TEMPORAL DYNAMICS OF PSYCHOACOUSTIC AND PHYSIOLOGICAL MEASURES OF COCHLEAR GAIN REDUCTION

by

William Bryan Salloom

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Speech, Language, and Hearing Sciences

West Lafayette, Indiana

May 2022

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. Elizabeth A. Strickland, Chair

Department of Speech, Language, and Hearing Sciences

Dr. Hari Bharadwaj

Department of Speech, Language, and Hearing Sciences

Dr. Michael Heinz

Department of Speech, Language, and Hearing Sciences

Dr. Jeffrey Lucas

Department of Biological Sciences

Approved by:

Dr. Jason R. Cannon

and

Dr. Preeti M. Sivasankar

To my father, Samir.

The most resilient person I know,

who fiercely loves and sacrificed so much for our family.

To my beautiful and loving wife Anne,

and her mom Tracy,

who have supported me in innumerable ways.

They have brought me so much happiness and joy to my life.

To my late grandparents, Arthur Roy and Bernice,

who provided care and stability to my early life,

especially in times of turbulence and difficulty.

ACKNOWLEDGMENTS

The quote, “It takes a village…” is certainly true as it pertains to where I am today, and completing my dissertation and doctoral degree. There are so many individuals that deserve recognition and thanks to help me along my journey. A promise I have made to myself is that I will try to mentor, teach, and help others in the way that I felt that I was so fortunate to have experienced. I have come to realize that there have been many instances in my life when someone took a chance on me and provided with opportunity, probably when they didn’t have to, but perhaps they wanted to see me succeed in whatever it is that I wanted to do. These moments stick with me, and I will never forget their kindness to me.

My advisor, Beth Strickland, has been a phenomenal mentor and someone whom I consider a great friend as well. Words do not do justice to what she has done for me in my life, professionally and personally. Before joining her lab, my experience as a 1st year graduate student without a permanent lab was the most difficulty and uncertainty I had felt in my professional life. There were many times I would not know if I would be continuing in research, or if those were my last days at Purdue and I would need to head back home to California. While I felt many windows were quickly closing around me, Beth had opened her doors to me, and helped slowly develop my skills as a researcher and build foundational knowledge of hearing science. Beth has afforded me with many more opportunities, including introducing me to many people in our field, and even feel welcome around them despite their stature. She has supported me to take opportunities and travel to present my research, and to network with other researchers for advice and collaboration. Beth has supported me to pursue my own professional goals, like to learn more about statistical analysis and coding, and understanding underlying mechanisms in normal and impaired systems. She is very supportive and thoughtful person outside of research, someone who takes the time to learn more about the person outside of the lab. Despite her kindness, she is also someone who is not afraid to challenge you, and to make sure that you are trying to improve yourself. I respect this honesty and sincerity, as they are necessary to grow as a person, and I am honored to have had her as my mentor and to have been a part of her lab.

I want to thank the members of my dissertation committee, whom have been very supportive and integral to my research. Hari Bharadwaj, who served as a co-mentor to me during my graduate studies, is a tremendous scientist and advisor. I knew that I wanted him on my

committee as soon as I heard his seminar at Purdue, because of his systems neuroscience approach, as well as his ability to take extremely complex topics and break them down into simple and explainable pieces. I have learned so much about auditory physiology, and have developed a foundation of computational skills from his guidance. Because of these skills and knowledge that I developed under Hari's guidance, as well as other factors, I have decided to pursue a postdoctoral research position in auditory physiology, studying otoacoustic emissions. Hari was very patient in his approach of mentoring, and was always kind enough to explain things thoroughly to me, or re-explain things that I didn't understand initially. His rigor and attentional to scientific detail have resonated with me as a researcher, and I strive to follow these qualities in my own research. I am grateful that I was able blend my research interests, from the behavioral side in the Strickland lab to the physiological and computational side in his lab.

I want to thank Mike Heinz, who has a great scientific mind and has always been willing to help me in my research endeavors. I have learned so much from Mike over the years, including his courses on the auditory periphery and biomedical signal processing. Although his courses were challenging, they were very stimulating, and I did not want to stop learning about his topics after the lectures ended. I think that is a key quality to have as a professor, to inspire and motivate others to learn about their topics. Despite all of his academic success, it is Mike's thoughtfulness that stands out to me. I remember how stressed I was early on in my doctoral program because of my preliminary exam, and Mike pointed me to different resources that would seem obvious to an observer but were not to me at the time. He also had no issue with lending materials to learn from, and has done so many times throughout my degree. I am thankful for his kindness, thoughtfulness, and generosity. I also want to thank Jeff Lucas, who has been a tremendous scientific mind and member on my committee. Jeff has always been willing help me or meet with me during my time at Purdue. I have referred to his expertise on scientific design, statistical analysis, and hypothesis testing. Jeff's perspective is often unique and persists with me beyond the moment of discussion, especially after hearing his questions after a research talk, or hearing his own research talks. Aside from that, he has an excellent sense of humor and wit, and I enjoy how he has blended that with the rigor of science. It has been an absolute pleasure to have learned from Jeff.

I want to thank the late Aïda Metzenberg, who was my research advisor as an undergraduate student at CSUN. Aïda was one of the sweetest individuals I've ever met, and definitely one of the smartest. At that time, I didn't have the biology and chemistry background

that would be needed to be successful in her lab, but she allowed me to take those courses while working as a research assistant. From my time in her lab, I learned what it meant to conduct research, to design a project, and to be a professional. To me, this is not trivial, and is probably the biggest reason I was able to begin my career as a researcher. From her belief and kindness toward me, I was able to pursue my passion. More importantly, she was this way to anyone who stepped foot in her lab, from undergraduate to graduate students, she always showed respect to anyone she interacted with. I also want to give thanks to Michael Gardner, a biopsychology professor at CSUN, who is the person who first got me interested in neuroscience research to begin with. From his course and enthusiastic personality, I found myself asking questions in class as well as in his office hours. He told me that I talk a lot and ask a lot of questions! I remember that I had felt bad about this initially, but he followed up by saying that he enjoys talking to students who are as enthusiastic about his field as he is. He further suggested that I should pursue a career in research if I want to answer some of the questions that I have, and that he'd write me a letter of recommendation if I wanted to pursue graduate school. This was a turning point in my life, and I am forever grateful for his kindness, support, and career advice that he provided me when I had no clue what a career in research entailed.

I want to thank former members of the Strickland lab who were beyond welcoming to me. Erica and Kristina in particular, who were in the final year of their PhD when I joined the lab, and made sure that I understood the logistics of the lab and any questions I had on research. Beyond lab mates, they are my friends, and I will not forget the kindness they showed me when I had no experience in the field. I also want to thank Skyler Jennings, who had already graduated and left the lab by the time I had arrived, but he made an effort to keep in touch with me and to help me understand anything research related, and support my research endeavors. I want to thank the current and former audiology students in the lab who made my work possible, and whom I've had enjoyable conversations with over the years, including: Miranda, Kristen Wade, Anna Hopkins, Alexis, Andrea. I have also been lucky to work in a department with a great crew of PhD and postdoc friends that I have learned from and laughed with, including Chandan Suresh, Satya Parida, Rav Singh, Francois Deloche, Subong Kim, Andrew, Vijay Muthaiah, and Andres Llico, Kevin Berry, and Andrew Sivaprakasam. I also want to thank my colleagues, staff, and others whom made my time at Purdue a special place I've made over the years in the SLHS department and hearing science community at Purdue, including: Robert Fujiki, Anumitha, Jordan, Fernando, Agudemu, Homeira, Varsha, Vibha, Ravi Krishnan, Joshua Alexander, Alex Francis, Jessica Huber, Jennifer Simpson, Preeti Sivasankar, Teasha McKinley, Vicki Black, Ed Bartlett, Donna Fekete, Keith Kluender, Mark Sayles, and Leroy Medrano. I also would like to thank all of my subjects I have tested over the years, as it would not be possible to do research without their participation. I am sure I am missing many names here, it is not intentional, and it speaks to the enormous amount of people whom I've had the pleasure to interact with at Purdue.

Lastly, I would like to thank my family for being supportive during my progress. I love my wife dearly, and I thank her and her mother for their support. I thank my father for providing the foundation and support needed to achieve my pursuits and dreams, and to be a loyal defender of our family.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

CEOAEs	click evoked otoacoustic emissions
dB	decibel
DPOAEs	distortion product otoacoustic emissions
GOM	growth of masking
Hz	hertz
IO function	input-output function
FM	forward masking
kHz	kilohertz
MEMR	middle ear muscle reflex
ms	milliseconds
MOCR	medial olivocochlear reflex
NB	narrowband
NP	nearest spectral peak
OAEs	otoacoustic emissions
peSPL	peak equivalent sound pressure level
PTC	psychophysical tuning curve
PSD	power spectral density
SFOAEs	stimulus frequency otoacoustic emissions
SL	sensation level
SNR	signal-to-noise ratio
SNHL	sensorineural hearing loss
SPL	sound pressure level
TEOAEs	transient evoked otoacoustic emissions
TMC	temporal masking curve
WAI	wideband acoustic immitance

ABSTRACT

Humans are able to hear and detect small changes in sound across a wide dynamic range despite limited dynamic ranges of individual auditory nerve fibers. One mechanism that may adjust the dynamic range is the medial olivocochlear reflex (MOCR), a bilateral sound-activated system which decreases amplification of sound by the outer hair cells in the cochlea. Much of the previous physiological MOCR research has used long broadband noise elicitors. In behavioral measures of gain reduction, a fairly short elicitor has been found to be maximally effective for an on-frequency, tonal elicitor. However, the effect of the duration of broadband noise elicitors on behavioral tasks is unknown. Additionally, MOCR effects measured using otoacoustic emissions (OAEs), have not consistently shown a positive correlation with behavioral gain reduction tasks. This finding seems counterintuitive if both measurements share a common generation mechanism. The current study measured the effects of ipsilateral broadband noise elicitor duration on psychoacoustic gain reduction (Chapter 2) and transient-evoked OAEs (TEOAEs) (Chapter 3) estimated from a forward-masking paradigm. Changes in the TEOAE were measured in terms of magnitude and phase. When phase was accounted for in the TEOAEs, the time constants were approximately equal to the psychoacoustic time constants, and were relatively short (~80 ms). When only changes in TEOAE magnitude were measured, and phase was omitted, the average time constants were longer (~172-ms). Overall, the psychoacoustic and physiological data were consistent with the timecourse of gain reduction by the MOCR. However, when the magnitudes from these data were directly compared in a linear mixed-effects model (Chapter 4), no positive predictive relationship was found, and in some cases there was a significant negative association between the physiological and psychoacoustic measures of gain reduction as a function of elicitor duration. The multitude of factors involved in this relationship are discussed, as are the implications of dynamic range adjustment in everyday listening conditions (noisy backgrounds) in both normal and hearing impaired listeners (Chapter 5).

CHAPTER 1: INTRODUCTION

1.1 Overview

Humans are able to hear over an extremely wide range of sound levels and discriminate changes as small as 1-dB across this range (Viemeister, 1983). From the faintest of sounds, such as a sewing needle hitting the floor in a quiet room, to very loud sound, such as a rock concert with thousands of screaming fans, humans can detect and discern sounds from one another even in the most complex and noisy environments. There appear to be multiple mechanisms that may optimize the coding of incoming sounds (Dean et al., 2005), which may support this broad perceptual dynamic range. One proposed mechanism is the medial olivocochlear reflex (MOCR), which reduces the gain of the cochlea (Murugasu and Russell, 1996; Cooper and Guinan, 2006). By reducing gain in the cochlea, it has been theorized that the MOCR can shift the dynamic range in everyday listening conditions. One area of particular interest is the time course of this effect by the MOCR and its relevance for auditory perception. For example, the adjustment by the MOCR would need to occur in a timely manner, or potentially important aspects of the sound of interest may be missed, such as detecting the fluctuating sound levels of speech in a noisy background. Studying the time course of the MOCR in carefully designed and controlled experiments can shed light onto one mechanism that the auditory system uses to adapt to sound over time. This dissertation explores the temporal dynamics of cochlear gain reduction, measured psychoacoustically and physiologically in the same subjects.

1.2 Auditory Peripheral Processing of Sound and The Cochlear Amplifier

A pivotal role that the mammalian peripheral auditory system has on hearing is the conversion of sound pressure waves to electrical neural signals, a process known as mechanotransduction. The peripheral auditory system has three primary regions, each of which acts to filter incoming sound: the outer, middle, and inner ear. After entering the outer ear, incoming sound will cause the tympanic membrane and the ossicles of the middle ear to vibrate. This vibrational energy will then be transferred to the cochlea of the inner ear, a snail shaped, fluid-filled structure that acts to separate sounds into its individual frequency components, similar to a bank of overlapping bandpass filters. This frequency analysis occurs on the basilar membrane, which resonates at different frequency locations along its length based on the average mass, stiffness, and damping of that particular region. These physical properties systematically vary along the length of the basilar membrane and determine the frequency response at any specific region. The basilar membrane vibrates maximally to high-frequency sounds in the basal region of the cochlea, whereas the basilar membrane vibrates maximally to low-frequency sounds in the apical region of the cochlea (Robles and Ruggero, 2001). Therefore, the vibrational energy propagates and travels across the basilar membrane (i.e., the traveling wave) at the region of its characteristic frequency (CF), which is the area of maximal vibration of the basilar membrane (Oghalai, 2005). These cochlear properties define its tonotopic organization.

Next, there are two sensory cells that are located along the basilar membrane in the organ of Corti, both of which are important in the mechanotransduction mechanism. This includes three rows of outer hair cells (OHCs) and a single row of inner hair cells (IHCs). OHCs are part of the "active process" or amplification of low level sounds, which in turn helps to sharpen the frequency specificity (i.e., the tuning) of the basilar membrane (Ruggero et al., 1997; Robles et al., 2001). Less gain or amplification is provided by the OHCs as the level of the sound increases, a property that contributes to the nonlinear response of the cochlea, and allows mammals to hear over an extraordinary range (Ruggero et al., 1997). Inner hair cells are responsible for converting the vibrational energy in the cochlear fluids into electrical signals that are then sent to auditory nerve fibers by synaptic transmission, and eventually to higher levels of the brain.

1.3 Cochlear Gain Reduction and the Medial Olivocochlear Reflex

As stated above, the OHCs provide gain to low intensity sounds by increasing the sensitivity of the basilar membrane close to and at it's CF. In doing so, the gain improves the sensitivity to soft sounds and sharpens the tuning of the basilar membrane, and enhancing the frequency selectivity the auditory system as a whole. While these characteristics allow for an extremely wide dynamic range of hearing in humans, there is surmounting neurophysiological and psychoacoustic evidence that adjustment of cochlear gain may be beneficial for auditory perception. One system that adjusts gain is the MOCR, a bilateral sound-activated feedback loop at the level of the brainstem, which projects to both cochleae and synapses directly on the base of the OHCs (Guinan, 1996). The MOCR decreases the gain produced by the OHCs in a frequency-specific manner relative to the elicitor, thereby decreasing the BM movement (Murugasu and Russell, 1996;

Cooper and Guinan, 2006). By reducing in the cochlea, it has been theorized that the MOCR can shift the dynamic range in everyday listening conditions. This is supported by the fact that MOCR activation can enhance auditory nerve responses to transient sound in a noisy background in animals (i.e., MOCR "antimasking"; Winslow and Sachs, 1988; Kawase et al., 1993), or increase sensitivity to changes in intensity in auditory perception (Almishaal et al., 2017; Strickland et al., 2018). It has also been hypothesized that it may enhance the fluctuation profile for complex sounds (Carney, 2018). While broadening cochlear tuning and reducing gain provided by the OHCs seems counterintuitive to improving perception, these concepts and their benefits can be conceptualized as a result of the dynamic range of hearing and the cochlear input/output (IO) function. For sounds on the compressive part of the cochlear IO function, small differences in sound intensity may be difficult to discern at the output of the filter due to the compressive nature in a healthy cochlea. However, if the cochlear gain were reduced (i.e., via MOCR activation) the range of compression of the cochlear IO function would become smaller (linearization of the IO slope) and result in an improvement in the signal-to-noise ratio at the output of the cochlear filter (e.g., Almishaal et al., 2017). This mechanism may also be important for speech intelligibility in background noise, where gain to the background noise is reduced by the MOCR and the SNR is improved.

