mössbauer technique. *The Journal of the Acoustical Society of America, 49*(4B), 1218-1231.

Röhl, M., & Uppenkamp, S. (2012). Neural coding of sound intensity and loudness in the human auditory system. *Journal of the Association for Research in Otolaryngology, 13*(3), 369-379.

Rose, J. E., Brugge, J. F., Anderson, D. J., & Hind, J. E. (1967). Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *Journal of neurophysiology, 30*(4), 769-793.

Rose, J. E., Hind, J. E., Anderson, D. J., & Brugge, J. F. (1971). Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey. *J. Neurophysiol, 34*(4), 685-699.

Rubio, M. E. (2006). Redistribution of synaptic AMPA receptors at glutamatergic synapses in the dorsal cochlear nucleus as an early response to cochlear ablation in rats. *Hearing research, 216*, 154-167.

Ruggero, M. A. (1992). Physiology and coding of sound in the auditory nerve. In *The mammalian auditory pathway: Neurophysiology* (pp. 34-93): Springer.

Ruggero, M. A., & Rich, N. C. (1991). Furosemide alters organ of Corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane. *The Journal of neuroscience, 11*(4), 1057-1067.

Ruggero, M. A., Rich, N. C., Recio, A., Narayan, S. S., & Robles, L. (1997). Basilar-membrane responses to tones at the base of the chinchilla cochlea. *The Journal of the Acoustical Society of America, 101*(4), 2151-2163.

Sachs, M. B., & Abbas, P. J. (1974). Rate versus level functions for auditory-nerve fibers in cats: tone-burst stimuli. *The Journal of the Acoustical Society of America, 56*(6), 1835-1847.

Salvi, R. J., Wang, J., & Ding, D. (2000). Auditory plasticity and hyperactivity following cochlear damage. *Hearing research, 147*(1), 261-274.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature reviews neuroscience, 10*(3), 211-223.

Scharf, B. (1959). Loudness of complex sounds as a function of the number of components. *The Journal of the Acoustical Society of America, 31*(6), 783-785.

Scharf, B. (1978). Loudness. *Handbook of perception, 4*, 187-242.

Scharf, B., & Hellman, R. P. (1966). Model of loudness summation applied to impaired ears. *The Journal of the Acoustical Society of America, 40*(1), 71-78.

Scheuren, J. (2014). ISO 532–Living and working with alternative loudness standards. *Proceedings of InterNoise 2014*.

Schlauch, R. S., DiGiovanni, J. J., & Ries, D. T. (1998). Basilar membrane nonlinearity and loudness. *The Journal of the Acoustical Society of America, 103*(4), 2010-2020.

Schlemmer, K. B., Kulke, F., Kuchinke, L., & Van Der Meer, E. (2005). Absolute pitch and pupillary response: Effects of timbre and key color. *Psychophysiology, 42*(4), 465-472.

Schneider, B., & Parker, S. (1990). Does stimulus context affect loudness or only loudness judgments? *Perception & Psychophysics, 48*(5), 409-418.

Schreiner, C. E., & Malone, B. J. (2015). Representation of loudness in the auditory cortex. *The Human Auditory System: Fundamental Organization and Clinical Disorders, 129*, 73.

Scollie, S., Seewald, R., Cornelisse, L., Moodie, S., Bagatto, M., Laurnagaray, D., . . . Pumford, J. (2005). The desired sensation level multistage input/output algorithm. *Trends in amplification, 9*(4), 159-197.

Sellick, P., Patuzzi, R., & Johnstone, B. (1982). Measurement of basilar membrane motion in the guinea pig using the Mössbauer technique. *The journal of the acoustical society of America, 72*(1), 131-141.

Serpanos, Y. C., O'Malley, H., & Gravel, J. S. (1997). The relationship between loudness intensity functions and the click-ABR wave V latency. *Ear and hearing, 18*(5), 409-419.