1.4 Measures of Cochlear Gain Reduction in Humans

There have been multiple techniques used to study gain reduction in humans, including physiological and psychoacoustic methods. In most of these studies, the magnitude (i.e., the strength) of the gain reduction is what has been studied. Both physiological and psychoacoustic methods are discussed here.

1.4.1 Otoacoustic Emissions

Otoacoustic emissions (OAEs) have been used to noninvasively measure MOCR effects in humans and nonhuman mammals (Collet et al., 1990; Backus and Guinan, 2006; Lilaonitkul and Guinan, 2009a, b, 2012). OAEs are sounds produced by the amplification of the OHCs that can be recorded in the ear canal (Kemp, 1978), and are routinely used clinically to assess cochlear health. Sounds called elicitors can be used to activate the MOCR. These typically reduce the magnitude of the OAEs compared to a condition without an elicitor, which is interpreted as an estimation of the MOCR strength (Collet et al., 1990). By using elicitors of different laterality with respect to the recording ear, the relative strength of each MOCR pathway can be measured. Stimulus-frequency OAEs (SFOAEs) have also been used to estimate MOCR gain reduction as a function of probe frequency. Data from SFOAEs have shown larger elicitor effects at 0.5 and 1 kHz compared to 4 kHz for all elicitor lateralities (Lilaonitkul and Guinan, 2009a, 2012). However, in these studies, it was difficult to estimate effects at higher probe frequencies because the signal-to-noise ratio (SNR) was low at these frequencies for most subjects. One thing to consider is that the majority of studies measuring MOCR effects from OAEs only measure the changes in OAE magnitude, but, changes in OAE phase can also be measured (e.g., Francis and Guinan, 2010). However, it is not clear whether magnitude or magnitude plus phase is more relevant for perception (Lilaonitkul and Guinan, 2012).

1.4.2 Psychoacoustic Methods of Gain Reduction

Early interest in psychoacoustic measures of gain reduction began with a phenomenon called overshoot (Zwicker, 1965), also known as the temporal effect (Hicks and Bacon, 1992), in which a signal presented at the onset of a broadband masker may be detected at a lower signal-to-masker ratio if the signal and masker are preceded by an additional sound. This additional sound can be either an extension of the masker or a separate broadband sound, and is typically referred to as precursor or an elicitor. While there may be multiple mechanisms involved in the temporal effect, several pieces of evidence link it to cochlear gain reduction. The temporal effect is reduced with temporary cochlear hearing loss caused by aspirin (McFadden and Champlin, 1990) or noise (Champlin and McFadden, 1989). In these conditions, quiet threshold increases, but the signal-to-masker ratio at threshold for the signal at the onset of the masker decreases and becomes similar to the threshold with preceding sound. The temporal effect also decreases in a graded way with permanent cochlear hearing loss (Bacon and Takahashi, 1992; Strickland and Krishnan, 2005). In this broadband masker and precursor condition, the temporal effect is largest for midlevel signals or maskers and is larger at higher frequencies than at lower frequencies (Zwicker, 1965; Bacon and Takahashi, 1992; Strickland, 2001, 2004). It has been hypothesized that this is due to higher compression at higher frequencies than at lower frequencies [summarized in Bacon and Savel (2004)]. Presenting the precursor contralaterally produces a smaller or no temporal effect (Turner and Doherty, 1997; Bacon and Healy, 2000). A few studies have investigated overshoot psychoacoustically as well as physiologically with otoacoustic emissions (OAEs) to determine if the two estimates share a common gain reduction mechanism (Keefe et al., 2009; Walsh et al., 2010). One study found that the overshoot effect could be measured physiologically with OAEs as well as psychoacoustically, suggesting that cochlear gain and the MOCR were linked to overshoot.

One issue that arises in the interpretation of these simultaneous masking studies is that they may also include the effects of two-tone suppression. Two-tone suppression is a phenomenon that where responses of a point on the basilar membrane which are evoked by a tone at or near the CF can be reduced (i.e., “suppressed”) by the presentation of another tone (Rhode, 1977). In this way two-tone suppression also decreases cochlear gain, but on a much faster time scale (nearly instantaneous; Kiang et al., 1965) than the sluggish MOCR, and thus the results may be complicated by the interaction of the two mechanisms (Strickland 2004, 2008; Hegland and Strickland, 2018).

Forward masking has been used to investigate cochlear gain reduction without the possibility of two-tone suppression. Short signals and maskers may be used to measure functions hypothesized to reflect cochlear processing without the influence of gain reduction. These functions may then be measured with a precursor before the signal and masker. If the masker frequency is approximately an octave below the signal frequency, the growth of masking is hypothesized to reflect the cochlear input-output function (Oxenham and Plack, 1997). Using this paradigm, several experiments have shown that an ipsilateral precursor shifts the lower leg of the input-output function to higher signal levels, consistent with a decrease in gain (Krull and Strickland, 2008; Jennings et al., 2009; Roverud and Strickland, 2010; Jennings and Strickland, 2012; Yasin et al., 2014; DeRoy Milvae and Strickland, 2018). These forward masking techniques rely on using maskers at the signal frequency (on-frequency) and maskers approximately an octave below the signal frequency (off-frequency). It is assumed that the listener attends to the auditory filter with the best signal-to-masker ratio, which will typically be at or near the signal frequency. Gain reduction at the signal frequency place is expected for both the signal and the on-frequency masker, but not the off-frequency masker. Therefore, the change in signal threshold with an off-frequency masker following a precursor can provide an estimate of gain reduction. This differential processing between on- and off-frequency maskers is the basis for studying behavioral gain reduction and is not consistent with other mechanisms such as temporal integration of the masker and elicitor (i.e., additivity of masking) (Yasin et al., 2014; DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021). If multiple masker frequencies are used to trace out a psychoacoustic tuning curve, adding a precursor decreases frequency selectivity (Jennings et al., 2009; Jennings and Strickland, 2012).

Psychoacoustic gain reduction effects in forward masking have also been studied using contralateral precursors. Kawase et al. (2000) and Aguilar et al. (2013) found small decreases in frequency selectivity for signals from 0.5 to 4 kHz with a contralateral precursor. Fletcher et al. (2016) estimated input-output functions using temporal masking curves for an on- and off-frequency masker for a signal frequency of 2 kHz and found a small decrease in estimated gain with a contralateral noise precursor. Salloom and Strickland (2021) found that cochlear gain reduction was largest for ipsilateral and bilateral precursors, and smallest for contralateral precursors, an effect that was consistent across multiple signal frequencies (1-, 2-, and 4-kHz) for their subjects. The results from the Salloom and Strickland (2021) study were then compared to previous studies measuring ipsilateral or contralateral effects of gain reduction in terms of magnitude, which found that the gain reduction with contralateral precursors are generally very small compared to gain reduction with ipsilateral precursors across studies, consistent with the results of that study.

1.5 Time course of the MOCR

For many reasons it is important to study the time course gain reduction. For example, because gain reduction has been theorized to translate to an improvement in speech perception in noise (e.g., Clark et al., 2012), it would be useful to know how fast gain adjustment by the MOCR takes place. Another reason to study the time course of gain reduction is that it can provide insight on the appropriate durations of the stimuli used to study the MOCR. Many of the studies investigating the time course of the MOCR have used OAE-based measurements (Liberman et al., 1996; Kim et al., 2001; Maison et al., 2001; Bassim et al., 2003; James et al., 2005; Backus and Guinan, 2006). Some of these studies measured the changes in the distortion-product OAEs (DPOAEs) with contralateral elicitation or adaptation of the distortion-product with time (Kim et al., 2001, Bassim et al., 2003; James et al., 2005). Kim et al., (2001) measured the rapid adaptation of the DPOAE, and fitted these data with a two-exponential functions and found two separate time constants: A faster time median constant of ~69 - 70 ms (10 ms - 330 ms), and a slower median time constant of 1.5 sec (350 ms - 5.5 sec). Bassim et al. (2003), a follow-up study to Kim et al., found very similar time constants with their DPOAE adaptation and contralateral elicitation on DPOAE magnitude: their faster time constant was ~72.9 ms (7 ms -350 ms), and a slower median time constant of ~2.1 sec (350 ms – 8 sec). These results are largely consistent Liberman et al. (1996), where the adaptation of the DPOAE was measured in anesthetized cats, where the rapid adaptation decayed with a time constant of ~60 - 100 ms, and a slower time constant of approximately 1 sec. James et al., (2005) measuring the change in DPOAE magnitude over time with contralateral elicitation found that the buildup of the MOCR is in the order of 100 ms, with an onset and offset delay of approximately 25 ms from the onset and offset of gain reduction. Lastly, Backus and Guinan (2006) measured MOCR time constants by the change of the stimulus-frequency OAEs (SFOAEs) with elicitor duration and found that gain reduction is characterized by three time constants: a faster time constant ~60 - 80 ms, a medium time constant 290 - 350 ms, and a slower time constant 10 sec. In summary, the OAE-based measures of gain reduction show relatively short time constants 60 - 100 ms for the fast effects, and these effects overall build up over the course of hundreds of milliseconds, and there is also a much longer time constant for the slow effects of gain reduction. It should be noted that studies using elicitors to activate the MOCR used broadband elicitors.

The time course of gain reduction has also been studied psychoacoustically with forward masking paradigms and tonal stimuli. Roverud and Strickland (2010) measured the time course of forward masking gain reduction by manipulating the duration of the on-frequency precursor and the delay between precursor offset and masker onset. The precursor duration ranged from 5 – 100 ms. For some subjects, gain reduction increased with precursor duration up to about 50-ms, but then decreased or rolled over with a 100-ms precursor. These findings are consistent with findings from Krull and Strickland (2008), which showed that for some subjects, an on-frequency precursor with a 40-ms duration resulted in a larger reduction in gain than for a 160-ms duration precursor of the same level. In a follow-up study to Roverud and Strickland (2010), Roverud and Strickland (2014) used a psychoacoustic measure of gain reduction and found differential effects of duration for on- and off-frequency tonal precursors on signal threshold. This study used a wider range of (10 – 150 ms) and finer steps between (20 - 30 ms increments) precursor durations. For the on-frequency elicitor, thresholds increased with increasing duration up to about 50 ms, and then plateaued. In contrast, thresholds with off-frequency elicitors continued to increase with elicitor duration. Time constants were fitted to the on-frequency data as part of a model that also included a temporal integration window. The time constants ranged from approximately ~28 ms to approximately ~76 ms. These results are consistent with cochlear gain reduction, possibly by the MOCR, in which the on-frequency elicitor is affected by gain reduction at the signal frequency place, but the off-frequency elicitor is not. In contrast to the OAE-based studies estimating MOCR time constants, all of these psychoacoustic studies used tonal precursors, and the effects of broadband noise duration in similar paradigms are not currently known. Furthermore, it is not clear what the relationship between psychoacoustic and physiological measure of gain reduction with broadband elicitor duration are unknown. Both of these areas were explored in this study.

1.6 Potential Clinical Application of MOCR Gain Reduction and its Temporal Dynamics

One of the main theorized roles for MOCR gain reduction is to improve speech perception in noisy backgrounds. This is supported by the fact that MOCR activation can enhance auditory nerve responses to transient sounds in noise in animals (Winslow and Sachs, 1988; Kawase et al., 1993), an effect that has been called MOCR “unmasking”. Perhaps not coincidently, the most common complaint among hearing aid and cochlear implant users is speech intelligibility in noisy environments (Kochin, 2000). Many of these device wearers may suffer from severe to profound OHC damage or loss. The dynamic range adjustment of OHC gain by the MOCR that occurs in normal-hearing individuals likely does not occur for individuals, as their OHCs may produce little to no gain or compression, and in turn their hearing may have a reduced dynamic range. In the case of hearing aids, compression and gain can be added to their devices, but dynamic range adjustment is not included. For cochlear implant users, sound is processed directly to the auditory nerve, thereby bypassing the cochlea completely, and thus there is no dynamic range adjustment provided by the MOCR. Because gain adjustment by the MOCR may have potential benefit to speech perception in noise (Brown et al., 2010; Clark et al., 2012), there is increasing interest in signal processing strategies trying to incorporate efferent effects into their devices, such as in hearing aids (Jürgens et al., 2016) and cochlear implants (Lopez-Poveda et al., 2016). Some recent modelling studies have used computer models that mimic MOCR effects on speech perception in noise, and then varied the efferent time constant and/or the SNR to evaluate speech intelligibility (Yasin et al., 2018; Liu and Demosthenous, 2020; Yasin et al., 2020). While the findings of these studies support the idea that implementing gain reduction in a way inspired by the efferent system may improve speech intelligibility, more work is needed to understand how cochlear gain reduction works in humans generally. For example, it is possible that certain parts of speech benefit more from a fast time constant of gain reduction (Yasin et al., 2020), while other parts of speech benefit from a longer time constant (Yasin et al., 2018), or some combination of both (Liu and Demosthenous, 2018). Furthermore, these results of those studies showed the benefits of gain reduction with their modelling results, but the majority of the studies had not tested their algorithms directly in humans psychoacoustically. The current research document emphasizes measuring the time constants of gain reduction psychoacoustically and physiologically in the same subjects.

1.7 Multiple Studies Have Compared Psychoacoustic and Physiological Measures of Gain Reduction

The relationship between psychoacoustic and physiological measures of gain reduction is murky at best. As described in the section 1.4.2, gain reduction measured psychoacoustically is achieved by measuring the changes in a signal or masker threshold with the presence of an additional sound, termed a precursor. These tasks include simultaneous masking (e.g., overshoot) and forward masking paradigms with short signals and maskers, and long precursors intended to activate the MOCR. There are also psychoacoustic tasks that are inferred to be related to gain reduction, such as intensity discrimination tasks of a signal temporally embedded by a masker, with and without the presence of a contralateral elicitor to activate the MOCR (Micheyl et al., 1997). Physiologically, the strength of the MOCR can be measured with various OAE paradigms, where the difference in magnitude of the OAE is estimated with and without an elicitor present (described in section 1.4.1). Studies that have used similar types of measures to compare psychoacoustic and physiological measures of gain reduction within a single study will be described. It is important to note that some of these studies have tried to directly correlate their psychoacoustic and physiological measures to determine their relationship (Wicher and Moore, 2014; Fletcher et al., 2016; Maruffo-Pérez et al., 2021; Micheyl et al., 1997; Kawase et al., 2000), while other studies took an observational approach due to the methodology used (Keefe et al., 2009; Walsh et al., 2010). Table 1 summarizes these studies. The majority of the studies found no relationship between their two measures (Keefe et al., 2009; Wicher and Moore, 2014; Fletcher et al., 2016; Maruffo-Pérez et al., 2021), while others have found a positive relationship between their two measures (Micheyl et al., 1997; Kawase et al., 2000; Walsh et al., 2010). Studies that have compared speech perception in noise tasks to changes in OAE by an elicitor were not compared here as it is beyond the scope of this research.

There are many factors that might explain why there may be such mixed results in the relationship between psychoacoustic and physiological measures of gain reduction. As shown in Table 1, there are considerable differences in the methodology used in these studies. It is not clear how much of a role the differences in methodology and stimuli had in the comparisons made. Further, it is hard to compare results across studies for the same reason. For example, multiple studies used simultaneous masking paradigms (Keefe et al., 2009; Walsh et al., 2010; Wicher and Moore, 2014) which have the inherent involvement of two-tone suppression in their data, making it is hard to determine how much of the measured effect is due to MOCR alone, or some combination of MOCR plus two-tone suppression. For the studies using simultaneous masking, two studies used overshoot tasks in both their psychoacoustic and physiological measures (Keefe et al., 2009; Walsh et al., 2010), while the other measured psychophysical tuning curves (PTCs) and the change in the DPOAE with an elicitor (Wicher and Moore, 2014). These two approaches use very different methodologies, stimuli, and analyses. And even when the two methodologies were quite similar in their stimuli and methods used, the results contrasted with one another (Keefe et al., 2009; Walsh et al., 2010). This is also true for the OAE paradigms, where some studies measured the change in the DPOAE ([2f1-f2 amplitude]; Kawase et al., 2000; Wicher and Moore, 2014;), and others measured the change in the click-evoked OAEs (CEOAE; Micheyl et al., 1997; Fletcher et al., 2016) or SFOAE (Keefe et al., 2009; Walsh et al., 2010), or a combination of CEOAE and SFOAE (Marrufo-Pérez et al., 2021). A potential issue when comparing MOCR effects from DPOAEs to CEOAEs/SFOAEs is that DPOAEs do not share a primary generation mechanism with CEOAEs and SFOAEs (Shera and Guinan, 1999), and it is not clear how these differences impact the conclusions of those studies.

Table 1. Methodological details of various psychoacoustic and physiologic measures of MOCR effects and how they were related to one another from previous studies. The studies are ordered by the behavioral and physiological methods used, and precursor laterality (Contra = contralateral; Ipsi = ipsilateral). FM, forward masking; SL, sensation level; SM, simultaneous masking; PTC, psychophysical tuning curve; TMC, temporal masking curve.

Study	**Behavioral method**	**Phys. Method**	**# of Subj**	**Elicitor laterality**	**Comparison (* = significant)**
Kawase et al (2000)	PTC; FM, off-freq masker	DPOAE	6, 12 ears	Contra	Correlation *
Wicher & Moore (2014)	PTC; SM, off-freq masker (m above or below the sig)	DPOAE	6	Contra	Correlation
Fletcher et al (2016)	TMC; FM, estimated ΔG	CEOAE	12	Contra	Correlation
Micheyl et al., (1997)	Intensity difference limen	CEOAE	20	Behavior:; Contra plus Ipsi compared to Ipsi alone Physiology: Contra only	Correlation *
Marrufo-Pérez et al (2021)	Quiet threshold	CEOAE and DPOAE	7 - 15	Contra	Correlation
Keefe et al. (2009)	Overshoot task; SM	SFOAE	Behavior = 12, Phys = 14	Ipsi	Qualitative comparison, not similar
Walsh et al. (2010)	Overshoot task; SM	SFOAE	7	Ipsi	Qualitative comparison, similar

Another point of consideration is that the majority of these studies used a contralateral elicitor to evoke the MOCR. Contralateral elicitors avoid the effects of excitatory masking of the probe in the ipsilateral ear, and any change of the probe or signal should be due to the MOCR alone. This setup is especially useful when making multiple measurements in the ipsilateral ear over periods of time while the long duration or continuous contralateral elicitor is on. However, a recent study showed that gain reduction by ipsilateral and bilateral precursors is substantially larger than gain reduction by a contralateral elicitor, at least when measured behaviorally (Salloom and

Strickland, 2021). This finding is also consistent with neurophysiological innervation ratios of MOCR fibers in many mammals (Warr, 1992). Salloom and Strickland (2021) reviewed the pertinent psychoacoustic gain reduction literature which corroborated their findings, that gain reduction is much weaker when contralaterally elicited versus ipsilaterally or bilaterally elicited. While the exact innervation of ipsilateral and contralateral MOCR fibers in human is unknown, it is likely that using a contralateral elicitor to activate the MOCR may be measuring the weakest of the responses. It is not clear if this alone would change the results in the comparisons from these studies, but it is very possible that those studies measured an overall smaller effect than if measured with an ipsilateral elicitor. One last point to make is that the majority of these studies had twelve or fewer subjects (Kawase et al., 2000; Fletcher et al., 2016), and a few had very small subject pools [Walsh et al. (2010): 7 subjects; Wicher and Moore (2014) 6 subjects]. While it is understandable that these experiments take considerable time to complete and it is therefore harder to have a relatively large subject pool, correlational analysis power is often dictated by the underlying distribution of the data. A small subject pool may not necessarily reflect the population being studied, and it is worth considering how many subjects would be needed to achieve a certain statistical power before conducting the experiment. By increasing the overall subject size the statistical power of the analysis can also increase, and in turn would also reduce the possibility of committing a type II error (i.e., a false negative) as statistical power and subject size are inversely related to the probability of the type II error. Again, it is not clear if these studies took this into account, or if their conclusion would have changed otherwise.

One last consideration is that individual subject variability is not accounted for in the analysis in any of these studies. It is known that correlational or observational analysis, used in all of those previous studies, can only account for group level differences. This is important as previous research has shown that gain reduction responses can vary quite a lot when measured psychoacoustically (e.g., Jennings et al., 2009) and with OAEs (e.g., Mertes and Goodman, 2016). While there are many other issues that arise when comparing directly psychoacoustic and physiological measures of gain reduction (for review see: Marrufo-Pérez et al., 2021; Jennings, 2021), the current research document will specifically address the issues outlined here in the design of the research, including: the similarity of methodology and stimuli between gain reduction measures, the laterality of the elicitor, the overall subject size, and subject variability.

CHAPTER 2: A PSYCHOACOUSTIC FORWARD MASKING PARADIGM TO ESTIMATE GAIN REDUCTION: TEMPORAL DYNAMICS

2.1 Introduction

Previous psychoacoustic studies have estimated the time course of effects that are consistent with cochlear gain reduction using tonal precursors either at the signal frequency (on-frequency) of far below the signal frequency (off-frequency) (Roverud and Strickland, 2010; Roverud and Strickland, 2014). In the Roverud and Strickland (2014) study, signal threshold increased with on-frequency precursor duration up to about 50-ms, and then plateaued or rolled over. In contrast, thresholds with off-frequency elicitors continued to increase with elicitor duration. These results are consistent with cochlear gain reduction, possibly by the MOCR, in which the on-frequency elicitor is affected by gain reduction at the signal frequency place, but the off-frequency elicitor is not. However, no study has estimated the effects of broadband elicitor duration in a similar way, yet many psychoacoustic studies use broadband sound to elicit the MOCR based on OAE data that show that medium MOCR time constants are in the hundreds of milliseconds (e.g., Backus and Guinan, 2006). Therefore, the current experiment can be thought of as a direct extension of these previous studies, but with the use of broadband elicitors and a larger range of precursor durations. The purpose of the current experiment was two-fold: to better understand how broadband elicitor duration affects auditory perception, and to explore the relationship between these data compared to physiologically measured gain reduction using an OAE paradigm (Chapter 4).

2.2 Methods

The psychoacoustic experiment detailed in this chapter used two forward masking techniques to measure cochlear gain reduction at 2 and 4 kHz, using a range of broadband precursor durations.

2.2.1 Subjects

Nineteen subjects (7 male and 12 female) completed the experiments in the current study. Their ages ranged from 19 to 35 years (median = 24 years) at the time of testing. This is a relatively large subject pool compared to the studies described in section 1.7 comparing psychoacoustic and physiological measures of gain reduction. All subjects had normal auditory function, determined through the use of a battery of audiologic measures. All subjects had clinically normal pure tone thresholds [≤ 15 dB hearing level (HL) at audiometric frequencies between 250 and 8000 Hz]. Distortion product otoacoustic emissions (DPOAEs) were present (Bio-logic system, Natus Medical Inc., Pleasanton, CA) from 1500 to 8500 Hz (minimum criteria of –6 dB SPL distortion product and 6 dB SNR for 10 of 12 frequencies tested with no consecutive absent responses). Tympanograms (Tympstar, Grason-Stadler, Inc.) were normal (type A), indicating normal middle-ear function. Ipsilateral acoustic reflex thresholds were measured using white broadband noise elicitors. Because the signal was always presented in the subject's right ear in the experiments of the current study, the clinical acoustic reflex thresholds were measured with respect to the probe in the right ear, except for subject 18, where acoustic reflex thresholds were measured from the subject's left ear. This was due to a notch in the audiometric threshold at and around the 2 kHz probe frequency in the right ear for that subject, but not for her left ear, and thus her left ear was the test ear during the experiments The clinical acoustic reflex thresholds were measured in dB HL and converted to corresponding dB SPL units for fair comparison to the experimental elicitors (i.e., precursors) used in the current study. To do so, noise levels from the immittance equipment were recorded from a sound level meter attached to a Zwislocki coupler mounted in a KEMAR ear. The noise levels in dB SPL were approximately 8 dB higher than the nominal levels in dB HL. Overall, no subject's acoustic reflex threshold was below 58 dB SPL, a level that is higher than the precursor level used in our experiments (50 dB SPL). These thresholds can be found in Table 4.

Two potential subjects fell out of the study. One subject had subjective tinnitus, which would have been confounding in our tone detection tasks, as subjects are required to detect quiet sounds in the presence of a precursor and/or masker. The other potential subject could not consistently perform the behavioral tasks, and had issues with differentiating the signal from the other sounds, and was thereby discontinued from the study after multiple days of unsuccessful practice sessions. None of their data are reported here. All subjects were paid for their time in the study except for S1, who is the first author. Other subjects were recruited via fliers on the Purdue campus. All research was conducted under a research protocol approved by the Institutional Review Board at Purdue University to safeguard the rights, safety, and well-being of our subjects.

2.2.2 Psychoacoustic Stimuli

Estimates of gain reduction were made at 2 and 4 kHz using two forward masking techniques that rely on the timing of cochlear gain reduction via the MOCR. The technique used to measure gain reduction with the use of short duration maskers ("masker present" conditions) will be explained first. The 2- and 4-kHz signals used in the following behavioral experiments were 10-ms sinusoids, including 5-ms $\cos^2$ onset and offset ramps. This signal duration is longer than the 6 or 8 ms used in some previous studies (e.g., Jennings et al., 2009; Roverud and Strickland, 2010; DeRoy Milvae and Strickland, 2018) to ensure that the spectral spread was within one auditory filter bandwidth (DeRoy Milvae and Strickland, 2018). Next, the masker levels needed to mask a signal fixed at 5 dB sensation level (SL) was determined. A 5-dB shift in signal threshold was desired so that the signal was fixed on the lower leg of the cochlear input-output function. Gain reduction is largest in this region of the input-output function (physiologically: Cooper and Guinan, 2006; psychoacoustically: Krull and Strickland, 2008; Roverud and Strickland, 2010). The masker duration was 20 ms (including 5-ms $\cos^2$ ramps), which should be too short to elicit the MOCR during the signal presentation (James et al., 2005; Backus and Guinan, 2006).

For gain reduction measurements, the maskers were on-frequency (at the signal frequency) and off-frequency (0.6 times the signal frequency) sinusoids. That is, 2- and 1.2-kHz maskers were used for the 2-kHz signal frequency, and 4- and 2.4-kHz maskers were used for the 4-kHz signal frequency. The signal was set at 5 dB SL, and the masker level was adjusted to find the level needed to just mask the signal. By measuring both on- and off-frequency masked threshold, we were able to test whether the effects of the precursor on signal threshold were consistent with cochlear gain reduction (explained below). To verify that the maskers were equally effective at masking the signal, the on- and off-frequency maskers were fixed at the thresholds determined earlier, and the signal was varied to check that the signal thresholds were raised to 5 dB SL. On- and off frequency masked signal thresholds were considered similar if the difference between the two conditions was less than 3 dB. If signal threshold with the fixed maskers differed by more than this amount, the masker level was adjusted, and the signal threshold was remeasured until signal thresholds were within 3 dB of one another. The configuration for these thresholds are shown in Fig. 1(A) (off-frequency) and Fig. 1(C) (on-frequency). Signal thresholds in this reference condition were then compared to a condition where a precursor intended to elicit the MOCR preceded both the masker and the signal [Fig. 1(B) - off-frequency and Fig. 1(D) – on-frequency, respectively].

The precursor was a 50-dB SPL pink broadband noise (0.25 – 10 kHz) which varied in duration [50, 65, 100, 200, 400, 800 ms], including 5-ms $\cos^2$ onset and offset ramps. Pink noise has a spectrum level that decreases by 3 dB per octave. This elicitor provides a more accurate comparison of gain reduction across frequency, as pink noise will excite auditory filters across the frequency range with approximately equal energy. Additionally, previous studies have found broadband noise stimuli to be particularly effective elicitors of cochlear gain reduction (Maison et al., 2000; Lilaonitkul and Guinan, 2009a; Wicher and Moore, 2014). During a single session, the pink noise was presented ipsilaterally with respect to the masker and signal. Ipsilateral precursors have shown to produce substantially larger effects (at many signal frequencies) that is consistent with cochlear gain reduction when compared to contralateral elicitors of the same sound pressure level, at least measured psychoacoustically (Salloom and Strickland, 2021). 50 dB SPL is a relatively low elicitor level compared to many other studies in humans measuring MOCR effects, which typically use 60-dB SPL elicitors. While many studies assume that a 60-dB SPL elicitor should be too low in level to elicit the MEMR, it is possible that the measures used to estimate the acoustic reflex thresholds in those studies underestimate the true threshold of reflex activation because of the sensitivity of the test. For example, it has been shown that MEMR thresholds using WAI tympanometry are on average ~14 dB lower than MEMR thresholds measured using typical clinical tympanometry (Keefe et al., 2010; Feeney et al., 2017). Therefore, we used WAI tympanometry to estimate each subject's ipsilateral MEMR threshold (see section, 3.3.1), and avoid MEMR elicitation. No subject in the current study had an ipsilateral WAI MEMR threshold ≤ 50 dB SPL for broadband noise (see Table 4). The range of precursor durations used in the current study cover a broader range than those used in the Roverud and Strickland (2014) study, and should overlap with the entirety of the buildup of the MOCR (James et al., 2005; Backus and Guinan, 2006).