Shi, L.-F., Doherty, K. A., & Zwislocki, J. J. (2007). Aided loudness growth and satisfaction with everyday loudness perception in compression hearing aid users. *Journal of the American Academy of Audiology, 18*(3), 206-219.

Shumway, R. H., & Stoffer, D. S. (2011). Time series regression and exploratory data analysis. In *Time series analysis and its applications* (pp. 47-82): Springer.

Siegle, G. J., Granholm, E., Ingram, R. E., & Matt, G. E. (2001). Pupillary and reaction time measures of sustained processing of negative information in depression. *Biological psychiatry, 49*(7), 624-636.

Siegle, G. J., Ichikawa, N., & Steinhauer, S. (2008). Blink before and after you think: blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology, 45*(5), 679-687.

Siegle, G. J., Steinhauer, S. R., Carter, C. S., Ramel, W., & Thase, M. E. (2003a). Do the seconds turn into hours? Relationships between sustained pupil dilation in response to emotional information and self-reported rumination. *Cognitive Therapy and Research, 27*(3), 365-382.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003b). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage, 20*(1), 114-124.

Sigalovsky, I. S., & Melcher, J. R. (2006). Effects of sound level on fMRI activation in human brainstem, thalamic and cortical centers. *Hearing research, 215*(1), 67-76.

Silk, J. S., Siegle, G. J., Whalen, D. J., Ostapenko, L. J., Ladouceur, C. D., & Dahl, R. E. (2009). Pubertal changes in emotional information processing: Pupillary, behavioral, and subjective evidence during emotional word identification. *Development and psychopathology, 21*(1), 7-26.

Silva, I., & Epstein, M. (2010). Estimating loudness growth from tone-burst evoked responses. *The Journal of the Acoustical Society of America, 127*(6), 3629-3642.

Silva, I., & Epstein, M. (2012). Objective estimation of loudness growth in hearing-impaired listeners. *The Journal of the Acoustical Society of America, 131*(1), 353-362.

Silva, I., & Florentine, M. (2006). Effect of adaptive psychophysical procedure on loudness matchesa). *The Journal of the Acoustical Society of America, 120*(4), 2124-2131.

Small, A. M., Brandt, J. F., & Cox, P. G. (1962). Loudness as a function of signal duration. *The Journal of the Acoustical Society of America, 34*(4), 513-514.

Stebbins, W. C. (1966). Auditory reaction time and the derivation of equal loudness contours for the monkey1. *Journal of the experimental analysis of behavior, 9*(2), 135-142.

Stebbins, W. C., Hawkins, J. E., Johnsson, L.-G., & Moody, D. B. (1979). Hearing thresholds with outer and inner hair cell loss. *American journal of otolaryngology, 1*(1), 15-27.

Steinhauer, S., Hakerem, G., & Spring, B. (1979). The pupillary response as a potential indicator of vulnerability to schizophrenia. *Psychopharmacology bulletin, 15*(1), 44-45.

Steinhauer, S., & Zubin, J. (1982). Vulnerability to schizophrenia: Information processing in the pupil and event-related potential. *Biological markers in psychiatry and neurology*, 371-385.

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International journal of psychophysiology, 52*(1), 77-86.

Stern, R. M., & Morgan, N. (2012). Features based on auditory physiology and perception. *Techniques for Noise Robustness in Automatic Speech Recognition*, 193227.

Stevens, J., & Hall, J. (1966). Brightness and loudness as functions of stimulus duration. *Perception & Psychophysics, 1*(9), 319-327.

Stevens, S. (1936). A scale for the measurement of a psychological magnitude: loudness. *Psychological Review, 43*(5), 405.

Stevens, S. (1955). The measurement of loudness. *The Journal of the Acoustical Society of America, 27*(5), 815-829.

Stevens, S. (1956). The direct estimation of sensory magnitudes: Loudness. *The American journal of psychology, 69*(1), 1-25.

Stevens, S. (1957). On the psychophysical law. *Psychological review, 64*(3), 153.

Stevens, S. (1959). Cross-modality validation of subjective scales for loudness, vibration, and electric shock. *Journal of Experimental Psychology, 57*(4), 201.

Stevens, S. (1961). To honor Fechner and repeal his law, A power function, not a log function, describes the operating characteristic of a sensory system. *American Association for the Advancement of Science, 133*(3446), 6.

Sun, W., Manohar, S., Jayaram, A., Kumaraguru, A., Fu, Q., Li, J., & Allman, B. (2011). Early age conductive hearing loss causes audiogenic seizure and hyperacusis behavior. *Hearing research, 282*(1), 178-183.

Sutton, S., Braren, M., Zubin, J., & John, E. (1965). Evoked-potential correlates of stimulus uncertainty. *Science, 150*(3700), 1187-1188.

Syka, J., & Rybalko, N. (2000). Threshold shifts and enhancement of cortical evoked responses after noise exposure in rats. *Hearing research, 139*(1), 59-68.

Thornton, A., Yardley, L., & Farrell, G. (1987). The objective estimation of loudness discomfort level using auditory brainstem evoked responses. *Scandinavian audiology, 16*(4), 219-225.

Thorson, M. J., Kopun, J. G., Neely, S. T., Tan, H., & Gorga, M. P. (2012). Reliability of distortion-product otoacoustic emissions and their relation to loudness. *The Journal of the Acoustical Society of America, 131*(2), 1282-1295.

Tryon, W. W. (1975). Pupillometry: A survey of sources of variation. *Psychophysiology, 12*(1), 90-93.

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, & Behavioral Neuroscience, 16*(4), 601-615.

Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature communications, 8*, 14637.

Van Der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., . . . Wartenburger, I. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology, 47*(1), 158-169.

Van Gerven, P. W., Paas, F., Van Merriënboer, J. J., & Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology, 41*(2), 167-174.

Vogler, D., Robertson, D., & Mulders, W. (2011). Hyperactivity in the ventral cochlear nucleus after cochlear trauma. *The Journal of Neuroscience, 31*(18), 6639-6645.

Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. *Journal of experimental psychology: Applied, 8*(4), 211.

Wang, C.-A., Boehnke, S. E., Itti, L., & Munoz, D. P. (2014). Transient pupil response is modulated by contrast-based saliency. *Journal of Neuroscience, 34*(2), 408-417.

Wang, C., Boehnke, S., White, B., & Munoz, D. (2012). Microstimulation of the monkey superior colliculus induces pupil dilation without evoking saccades. *Journal of Neuroscience, 32*(11), 3629-3636.

Wang, C. A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience, 40*(5), 2822-2832.

Wang, F., Zuo, L., Hong, B., Han, D., Range, E., Zhao, L., . . . Liu, L. (2013). Tonotopic reorganization and spontaneous firing in inferior colliculus during both short and long recovery periods after noise overexposure. *Journal of biomedical science, 20*(1), 1.

Whilby, S., Florentine, M., Wagner, E., & Marozeau, J. (2006a). Monaural and binaural loudness of 5- and 200-ms tones in normal and impaired hearing. *J Acoust Soc Am, 119*(6), 3931-3939.

Whilby, S., Florentine, M., Wagner, E., & Marozeau, J. (2006b). Monaural and binaural loudness of 5-and 200-ms tones in normal and impaired hearinga). *The Journal of the Acoustical Society of America, 119*(6), 3931-3939.

Whiting, B., Moiseff, A., & Rubio, M. E. (2009). Cochlear nucleus neurons redistribute synaptic AMPA and glycine receptors in response to monaural conductive hearing loss. *Neuroscience, 163*(4), 1264-1276.

Wiemer, J., Mühlberger, A, Pauli, P. (2014). Illusory correlations between neutral and aversive stimuli can be induced by outcome aversiveness. *Cognition & emotion, 28*(2), 193-207.