A reduction in gain by the precursor is predicted to shift signal threshold more following the off-frequency masker than the on-frequency one (Kawase et al., 2000; Jennings et al., 2009; Yasin et al., 2014). This is because the off-frequency masker is processed linearly at the signal frequency place and thus not affected by gain reduction, whereas the signal and the on-frequency masker are nearly equally affected by gain reduction. This contrasts with the prediction of temporal integration, also called additivity of masking, where adding the precursor would produce equal shifts in signal threshold in the two conditions (Penner and Shiffrin, 1980; Plack and O'Hanlon, 2003; Oxenham and Moore, 1994). For forward maskers presented sequentially, temporal integration posits that masking occurs via a neural mechanism that integrates the energy of the stimuli within a temporal window (Oxenham and Moore, 1994; Oxenham, 2001). In this case, the intensities of the precursor, masker, and signal would be integrated at some level of the auditory system. With temporal integration (i.e., additivity of masking), it would be expected that the on- and off-frequency conditions should produce equal shifts in thresholds with the addition of the precursor. In contrast, in this study, it is proposed that masking by the masker may occur within a temporal window (because the duration is too short for gain reduction to affect the signal) but masking from the precursor occurs by gain reduction. Differences between gain reduction and additivity of masking in forward masking are outlined in detail in the Discussion section of Salloom and Strickland, 2021.

Off-frequency masked effects produce larger effects than on-frequency masked effects, an effect that is consistent with gain reduction (Milvae DeRoy and Strickland, 2018; Salloom and Strickland, 2021). Therefore, signal thresholds with an off-frequency masker were used for the analyses (magnitude and time constants), as they have been interpreted as a change in gain. Therefore, the difference in signal threshold between the off-frequency masked signal with [Fig. 1(B)] and without [Fig. 1(A)] the precursor will be referred to as the "masker present" gain reduction estimate used in the current study. A second method was used to estimate gain reduction at 2- and 4-kHz signal frequencies. Instead of using an off-frequency masker to fix the signal on the lower leg of the input-output function, quiet threshold of the signal served as the baseline condition [Fig. 1(E)] and was compared to signal threshold with a precursor and a 20-ms delay between the precursor offset and signal onset [Fig. 1(F)]. This will be referred to as "masker absent" estimate of gain reduction. Previous studies using masker present vs masker absent conditions found no significant differences in the magnitude of the gain reduction between the two methods (DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021) for signal frequencies of 2 and 4 kHz. We used this masker absent condition to determine if it had a similar growth pattern to the off-frequency masked condition, as well as because it is more similar in stimuli to the TEOAE experiment than our masker present conditions. The latter point was reasoned as we seek to compare behavior to physiology in other analyses, and therefore it may be more of a fair comparison between measures. To better understand gain reduction effects, a schematic of our listening conditions shows the assumed underlying cochlear input/output (I/O) functions in the next subsection. In this study, the growth in signal threshold with precursor duration was measured for each signal frequency. From these thresholds, gain reduction was measured in terms of magnitude and the time constants.

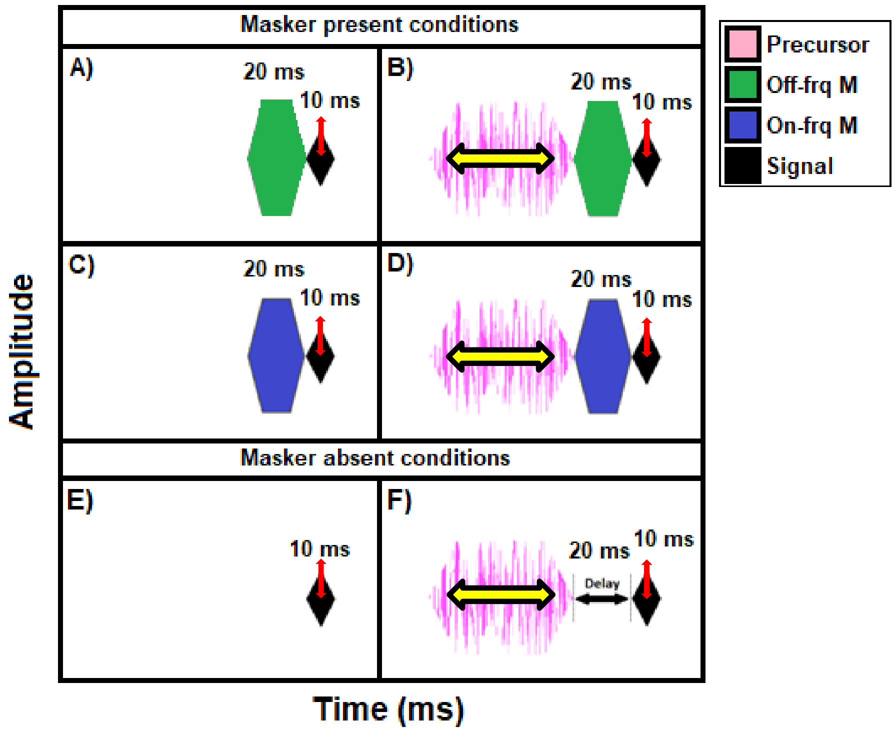

Figure 1. Schematic of the stimuli used for the masker present method [A and B, off-frequency; C and D, on-frequency] and the masker absent method [E and F], respectively. On- and off-frequency maskers were always 20 ms, as was the delay between the precursor and signal in the masker absent condition. The precursors were presented ipsilaterally with respect to the signal ear. Precursor durations ranged from 50 to 800 msec, shown by the yellow arrows on the precursor. The double-headed arrow (red) indicates that the signal was adaptively varied, while the masker was fixed at a level that shifted the signal by 5 dB with no precursor present.

2.2.3 Gain Reduction I/O Function Schematics

Fig. 2 is a schematic representation of how a reduction in gain by a precursor would affect the signal (and masker, if present) for each listening condition [off-frequency – panels (A) and (B); on-frequency - panels (C) and (D); masker absent – panels (E) and (F)]. Included in each panel is the time-frequency representation of the stimuli, and a theoretical representation of how the masker or precursor plus masker affect the signal. For the conditions without a precursor, there are some assumptions that should be addressed to better understand the schematic. First, it is assumed that the listener makes use of an auditory filter with a center frequency at or close to that of the signal frequency. Second, only the components of the masker passing through this filter have any effect in masking the signal. Third, signal threshold is assumed to correspond to a certain signal-to-masker ratio at the output of the auditory filter. These assumptions are foundational to our work and are based on the power spectrum model of forward masking (Fletcher, 1940).

The input-output (I/O) functions for the signal frequency (e.g., 2 or 4 kHz in this study) are shown by the solid line in each panel. These I/O functions approximate the growth of excitation curves measured physiologically on the basilar membrane (e.g., Ruggero et al., 1997). For each listening condition, baseline conditions (i.e., no precursor) are shown in the left panels while corresponding precursor present conditions are shown in the right panels. Both the on- and off-frequency conditions have the signal at 5 dB SL, while the no masker condition tracks the signal at quiet threshold.

Panel A shows the off-frequency masked condition where the signal is detected at a constant signal-to-masker ratio, which is the difference between the signal and masker levels at the output of the filter, shown by the double-headed arrow between the output levels for the signal and the masker. When the precursor is added to this condition [panel (B)], gain will be decreased, shown by the dashed line, and the signal will be affected (shown by the downward arrow and italicized '*s*'), however, notice that the off-frequency masker will not. This is because the off-frequency masker is processed linearly at the signal frequency place, and only the signal is affected by the precursor. Because the signal output decreased, the input of the signal level must be increased in order to re-establish the original threshold signal-to-masker ratio, shown by the right facing arrow on the x-axis. Overall, this will lead to an increase in the signal threshold based on this additional input.

For the on-frequency masked conditions, shown in panel (C), the signal and the masker share a common frequency, so are both on the solid line. The signal is also at 5 dB SL, and a constant signal-to-masker ratio is established at the output of the auditory filter for detection shown by the double arrow at the output. The addition of the precursor, in panel (D), shows that gain reduction does occur indicated by the dashed line, however, it affects the signal and masker equally, resulting in no change in the signal-to-masker ratio at the output of the filter and thus no change in signal threshold. The difference in processing between the on- [panels (A) and (B)] and the off-frequency [panels (C) and (D)] maskers is the basis of our gain reduction experiments, and is not consistent with excitatory masking. If masking by the precursor were excitatory, and equal increase in signal threshold would be seen in the on-frequency and off-frequency masker conditions.

Lastly, for the masker absent condition, in panel (E), the signal is detected at quiet threshold. In panel F, the precursor reduces the gain of the signal, indicated by the dashed line. A downward arrow and italicized ‘s’ indicate that the output level of the signal is lower and is not detectable at this level. In order to achieve the output level needed for detection, the input of the signal needs to be increased, indicated by the rightward arrow on the x-axis, and this causes the output of the signal to return to its original level and become detectable.

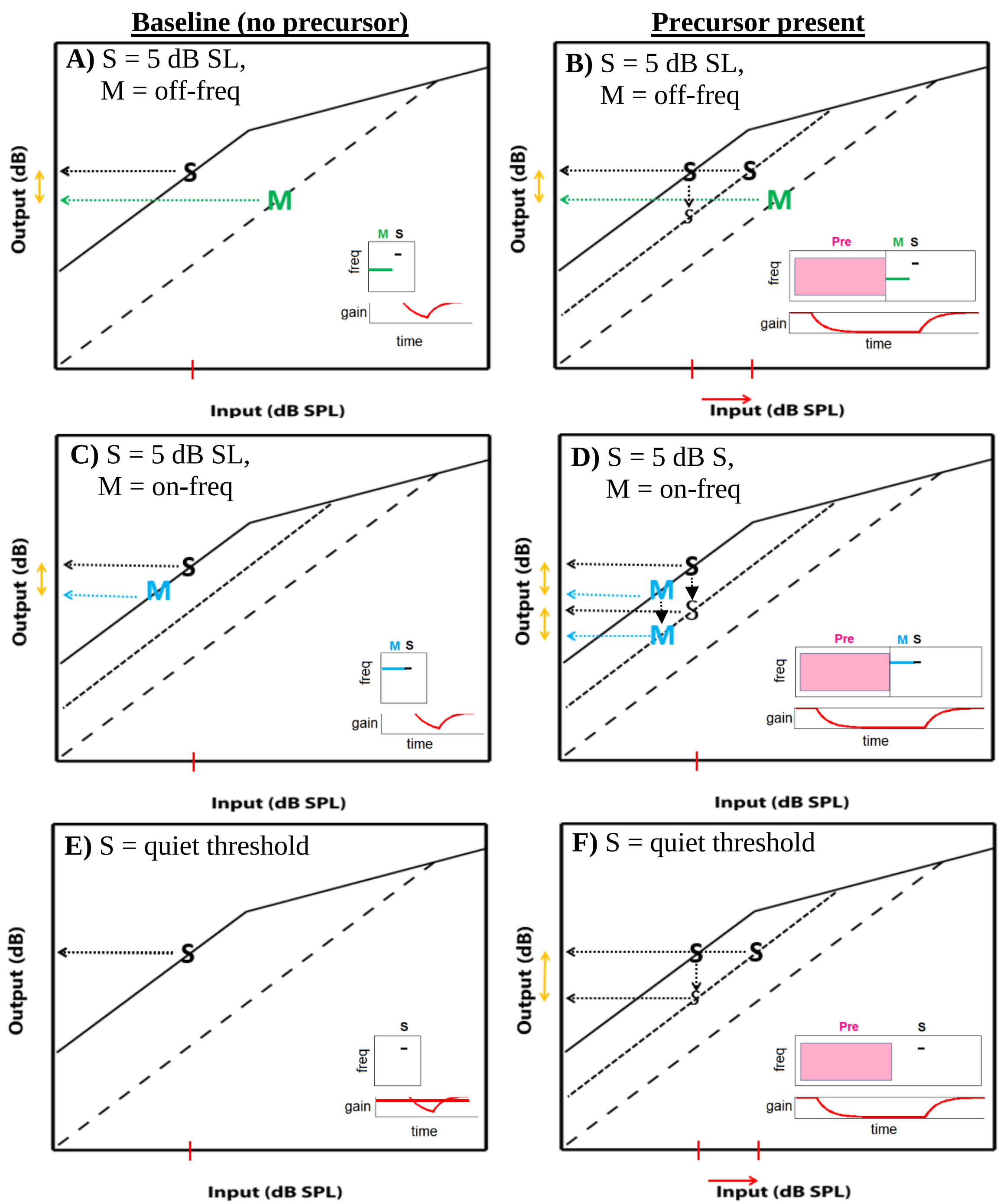

Figure 2. Schematic of gain reduction effects on the signal and masker for each listening condition depicted with cochlear I/O functions. Baseline conditions are the panels to the left, and with the corresponding precursor present conditions in the panels to the right. The solid line represents the responses to the signal within a filter at or near the signal frequency. Responses to the signal with gain reduction are indicated by the dashed line. Off-frequency conditions are in panels A and B, on-frequency conditions are in panels C and D, and masker absent conditions are in E and F.

2.2.4 Procedure

All psychoacoustic measures were completed in a double-walled sound-attenuating booth. Stimuli were generated with custom MATLAB software (Bidelman et al., 2015) with a Lynx TWO-B sound card (Lynx Studio Technology, Inc., Costa Mesa, CA). The stimuli were then passed through a headphone buffer (TDT HB6, Tucker-Davis Technologies, Alachua, FL) and delivered to one ear through an Etymotic ER-2 (Etymotic Research, Inc., Elk Grove Village, IL) insert earphone. The subjects had insert earphones in both ears. The insert earphones had a flat frequency response at the eardrum from 250 to 8000 Hz. High pass noise (from 1.2 times the signal frequency to 10 kHz) was used to reduce the possibility of off-frequency listening (Nelson et al., 2001) for all parts of the experiment except during quiet threshold measurements. The high pass noise began 50 ms before the onset of the stimuli and ended 50 ms after the signal offset and was 50 dB below the signal level.

All psychoacoustic measurements utilized a three-interval forced-choice (3IFC) task using a MATLAB GUI, in which only one of the choices contained the signal. Each interval was visually marked on the computer screen, and intervals were separated by 500 ms of silence. Subjects could use either a mouse or the keyboard to indicate which interval contained the signal. Visual feedback was given for correct and incorrect responses. Signal and masker levels were adjusted to estimate a response threshold of 70.7% correct (Levitt, 1971). For signal threshold measures (quiet thresholds and measures of gain reduction), if the subject chose correctly over two consecutive trials, the level of the signal decreased, while an incorrect response would cause the level of the signal to increase (two down, one up). For masking thresholds, if the subject chose correctly over two consecutive trials, the level of the masker increased, while an incorrect response would cause the level of the masker to decrease (two up, one down). The step size was 5 dB for the first four reversals and then decreased to 2 dB for the remaining reversals. The last eight reversals were averaged to produce a final threshold for each run.

Subjects had approximately 1 h of training before data collection began in order to help them understand the task. Each session was 1–1.5 h to prevent attentional fatigue. Each condition was tested at least twice per session, and thresholds are an average of the last two thresholds recorded for that condition. These final thresholds served as the data reported throughout the current study in the figures and the statistical analysis. Runs with a standard deviation (SD) greater than 5 dB were discarded from the overall averages and repeated if necessary. Data from each subject were collected for a minimum of one session for each psychoacoustic task, and additional sessions were conducted if large variability or learning effects occurred. The order of presentation of the signal frequency and masking condition of the precursor was interleaved across subjects. All statistical and post hoc analyses of gain reduction were calculated with IBM SPSS 28 statistical software.

Before any statistical test was conducted, all of the data sets were tested for the assumption of normality in SPSS using the Shapiro–Wilk test of normal distribution and the corresponding normal Q-Q plots. With this test, any subset of data tested for this assumption with a p-value equal to or greater than 0.05 would meet the criterion to assume normal distribution.