Wilson, K. G., & Stelmack, R. M. (1982). Power functions of loudness magnitude estimations and auditory brainstem evoked responses. *Perception & psychophysics, 31*(6), 561-565.

Withnell, R. H., & Yates, G. K. (1998). Onset of basilar membrane non-linearity reflected in cubic distortion tone input-output functions. *Hearing research, 123*(1), 87-96.

Woolf, N. K., Ryan, A. F., & Bone, R. C. (1981). Neural phase-locking properties in the absence of cochlear outer hair cells. *Hearing research, 4*(3), 335-346.

Xu, H., Kotak, V. C., & Sanes, D. H. (2007). Conductive hearing loss disrupts synaptic and spike adaptation in developing auditory cortex. *The Journal of Neuroscience, 27*(35), 9417-9426.

Yates, G. K. (1990). Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions. *Hear Res, 50*(1-2), 145-162.

Yates, G. K., Winter, I. M., & Robertson, D. (1990). Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range. *Hear Res, 45*(3), 203-219.

Zapala, D. A., Criter, R. E., Bogle, J. M., Lundy, L. B., Cevette, M. J., & Bauch, C. D. (2012). Pure-tone hearing asymmetry: a logistic approach modeling age, sex, and noise exposure history. *Journal of the American Academy of Audiology, 23*(7), 553-570.

Zekveld, A. A., & Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: Insights from pupillometry. *Psychophysiology, 51*(3), 277-284. doi:10.1111/psyp.12151

Zekveld, A. A., Kramer, S. E., & Festen, J. M. (2011). Cognitive load during speech perception in noise: The influence of age, hearing loss, and cognition on the pupil response. *Ear and hearing, 32*(4), 498-510.

Zeng, F.-G., & Turner, C. W. (1991). Binaural loudness matches in unilaterally impaired listeners. *The Quarterly Journal of Experimental Psychology, 43*(3), 565-583.

Zhang, M., Palmer, C., & Mueller, H.G. (2012). *Gain learning hearing aids: does timing matter?* Paper presented at the Poster at the annual meeting of the American Auditory Society, , Scottsdale, AZ.

Zhang, M., & Zwislocki, J. (1992). *Hair cell magnitude and phase transfer functions: effects of overstimulation and their correlation with loudness functions.* Paper presented at the ARO Abstracts.

Zwicker, E. (1956). Die elementaren Grundlagen zur Bestimmung der Informationskapazität des Gehörs. *Acta Acustica united with Acustica, 6*(4), 365-381.

Zwicker, E. (1958). Über psychologische und methodische Grundlagen der Lautheit. *Acta Acustica united with Acustica, 8*(Supplement 1), 237-258.

Zwicker, E. (1961). Subdivision of the audible frequency range into critical bands (Frequenzgruppen). *The Journal of the Acoustical Society of America, 33*(2), 248-248.

Zwicker, E. (1977). Procedure for calculating loudness of temporally variable sounds. *The Journal of the Acoustical Society of America, 62*(3), 675-682. doi:doi:http://dx.doi.org/10.1121/1.381580

Zwicker, E. (1984). Dependence of post-masking on masker duration and its relation to temporal effects in loudness. *The Journal of the Acoustical Society of America, 75*(1), 219-223.

Zwicker, E., & Fastl, H. (2013). *Psychoacoustics: Facts and models* (Vol. 22): Springer Science & Business Media.

Zwicker, E., Flottorp, G., & Stevens, S. (1957). Critical band width in loudness summation. *The Journal of the Acoustical Society of America, 29*(5), 548-557.

Zwicker, E., & Scharf, B. (1965). A model of loudness summation. *Psychological review, 72*(1), 3.

Zwislocki, J. J. (1965). Analysis of some auditory characteristics. In R. R. B. R. D. Luce & a. E. Galanter (Eds.), *Handbook of Mathematical Psychology* (Vol. 3). Wiley, New York!,.