Aside from two subjects (subject 10 and subject 16), whose off-frequency threshold shifts at 2 kHz were considered statistical outliers and therefore excluded from the analysis and overall averages, all other data from our subjects met the assumption of normality. A statistical outlier in this case was defined as any value that lies outside 1.5 times the interquartile range (i.e., above or below the 75^{th} and 25^{th} percentiles respectively). While data from these two subjects in the off-frequency conditions were considered statistical outliers and not included in the analysis, their corresponding on-frequency and no-masker threshold shifts at 2 kHz met the assumption of normality, and there were large differences between their off-frequency and on-frequency threshold shifts indicating that their data are consistent with gain reduction.

2.3 Results and Analysis

2.3.1 Magnitude of Gain Reduction

The individual and average threshold shifts as a function of precursor duration for the masker present and masker absent conditions can be seen in Fig. 3 and Fig. 4, respectively. The overall qualitative pattern, for both the individual and group data at 2 and 4 kHz is that threshold shifts with an off-frequency masker (square symbol) are consistently larger than threshold shifts with an on-frequency masker (diamond symbol), which is consistent with gain reduction and not excitatory masking. Additionally, the off-frequency masked conditions produce a growth function that is qualitatively similar in magnitude and overall shape to the masker absent conditions (square symbol). Both of these observations were specifically tested in the section below. Furthermore, these data will serve as the basis for the magnitude and time constants of gain reduction in the next sections of the chapter. Note that all 19 subjects participated in the 2 kHz conditions, while 9 subjects (subjects 1-8 and 11) completed the 4 kHz conditions.

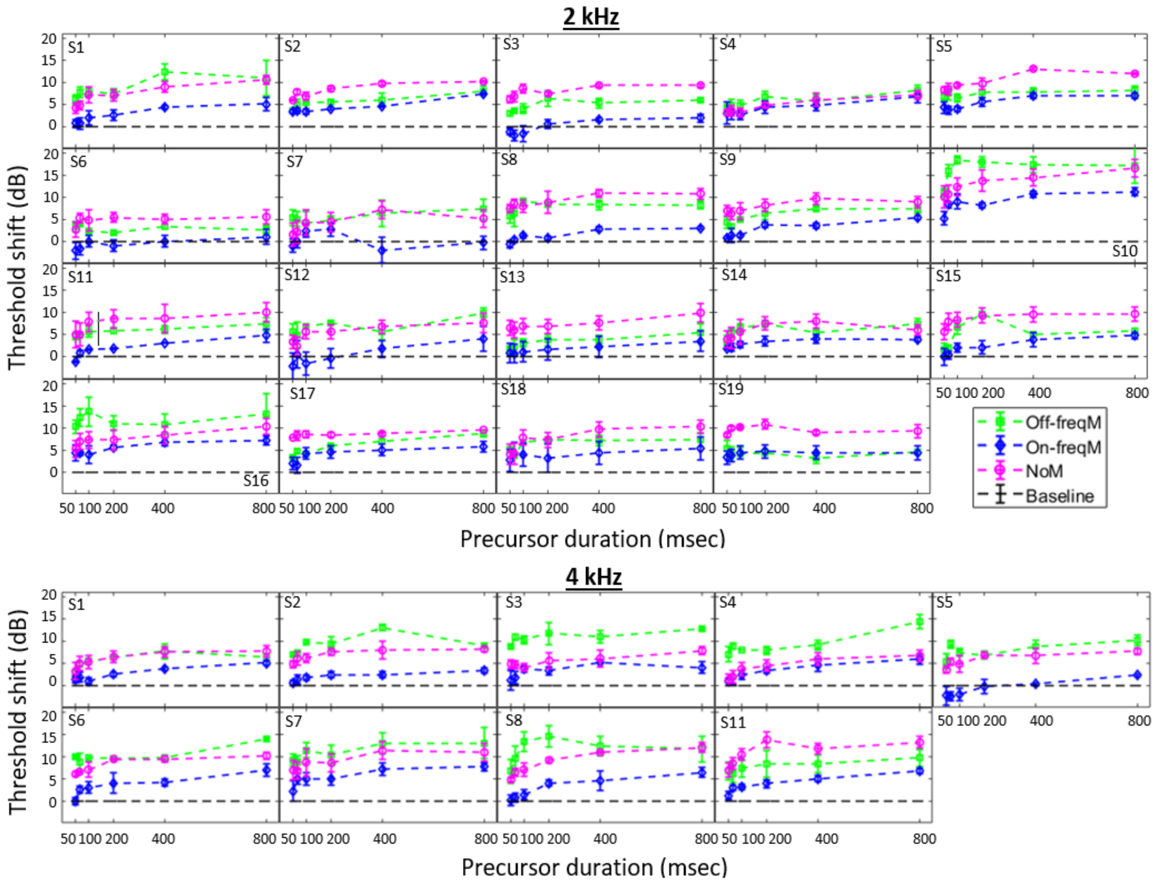

Figure 3. Signal thresholds as a function of precursor duration for individual subjects for signal frequencies of 2 kHz (top) and 4 kHz (bottom). Symbols indicate listening condition, while dotted lines indicate the reference condition with no precursor. SD of signal thresholds is indicated by the error bars.

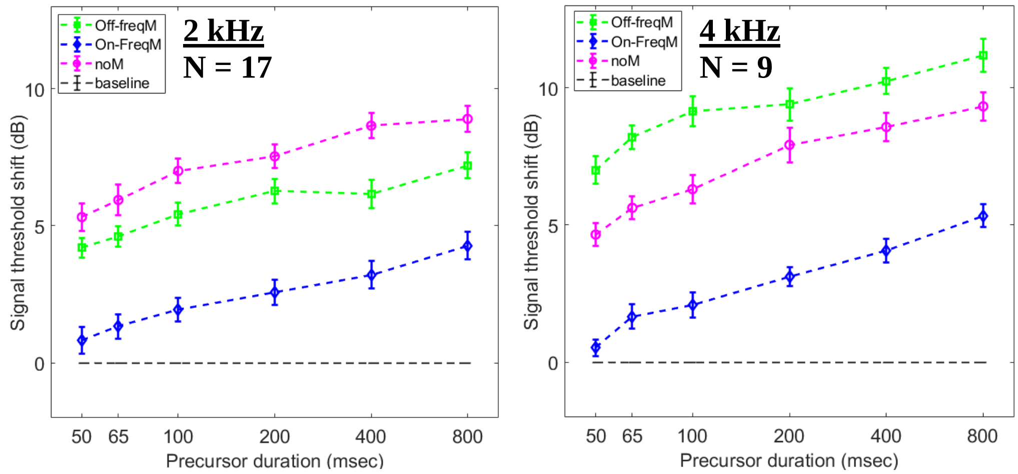

Figure 4. Average signal threshold shifts as a function of precursor duration for 2-kHz (left) and 4-kHz (right) signal frequencies, estimated from the data in Fig. 3 (and identically labeled). For 2 kHz data, subjects 10 and 16 were omitted from the overall average and statistical analysis as the off-frequency masked conditions did not meet the assumption of normality. Symbols indicate listening condition, while the horizontal dashed line represents the reference condition with no precursor. SEM of signal thresholds is indicated by the error bars.

Two-way 3 x 6 repeated measures ANOVAs were conducted to determine if the mean signal threshold shifts (dependent variable) significantly varied with the independent variables masking type (off-frequency, on-frequency, and no-masker) and precursor duration (50, 65, 100, 200, 400, and 800 msec). This analysis was done for both signal frequencies, 2 and 4 kHz, but not combined in a single analysis due to unequal subject sizes (2 kHz: N = 17; 4 kHz: N = 9). The results of these tests are summarized in Table 2. The 2-kHz results will be discussed first.

The results for the two-way 3 x 6 repeated measures ANOVA for the 2-kHz signal indicated that there was a significant main effect of masker type. Bonferroni corrections revealed that threshold shifts for the off-frequency masked conditions were significantly different than those for the on-frequency masked conditions, and similarly, there was a significant difference in threshold shifts between the no-masker and the on-frequency conditions. Lastly, there was a significant difference found between the off-frequency and no-masker conditions. These differences between listening conditions as a function of precursor duration can be found in Fig. 4 (2 kHz – left panel).

Next, a significant main effect was found with precursor duration, indicating that the magnitude of these shifts generally increased with precursor duration. Furthermore, the relative pattern of these differences was approximately constant across precursor duration (Fig. 5 – left panel). Across precursor durations, the average difference between the off-frequency and on-frequency masked conditions is 3.29 dB (SD = 0.30 dB), the no-masker and on-frequency masked conditions is 4.86 dB (SD = 0.36 dB), and the off-frequency and no-masker conditions is 1.57 dB (SD = 0.49 dB). This is to say that increasing the precursor duration shifted each listening condition by roughly the same amount. No interaction was found between the main effects. Lastly, in terms of overall magnitude of our gain reduction conditions, the off-frequency masked condition produced threshold shifts that ranged from 4.21 dB – 7.21 dB across precursor duration. The no-masker condition produced threshold shifts ranging from 5.32 dB – 8.90 dB. In either condition, threshold shifts increased by about 1 dB or less per doubling of precursor duration.

A two-way 3 x 6 repeated measures ANOVA was also conducted for the 4-kHz frequency, and indicated that there was a significant main effect of masker type. Bonferroni corrections revealed that threshold shifts for the off-frequency masked conditions were significantly different than those for the on-frequency masked conditions, and there was a significant difference in threshold shifts between the no-masker and the on-frequency conditions. However, there was no significant difference found between the off-frequency and no-masker conditions, which contrasts from this comparison from the 2 kHz results, and is consistent with previous work where these listening conditions produced similar thresholds (DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021).

A significant main effect was found with precursor condition at 4 kHz. Across precursor durations, the average differences between the off-frequency and on-frequency masked conditions is 6.39 dB (SD = 0.40 dB), the no-masker and on-frequency masked conditions is 4.26 dB (SD = 0.33 dB), and the off-frequency and no-masker conditions is 2.12 dB (SD = 0.54 dB). Also similar to 2 kHz, the relative pattern of the differences between conditions was very consistent across precursor duration (Fig. 5 – right panel). No interaction was found between the main effects. The off-frequency masked condition produced threshold shifts that ranged from 7.0 dB – 11.18 dB across precursor duration. The no-masker condition produced threshold shifts ranging from 4.65 dB – 9.32 dB. Similar to the 2 kHz results, threshold shifts increased by about 1 dB or less with doubling of the precursor duration. We do note that the no-masker condition produced larger than off-frequency masked conditions for 2 kHz conditions, but this relationship switched at 4 kHz where the no masker conditions were smaller than the off-frequency masked conditions. It was shown that off-frequency and no-masker conditions produced substantially larger shifts with precursor duration compared to the on-frequency masked conditions (Fig. 5), and these differences were statistically significant (Table 2).

To summarize, for both 2 and 4 kHz signal frequencies, the off-frequency and no-masker conditions produced substantially larger shifts than the on-frequency masked conditions (Fig. 4 and Fig. 5). The large shifts in off-frequency versus on-freuqency masked conditions are consistent with gain reduction (DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021) and not excitatory masking. While the no-masker conditions produced statistically larger threshold shifts than the off-frequency masked conditions at the 2 kHz frequency (1.57 dB average), the data at 4 kHz and previous studies have shown that the no-masker condition has comparable magnitude effects to the off-frequency masked condition in terms of magnitude, and thus has also been interpreted as cochlear gain reduction (DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021). Overall, the effect of precursor duration was significant, increasing threshold shifts by 1 dB or less for both signal frequencies.

Table 2. Results of the 3 x 6 repeated measures ANOVAs on the magnitude of gain reduction as a function of masking type (off-frequency, on-frequency, no-masker) and precursor duration (50, 65, 100, 200, 400, 800 msec). Note that the degrees of freedom in the F-values for precursor duration are not integers, which is because sphericity could not be assumed for these data. The more conservative Greenhouse-Geisser critical F-value instead, which helped correct for violation of sphericity. Asterisks indicate significance.

Signal frequency		**Masker type**			**Precursor duration**		**Masker type * Precursor duration interaction**	
	F-statistic	p-value	Paired comparisons	p-value	F-statistic	p-value	F-statistic	p-value
2 kHz	*F*(2,32) = 56.269	p < 0.001*	Off-freq vs On-freq	p < 0.001*	*F*(2.823,45.168) = 58.865	p < 0.001*	*F*(10,160) = 1.425	p = 0.174
			No-M vs On-freq	p < 0.001*				
			Off-freq vs No-M	p = 0.019*				
4 kHz	*F*(2,16) = 39.635	p < 0.001*	Off-freq vs On-freq	p < 0.001*	*F*(3.014,24.116) = 60.339	p < 0.001*	*F*(10,80) = 0.832	p = 0.599
			No-M vs On-freq	p < 0.001*				
			Off-freq vs No-M	p = 0.116				

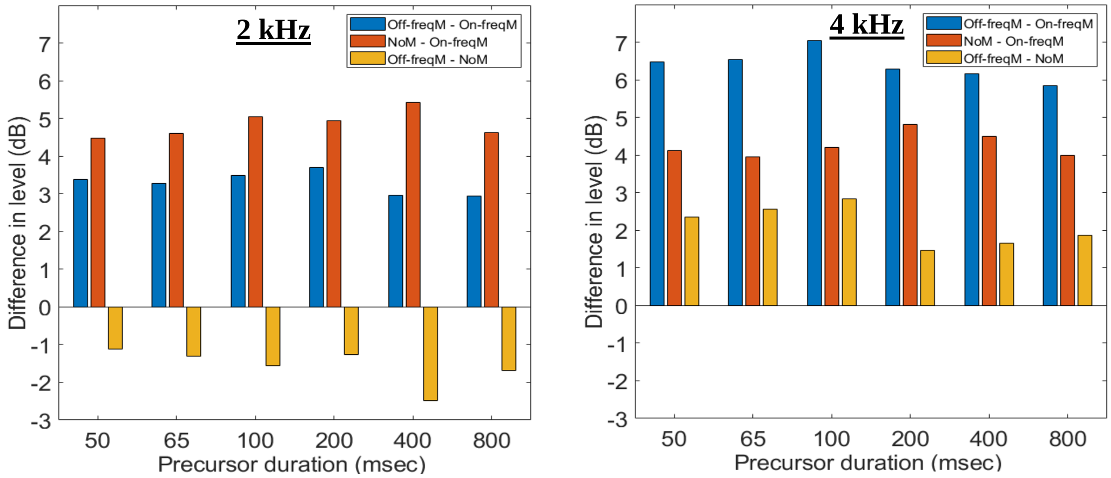

Figure 5. Average difference between listening conditions as a function of precursor duration for 2-kHz (left panel) and 4-kHz (right panel) signal frequencies. These differences were calculated from the data in Fig. 4. For all precursor durations, the off-frequency and no-masker conditions produced large significant differences compared to the on-frequency masked condition at 2 and 4 kHz. These results are consistent with the off-frequency and no-masker condition being estimates of gain reduction.

2.3.2 Time Constant Estimation

Time constants were estimated from both of the gain reduction growth functions (off-frequency masked and no-masker) in the previous section. An inverse exponential function was fit to individual and group data, and an overall time constant (τ) and a corresponding variance accounted for value were estimated. The formula for this function is,

$$Y(t) = Y_{max}(1 - e^{-t/\tau})$$

which estimates the time at which approximately 63% of the maximal effect of the growth function is achieved. In this formula, 'Y' represents the event or response of interest, 't' represents elapsed time, and τ represents the time constant. With respect to our behavioral experiments, Y is the signal threshold shift between the signal with and without the precursor condition (i.e., gain reduction). Then 't' equates to the total duration from the onset of the precursor to the onset of the signal. As mentioned in the behavioral stimuli section, there was always 20 milliseconds between the offset of the precursor and the onset of the signal, whether a masker was present or absent. Therefore, each data point in the time constant estimation corresponded to a total duration of the precursor plus 20 milliseconds to account for the masker or silence between the offset of the precursor and the onset of the signal ([50, 65, 100, 200, 400, and 800 msec] -> [70, 85, 120, 220, 420, and 820 msec]). Time constants estimated with this function have previously been used in human behavioral and physiological measures of overshoot (Walsh et al., 2010), an effect that has been posited to relate to gain reduction via the MOCR (e.g., Strickland, 2001).

In order for individual time constants to be included in the group averaged time constant, at least 60% of the variance in the data needed to be accounted for in the individual fit. Additionally, as reported in Walsh et al., (2010), some individual growth functions end up being flat (i.e., no effect of precursor duration on the signal), and therefore the time constant could not be derived due to having a horizontal slope, which would result in a time constant at or near zero milliseconds. These time constants were also not included in the group averaged time constant.

2.3.3 Psychoacoustic Time Constants

Individual data fitted with the function can be found in Fig. 6, where 2-kHz data is in the top portion, and 4-kHz data is in the bottom portion. The symbols are closed for off-frequency masked conditions and open for no-masker conditions, and the fitted curve was solid for off-frequency masked and dashed for no-masker conditions. The corresponding subject number for each subject in the experiment can be found in each cell of Fig. 6. Generally, these functions gradually build over the course of tens of milliseconds and then asymptote. This finding is reflected in most of the time constants for both signal frequencies, tabulated in Table 3. In some cases, in the 2-kHz (off-freq: S6, S10, S16, S19) and at 4-kHz (off-freq: S4, S5, S6, S7, S8; no masker: S11) conditions, there is an oscillating effect where the functions slightly increase and decrease as a function of elicitor duration. The latter observation has been documented before with tonal precursors that share the same frequency as the signal (on-frequency precursor), which has been modeled as the precursor turning down the gain at the signal frequency place and decreasing its own effectiveness (Roverud and Strickland, 2014). There was quite a wide range of values for individual time constants, from ~27 – 152 ms for 2 kHz conditions, and ~43 – 217 ms for 4 kHz conditions. This finding is similar to the wide range of time constants for individual subjects eported from the behavioral overshoot tasks in the Walsh et al. (2010) study (19.5 – 141.8 ms).

Off-frequency masked data for Subjects 10 and 16 at 2 kHz are shown as filled star symbols to indicate they were considered statistical outliers, and therefore not fitted to estimate a time constant. However, as stated earlier, their no-masker conditions met the assumption of normality, and time constants for that condition were estimated. Also noted earlier, flat functions were not fitted as they would produce time constants that approximate zero milliseconds, as was the case for subjects 6 and 19 for the 2 kHz off-frequency masked conditions. This was also true in the Walsh et al. (2010) study, where one of their subjects produced flat effects in their overshoot task and a time constant was not estimated (see their Fig. 7, subject KW).

The average threshold shifts fitted with the function can be found in Fig. 7. For 2 kHz, the average time constant for the off-frequency masked conditions (N = 15) was 84.68 ± 8.52 msec ($R^2 = 0.91$), and for the no masker conditions (N = 19) was 82.76 ± 8.52 msec ($R^2 = 0.96$). For 4 kHz, the average time constant for the off-frequency masked conditions (N = 9) was 61.82 ± 5.42 msec ($R^2 = 0.91$), and the no masker conditions (N = 9) was 103.70 ± 15.79 msec ($R^2 = 0.97$). It is clear that the 2 kHz time constants were similar, both approximately ~83 msec, while the 4-kHz listening conditions were slightly shorter (off-frequency) and longer (no masker) than the 2-kHz conditions. However, the average of the two time constants at 4 kHz is close to 80 ms and it is possible that the uneven sample size played a role in the overall difference between the two frequencies. Despite this, all of the constants were fairly short, where the buildup effect is maximal within approximately 100 milliseconds of elicitor activation. Lastly, the exponential function fit the individual and the average behavioral data quite well, with over 90 percent of the variance accounted for in each fit.

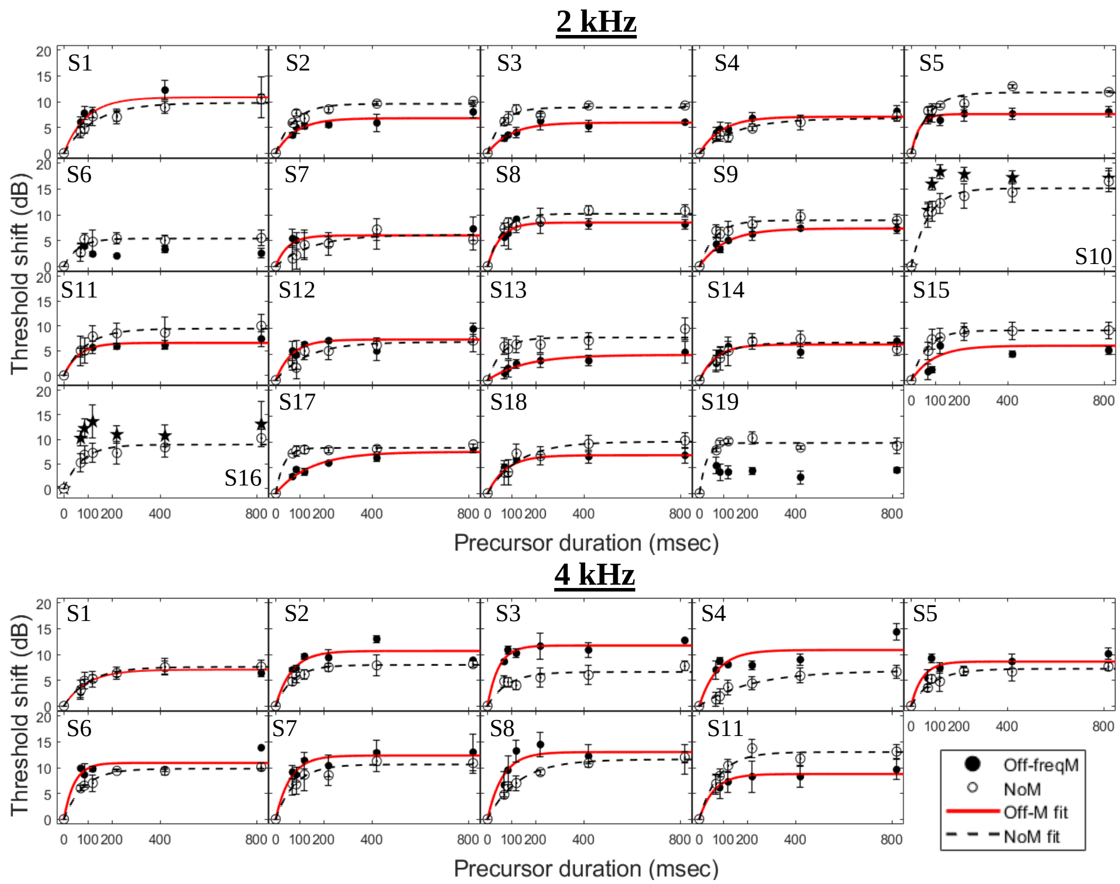

Figure 6. Signal threshold shifts as a function of precursor duration at 2 kHz (top) and 4 kHz (bottom). Each separate cell is an individual subject's data. These data are the same as in Fig. 3. Filled symbols are the off-frequency masked data, and open symbols are the no masker data. The solid and dashed curves are the exponential curve fits for the off-frequency masked and no masker conditions, respectively. SD of signal thresholds is indicated by the error bars.

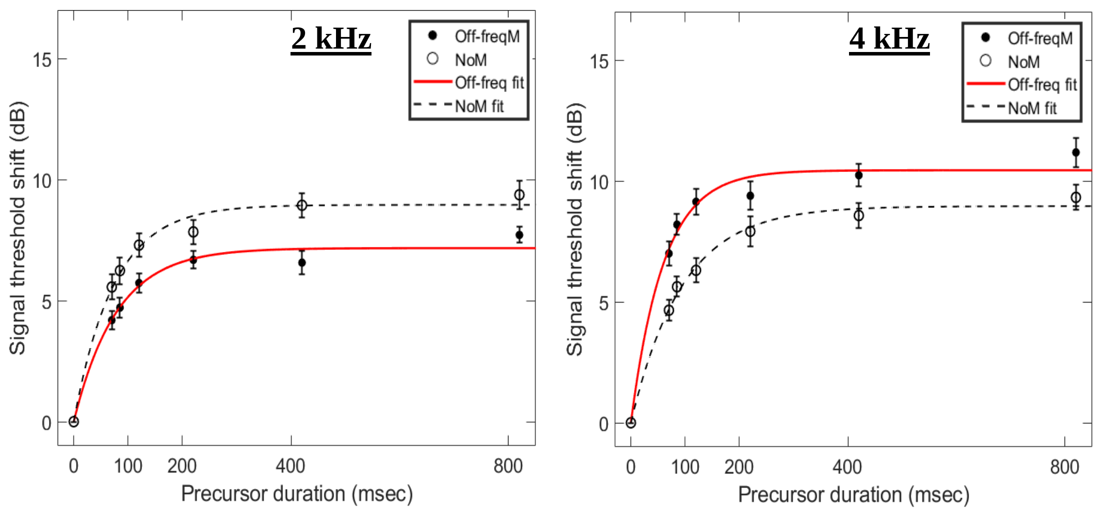

Figure 7. Average signal threshold shifts as a function of precursor duration for 2 kHz (left) and 4 kHz (right) signal frequencies, estimated from the data in Fig. 6 (also identically labeled). As noted, the averaged time constant for the off-frequency masked conditions at 2 kHz (left side - filled symbol and solid fit) had a sample size of N = 15, as subjects 10 and 16 were statistical outliers, and subjects 6 and 19 were not fitted due to an overall flat response with precursor duration (see Table 3). SEM of signal thresholds is indicated by the error bars.

Table 3. Individual and averaged time constants and corresponding variance accounted for (R^2) when fitting the exponential function for each listening condition. $\boldsymbol{\tau}$ units are in milliseconds, and the ± error of the averaged time constants is the SEM of the individual time constants. Yellow boxes indicate a time constant could not be estimated. Nine subjects participated in the 4 kHz conditions, and all had a corresponding time constant.

	2 kHz off-frequency (N = 15)		**2 kHz no masker (N = 19)**		**4 kHz off-frequency (N = 9)**		**4 kHz no masker (N = 9)**	
Subj	τ	R^2	τ	R^2	τ	R^2	τ	R^2
1	***91.08***	***0.92***	***125.57***	***0.96***	***95.67***	***0.95***	***104.47***	***0.98***
2	***84.38***	***0.92***	***72.83***	***0.96***	***66.76***	***0.91***	***72.37***	***0.99***
3	***94.47***	***0.96***	***57.05***	***0.96***	***46.35***	***0.97***	***82.4***	***0.89***
4	***92.2***	***0.93***	***152.05***	***0.96***	***73.34***	***0.81***	***217.22***	***0.98***
5	***33.52***	***0.97***	***70.1***	***0.96***	***47.47***	***0.86***	***90.29***	***0.96***
6	-	-	***58.44***	***0.91***	***42.91***	***0.88***	***80.8***	***0.99***
7	***50.61***	***0.85***	***151.59***	***0.89***	***59.91***	***0.96***	***78.58***	***0.97***
8	***52.58***	***0.96***	***59.37***	***0.96***	***63.24***	***0.91***	***127.19***	***0.99***
9	***108.16***	***0.97***	***64.36***	***0.99***	-	-	-	-
10	-	-	***70.35***	***0.98***	-	-	-	-
11	***61.69***	***0.97***	***91.1***	***0.97***	***60.72***	***0.97***	***79.99***	***0.97***
12	***66.4***	***0.85***	***130.24***	***0.94***	-	-	-	-
13	***153.73***	***0.94***	***58.7***	***0.92***	-	-	-	-
14	***67.72***	***0.89***	***84.01***	***0.94***	-	-	-	-
15	***93.63***	***0.66***	***63.29***	***0.98***	-	-	-	-
16	-	-	***73.31***	***0.94***	-	-	-	-
17	***144.69***	***0.95***	***32.99***	***0.99***	-	-	-	-
18	***75.3***	***0.95***	***128.64***	***0.99***	-	-	-	-
19	-	-	***28.53***	***0.97***	-	-	-	-
Average	***84.68 ± 8.52***	***0.91***	***82.76 ± 8.52***	***0.96***	***61.82 ± 5.42***	***0.91***	***103.70 ± 15.79***	***0.97***

CHAPTER 3: PHYSIOLOGICAL MEASURES OF GAIN REDUCTION

3.1 Introduction

The time course of the MOCR has been studied using OAE based paradigms (Puria et al., 1996; Kim et al., 2001; Bassim et al., 2003; James et al., 2005; Backus and Guinan, 2006). From these studies, the MOCR time course has been shown to buildup and decay over the course of hundreds of milliseconds, with a 25 ms onset and offset delay before the effects can build or decay, respectively (James et al., 2005). For studies using elicitors to activate the MOCR, broadband noise is typically used as it has been shown to be a powerful MOCR activator (Maison et al., 2000). These elicitors are presented contralaterally with respect to the probe ear, which avoids the effects of excitatory masking on the probe stimulus itself. These overall findings have led to the types of paradigms and stimuli used in both physiological and psychoacoustic studies of MOCR gain reduction: long or continuous elicitors (> 100s of milliseconds) that are presented contralaterally. As discussed in Chapter 2, it was unknown how broadband elicitor duration affected auditory perception. The next step in the present study was to measure MOCR time constants in an OAE based paradigm to compare to the psychoacoustic data. I hypothesized that there should be similarity between physiological and psychoacoustic gain reduction time constants, as they should share a common underlying mechanism: gain reduction mediated by the MOCR. In order to test this hypothesis, the same 19 subjects who completed the psychoacoustic experiments also completed a forward masking TEOAE paradigm that closely mirrored the psychoacoustic paradigms. This chapter details how the MOCR was measured from subjects using many different TEOAE parameters, and verifying that the effects were consistent with MOCR, and were not affected by MEMR or artifacts.

3.2 Methods

The physiological experiment detailed in this chapter used a forward masking TEOAE paradigm to measure MOCR magnitude using a range of broadband elicitor durations. One major part of the TEOAE analysis was to choose ‘optimal’ parameters to find the most stable and consistent MOCR effects, thereby providing the best responses to fit corresponding time constants. These are detailed in the sub-chapters of the Methods section (3.2). Changes in the TEOAE were measured in terms of magnitude and phase. Multiple methods were used to validate that the MEMR was not active during the TEOAE experiment, including the use of wideband acoustic immittance (WAI) tympanometry.

3.2.1 Subjects

All 19 subjects who completed the psychoacoustic experiment also completed TEOAE and WAI measures. The data from this chapter and Chapter 2 will be directly compared in Chapter 4.

3.2.2 Equipment and Calibration

All TEOAE and WAI measures were made with an ER-10X Extended-Bandwidth Acoustic Probe System (Etymotic Research, Elk Grove Village, IL) which utilizes integrated Forward Pressure Level (FPL) systems, allowing for the prediction of SPLs up to 20 kHz in humans. This system allowed for probe stimuli (clicks) and ipsilateral MEMR or MOCR-eliciting stimuli to be presented from separate speakers to limit interchannel interactions and distortions, and a microphone to measure sound pressure near the ear canal. The TEOAE and WAI stimuli were generated with two Tucker-Davis Technologies (TDT) RZ6 programmable DSP processors, and these signals were sent to and amplified by two separate TDT HB7 headphone drivers, which then fed to the ER-10X system and probe.

Before any TEOAE or WAI measurements were collected from a subject, the ER-10X probe needed to be calibrated by determining the Thevenin-equivalent impedance and pressure characteristics of the sound sources (Allen, 1992; Keefe et al., 1992). These measurements are necessary to determine the acoustic impedance at a location in the ear canal. The Thevenin-equivalent source and impedance for the click probe were estimated by measuring the acoustic response at the ER-10X microphone when the eartip was coupled to loads whose acoustic impedance values can be approximated using theoretical calculations (closed brass tube cavities of 8 mm diameter and five different lengths, which is the “ER-10X calibrator”). The estimates were refined until the so-called “calibration error” (a dimensionless energy ratio scaled by a factor of 10,000, and averaged over 2-8 kHz) was minimized. Error values of < 1 are typically considered good quality calibration (Neely and Liu, 1994); we consistently obtained errors within a range from 0.01 – 0.04. With the probe properties calibrated, the same click stimulus was then used to estimate the immittance properties of each subject's ear for each insertion of the probe tip. Because both TEOAE and WAI measures rely on ear canal probes that are tightly sealed in an individual's ear canal, probe tips that are not fully secure can cause air-leaks or poor fitting that compromise the validity of the measurements and calibrations. Air-leaks can cause changes of absorbance that increase with the degree of the leak, especially at lower frequencies where air leaks reliably cause an increase in absorbance (Groon et al., 2015). Therefore, the criteria of low-frequency (0.2 kHz) absorbance that was used to detect air leaks was less than 29%, and the admittance phase was greater than 44 ° but no greater than 90 °, as established by Groon et al. (2015).

3.2.3 Procedure

All subjects were tested in a single-walled sound booth in a chair and passively watched a muted video of their choice on Netflix or HULU with subtitles, and they did not need to respond to the sounds. Subjects were asked to remain still during the testing and were always allowed to take breaks as needed in between periods of data acquisition. For both TEOAE and WAI measures, the ER-10X probe was always fit to the subject's right ear except for subject 18. As stated in Chapter 2, subject 18's left ear was the test ear (probe ear) due to an audiometric threshold notch around the 2 kHz region in her right ear, but normal thresholds in her left ear. The ER-10X probe cables were fixed to the ceiling of the booth with a magnet to eliminate any contact with the subject's body and minimize any recorded noise due to movements. TEOAE and WAI experiments were made using a PC with custom MATLAB software (The MathWorks, Inc., Natick, MA). WAI measures and thresholds were always collected before beginning a TEOAE experiment, which allowed for quick verification that the elicitor stimuli would not activate the MEMR in the subsequent TEOAE and behavioral experiments. Thus, between the clinically measured acoustic reflex thresholds and the WAI MEMR thresholds, we had multiple ways to verify that the MEMR was not activated during our experiments, which increases the validity of our findings.

3.2.4 Wideband Acoustic Immittance (WAI) MEMR Measures and Stimuli

As noted in Chapter 2, clinical acoustic MEMR thresholds were measured as part of the standard audiological battery for testing subjects. These thresholds were measured using a 226 Hz probe and a broadband elicitor. Because the MEMR could affect higher frequencies and would affect the sound level going into the ear as well as the sound measured coming out, WAI measurements were made in each ear. WAI measures use click probes to assess middle-ear function over a wide range of frequencies (0.25 to 8.0 kHz). While typical clinical tympanometry can provide reliable acoustic reflex thresholds, they may also underestimate them by ~14 dB when compared to WAI MEMR thresholds (Feeney and Keefe, 2001; Feeney et al., 2017). MEMR thresholds were measured for each subject using a WAI paradigm adapted from Keefe et al. (2016), which was virtually identical to the paradigm used by Bharadwaj et al. (2021). A 90-dB peak-equivalent SPL (peSPL) click probe with a flat incident power spectrum in the 0 – 10 kHz range was used to measure the acoustic immittance properties of the ear-canal and middle-ear system. Each WAI MEMR measurement trial consisted of a series of seven clicks alternating with a 120-ms ipsilateral elicitor (including 5-ms $\cos^2$ onset and offset ramps) (see Bharadwaj 2021; Fig. S-3A). The elicitor was a pink broadband noise (0.5 – 8 kHz), and ranged from 34 – 88 dB FPL. The gap between the peak of the click and the onset of the noise elicitor (0 voltage point) was 27.94 ms, and the gap between the offset of the noise (0 voltage point) and the next click was 13.97 ms, for a total of ~41.91 ms between elicitor presentations (e.g., the click window). This trial structure was used for each elicitor level and the level was repeated 32 times with an inter-trial interval of 1.5 seconds to allow for the MEMR to relax back to baseline levels. For each elicitor level, the immittance measured using clicks numbered two through seven in the sequence were averaged together and the change relative to the first click was calculated as the WAI MEMR metric. The dB-change in ear-canal pressure induced by the MEMR was quantified as a function of frequency to yield a pattern of alternating negative and positive peaks at different frequencies. Additional calibrations which help reduce extraneous variance were leveraged in these estimations (described in the Physiology Equipment & Calibration section). The dB-change in ear-canal pressure was added to the dB-change in ear-canal conductance to yield a prototypical pattern of dB-change in absorbance power (see Keefe et al., 2017, their Fig. 6). The absolute value of this dB-change in absorbed power was averaged over 0.5 – 2 kHz to yield a single number per elicitor level (Fig. 8). A thresholding procedure was used to reject artifactual trials before averaging. Fig 8 shows ipsilateral WAI MEMR data from a single subject to illustrate how WAI MEMR thresholds were estimated. The top panel shows the shifts in power absorbance (ΔA) across frequency as a function of ipsilateral elicitor level, relative to the absorbance of the click alone. The bottom panel shows the shift in absorbed power as a function of elicitor level over a narrow frequency band (1/3rd-

octave band centered at 1 kHz), derived from the data in the top panel. A change in absorbance power greater than 0.1 dB (dashed line) indicates significant MEMR activation (WAI MEMR threshold). This subject's ipsilateral WAI MEMR threshold is approximately 64 dB SPL with the pink noise elicitor, shown in the bottom panel. WAI MEMR thresholds were collected for all subjects in the study. The total duration of this experiment was approximately 16-17 minutes for a single subject.

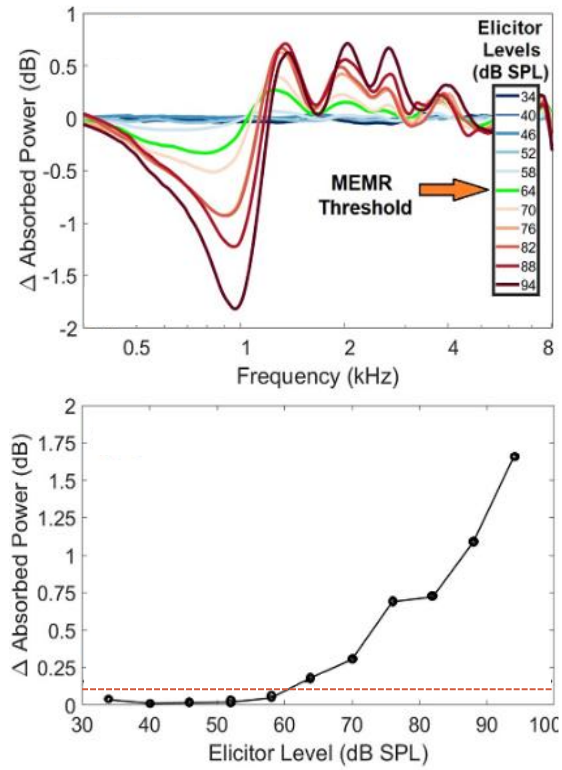

Figure 8. WAI data with an ipsilateral elicitor from a single subject. Top: Change in absorbed power across frequency. Bottom: Change in absorbed power in a 1/3rd octave band at 1 kHz. The level at which the fitted function crossed 0.1 dB is considered the WAI MEMR threshold in the current study.

3.2.5 TEOAE Stimuli

In this experiment transient evoked otoacoustic emissions (TEOAEs) were measured with and without the presence of a pink noise elicitor of various durations to estimate the temporal properties of the MOCR. To generate cochlear emissions, we used 54 dB peSPL click probes, with a flat incident power spectrum in the 0-10 kHz range. Using a relatively low-intensity click should place the response on the lower leg of the IO function where gain reduction should be largest (Cooper and Guinan, 2006), and parallels the design of psychoacoustic experiments. This click level was chosen to generate emissions significantly higher than the noise floor to provide a good signal-to-noise ratio, while likely being too low to elicit the MOCR or MEMR. This was an important consideration in the design, as it has been shown that commonly used stimuli such as clicks (and other probe stimuli) can elicit the MOCR and the MEMR at relatively moderately high sound levels (Guinan et al., 2003). MOCR or MEMR activation would confound interpretation of the actual effect of the elicitor on the OAE. For example, the efferent activity evoked by the probe may increase the overall effect of the elicitor by adding the efferent activation from the probe and elicitor together, or, it is possible that the elicitor becomes less effective because the efferent activity turns down its response. Another scenario is that the efferent effects may fully saturate from probe activation, and the response from the elicitor on the probe is negligible. One previous study using a very similar TEOAE paradigm used similar click levels (Mertes and Goodman, 2015; mean RMS click level 55 dB SPL) to the current study, as have other studies with using different TEOAE paradigms (Boothalingam and Purcell, 2015; Boothalingam et al, 2018; 55 dB peSPL) and have generated TEOAEs with large signal-to-noise ratios. The elicitor was a 50-dB SPL pink broadband noise (0.25 – 8 kHz) which varied in duration [50, 100, 200, 400 ms, including 5-ms $\cos^2$ onset and offset ramps], and preceded the probe click by 5 ms. These elicitors were essentially identical to those used in the behavioral experiment in the overall sound pressure level, durations, and frequency content (pink noise), and should provide a fair comparison between the two data sets. The 65-ms and 800-ms elicitor durations that were used in the behavioral experiments were omitted because it would have considerably increased the time necessary to complete data collection. This elicitor level is slightly lower in sound level compared to many studies evoking the MOCR, which typically use 60-dB SPL elicitors, to ensure that the MEMR would not be elicited. Both the click and the elicitors were generated using a sampling rate of 48,828.125 Hz. This sampling rate satisfies the Nyquist Theorem requirement of choosing a sampling frequency double that of the highest frequency components of the stimuli.

3.2.6 TEOAE Testing

Fig. 9 shows a schematic of the stimulus conditions used in the TEOAE experiments, and introduces terms that will be used throughout the text. Stimuli consisted of clicks to generate TEOAEs and pink broadband noise of different durations to activate the MOCR. In the top panel of Fig. 9, the click is presented approximately 405 milliseconds from the onset of the stimulus presentation window. This click alone condition will generate a TEOAE without the influence of an elicitor, and is the baseline condition of the experiment. The total duration of this window is approximately 451 ms, which includes the pre- and post-click duration. A single presentation of the click (top panel) or the click plus elicitor (middle and bottom panels) is termed a single “buffer”. Importantly, as shown in Fig. 9, the buffer window was always fixed at this duration for any of the listening conditions. Notice in this figure, the relative temporal position of the click onset, the duration post-click, and the elicitor offset relative to the click onset never change from one condition to another. However, it is the onset of the elicitor that adjusts relative to the onset of the buffer window depending on how long the elicitor duration is. For example, for the 50-ms elicitor condition (middle panel), 350 ms precede the onset of the elicitor. Then, for the 400-msec elicitor condition, there are zero ms between the onset of the elicitor and the onset of the buffer window. In the current study, MOCR effects for a given listening condition were always recorded in a 21 msec post-click window, which will be called the “MOCR response window”. This window starts 5 ms post click and ends approximately 26 ms post click. The response from the first 4-ms including the click was not included, and only effects from within the MOCR response window will be averaged in the overall output.

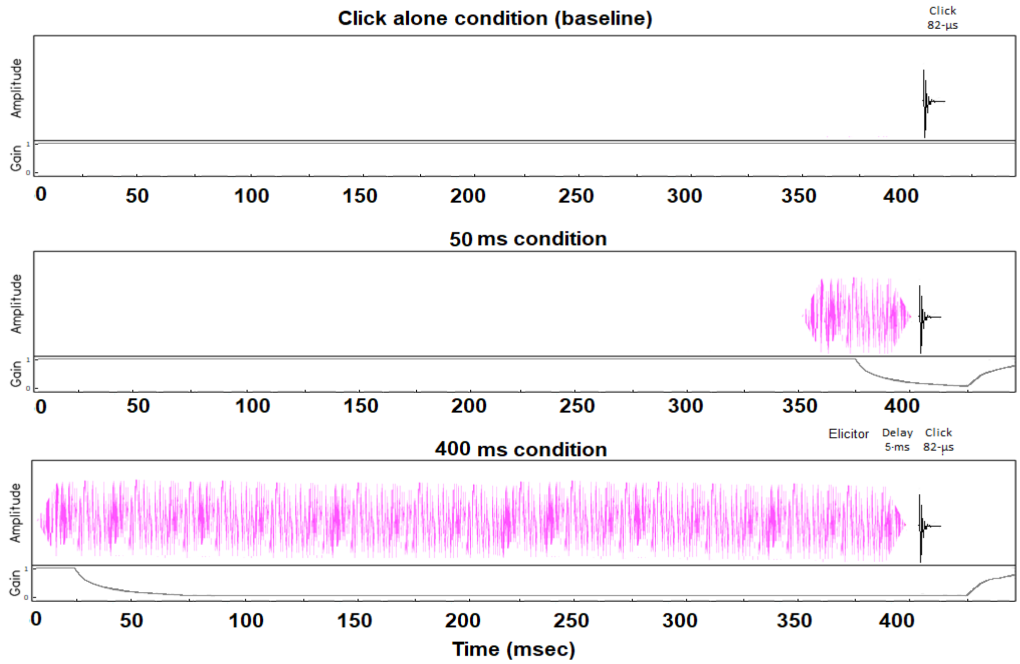

Figure 9. Configuration for the ipsilateral forward masking TEOAE experimental conditions (top of each panel), and a hypothetical change in gain (bottom of each panel). Each panel shows a listening condition, where the top panel is the no-elicitor condition (click alone; baseline), the middle panel shows a condition with a 50-ms elicitor (shortest elicitor), and the bottom panel shows a condition with a 400-ms elicitor (longest elicitor). Note that the offset of the pink broadband noise elicitor (except for the no-elicitor condition) and the onset of the click are consistently placed at their temporal location within the buffer for all listening conditions. The total duration of each buffer is ~451 ms.

Figure 10 illustrates the presentation of the listening configurations (no elicitor, 50 ms elicitor, 100 ms elicitor, 200 ms elicitor, and 400 ms elicitor) in the TEOAE experiment. For each listening condition, there were 40 sequential buffer presentations, and this constituted a “block” of buffers. There was always two seconds of silence following the completion of a block, in an effort to ensure that short-term MOCR effects would not affect the next block. The baseline condition (click only) was always the first block in the sequence, followed by a block with the shortest elicitor (50 msec), and elicitor duration increase in each subsequent block. This sequence of 5 blocks is called a “set”. A total of 16 sequential sets were completed for a single “array” of TEOAE data, which took approximately 27 minutes. A total of 8 arrays of data were collected from each subject to complete the experiment, which required 3.5-4 hours of total data collection time. Typically, a subject would complete the entire experiment over 2-3 lab visits, and a 15–20-minute break was always given if more than two arrays were collected in a single visit. The first 8 buffer presentations were discarded from each block (leaving 32 responses in the block) to allow the microphone and the amplifier settle and avoid artifact responses. This results in 512 MOCR responses for each listening condition for averaging in the single "array" of data. Therefore, completing the experiment (8 arrays of TEOAE data) equated to an average of 4096 responses for each listening condition.

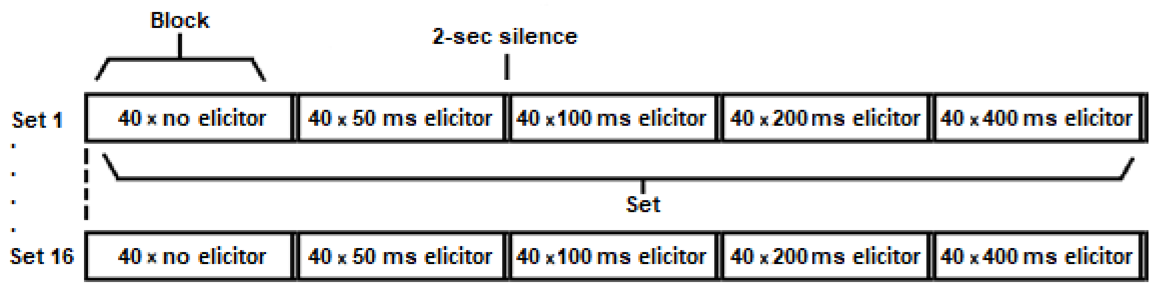

Figure 10. Schematic of the presentation of conditions (a single array of data) in the current study. Each block was presented sequentially, followed by 2 seconds of silence and the elicitor duration always increased across conditions. Completion of blocks in this order forms a set. Completion of 16 sets, as depicted here, would generate one array of MOCR responses. For each block, the initial 8 buffers were removed leaving 32 buffers to avoid potential microphone and amplifier artifacts. Therefore, each array always provided 512 total responses per condition. In the experiment, a total of 8 arrays of data were collected, resulting in responses averaged from 4096 buffers per condition.

Additionally, a direct verification of MEMR activation during TEOAE recordings was implemented in which time window containing the click (-0.2 to 1.16 msec) was monitored. A difference in sound pressure level (larger than 0.1 dB) between the condition with an elicitor and a condition with the click presentation by itself (baseline) would indicate potential MEMR activation by the elicitor. This is because the MEMR would cause some of the click energy to be reflected back to the microphone, particularly at the lower frequencies. A similar test has been implemented in a previous study (Mertes and Goodman, 2015). I tested this between the longest duration elicitor (400 ms), the most effective elicitor used in the TEOAE experiment, and the baseline (click only) condition. Fig. 11 shows an individual subject's time windowed response to the averaged click alone and the click in the 400-ms elicitor condition.

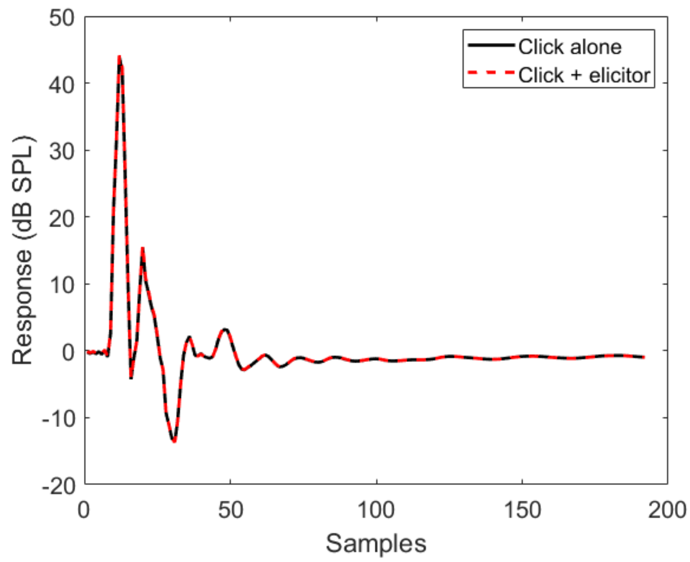

Figure 11. Results for a single subject (S3) from the TEOAE MEMR verification test windowed to just before and after the click stimulus presentation. The responses shown were averaged over the complete TEOAE data set (i.e., 4096 responses for each condition). Notice that the click does not change in level in the presence of the elicitor, as they are completely overlapped. The results show that the MEMR was not active during the TEOAE experiment for this subject.

3.2.7 Optimal Parameterization #1: TEOAE Magnitude and Magnitude Plus Phase

It will be necessary to define certain terminology and abbreviations in order to understand MOCR effects from the TEOAE data, so they will be explained here. MOCR strength is most commonly estimated by taking the arithmetic difference in ear canal sound pressure level between the OAE amplitudes with and without the presence of an elicitor. However, this metric does not include any of the phase information in the response. MOCR effects from OAEs that include phase in the response may provide more reliable responses (Backus and Guinan, 2007; Marshall et al., 2014). Additionally, there is still debate about OAE phase and its interpretation in relation to the MOCR (e.g., Francis and Guinan, 2010; Lilaonitkul and Guinan, 2012), and it is not clear which metric is more relevant for perception. Therefore, MOCR effects were evaluated as both changes in magnitude alone and magnitude plus phase of the TEOAE. To do so, MOCR effects were estimated as the vector difference between the TEOAE for the click alone and the TEOAE with an elicitor present. These differences were represented either in terms of change of the TEOAE magnitude alone ($\Delta TEOAE_{m}$) or change of the TEOAE magnitude plus phase ($\Delta TEOAE_{m+p}$), both schematized in Fig. 12. These metrics have been used to investigate MOCR effects from stimulus-frequency otoacoustic emissions (e.g., SFOAEs; Guinan et al., 2003; Backus and Guinan, 2006, 2007; Lilaonitkul and Guinan, 2012) and transient evoked otoacoustic emissions (e.g., TEOAEs; Mertes and Goodman, 2015), or both SFOAE and TEOAEs in a single study (Marshall et al., 2014). Interpreting MOCR effects with these metrics from studies using either SFOAEs or TEOAEs is a fair comparison as both measures share a common reflection generation mechanism (Shera and Guinan, 1999).

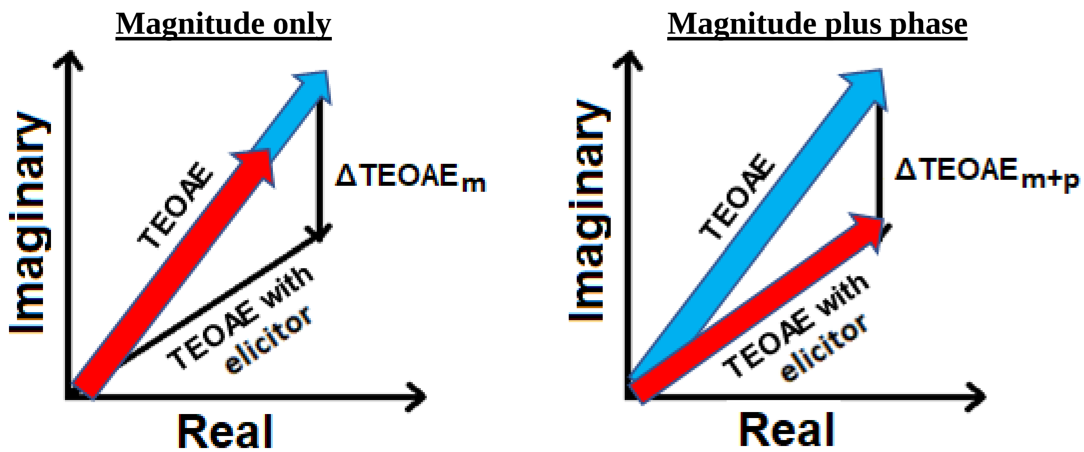

Figure 12. Vector diagram of MOCR effects on sound pressure in the ear canal for transient evoked otoacoustic emissions (TEOAE) measurements. The arrows are in the complex plane with length representing magnitude and direction representing phase. The left diagram is $\Delta TEOAE_m$ and the right diagram is $\Delta TEOAE_{m+p}$. In both diagrams, the longest arrow (blue) represents the TEOAE to the click presentation alone, which is the baseline condition. The shorter arrow (red) represents the TEOAE when the elicitor is present, due to MOCR-induced change in the TEOAE magnitude or magnitude plus phase. The change in the TEOAE is the vector difference between these two arrows, indicated by ΔTEOAE.

3.2.8 Optimal Parameterization #2: A Multitaper Approach to Estimate Power Spectral Density of the TEOAE

In this study, I was particularly interested in observing the changes of the TEOAE waveform (ΔTEOAE) within a frequency channel, and to quantify these changes as a function of elicitor duration. To do so, a multitaper spectral analysis was used to estimate the power spectral density (PSD) of the TEOAE waveforms in the current experiment. Generally, the goal of spectral density estimation is to separate a signal into its different frequency components and estimate the power of the signal across these frequency bands. To do this in the frequency domain, it is assumed 1) that the sample is infinite in duration, 2) the signal is periodic, and 3) the signal can be decomposed into pure sinusoids. However, most real data do not follow these ideal conditions, where real data is typically finite, discrete, aperiodic, and time-varying. Because of this discrepancy, finite data (real data) typically produces estimates that differ greatly from the ideal spectral estimates, and are both 'biased' and suffer from high error 'variance' (Babadi and Brown, 2014). For example, data that is finite in duration consequentially leads to an inaccurate (i.e., biased) frequency representation, where global frequency interactions occur, and spectral splatter misrepresents the signal's true PSD at each frequency band. This is especially true for analyses that use rectangular windows at the start and end of the data in their spectral estimation, which has consequences in the frequency domain because of these abrupt transitions. The other issue in PSD estimates is high variance at a given frequency. It seems intuitive that a single sample of data may not represent the characteristics of the population or system that is being studied, which is why many samples are typically collected before analysis. Therefore, more samples in the analysis tends to reduce some of the variance and provide a better estimate of the signal of interest. However, PSD estimation of a discrete and finite signal may use methods that tend to be variable and noisy. The reason for this is that while increasing the sample size of the data does bring the estimate of the PSD closer to the to the "true" PSD of the signal, nonintuitively, increasing the sample size does not decrease the variance of the estimate for some methods (see Todd and Cruz, 1996). One last issue is that trying to correct for the high variance can lead to an increase in bias, or conversely, trying to reduce the bias can increase the variance. This well-known variance-bias tradeoff is always a concern in signal processing, where resolution (frequency or temporal) and variance of the signal of interest are pitted against each other.

Fig. 13 shows TEOAE data from a single subject estimated in terms of raw amplitude (in dB SPL) as a function of frequency for all the listening conditions. Data in the left panel shows the TEOAE response across a broad frequency range to illustrate the 'OAE fine structure', which are the quasi-periodic fluctuations in the OAE amplitude and phase. In the right panel, the variance of the OAE fine structure is even more apparent when zoomed in. It is hard to determine if there is a clear trend of a reduction in the TEOAE magnitude with elicitor duration as the relative shifts in amplitude fluctuate somewhat randomly with elicitor duration. This could be partly explained by the variance in the data (OAE fine structure) and when analyzing raw (unsmoothed) data.

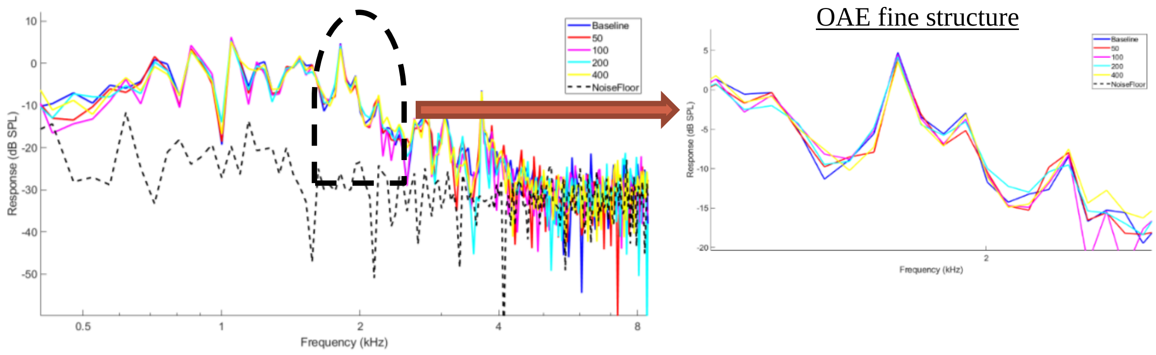

Figure 13. Raw TEOAE responses (dB SPL) across listening conditions from a single subject. Left panel shows the data across a wide frequency range, and the right panel shows a segment of the data across a 1/3rd octave band centered at 2 kHz to highlight the TEOAE fine structure. Note that the responses are variable across this narrow frequency region and may not be a wholly accurate estimation of the PSD of the TEOAE responses.

In order to reduce some of this variance in the TEOAE data, I used a multitaper spectral analysis to estimate the PSD of the signal. The multitaper analysis is a non-parametric method that can reduce both the bias and variance issues described above (see Babadi and Brown, 2014; Prerau et al., 2015). As mentioned before, finite data has the inherent issue of biases due to spectral splatter that is introduced when analyzed in the frequency domain with a rectangular window. When a multitaper approach is used, a taper (or window function) is applied to the finite data before performing spectral estimation, and as the name implies the taper reduces or "tapers" the spectral splatter of the side lobes because the taper smooths the discontinuities at the data ends and increases the energy concentration near the main lobe (see Prerau et al., 2015; their Fig 4 for an excellent visual description of this process). The general steps of the multitaper analysis involve some noisy data (i.e., the signal of interest), and the various tapers selected are then multiplied to the original signal, thereby estimating multiple single-taper spectra of the original signal. Then the mean of the single-taper spectrum is taken to calculate the multitaper spectral estimate of the signal of interest.

A key factor in the multitaper method is that it uses a selected number of these special taper functions, which are orthogonal to one another. This reduces the variance in the estimation, however, there is still a variance-bias tradeoff. The general parameters of the multitaper method involve: N = the size (length) of the data, TW = the time-half-bandwidth product, and L = the number of tapers. Understanding these parameters allows control of the features of the multitaper estimate. To select the number of tapers for the analysis, a Δf is needed where, Δf = the bandwidth (Hz) of the main lobe in the spectral estimate, which controls the minimum distance between peaks that can be resolved. A large Δf will produce smooth, low-resolution peaks, and in contrast, a small Δf will produce higher resolution peaks with greater detail. With the Δf and N, time-half-bandwidth (TW) can be calculated as,

$$TW = \frac{N\Delta f}{2},$$

where, TW > 1. The time-half-bandwidth product is a parameter used in calculating the number of tapers that relate the frequency resolution to the data window size, and is the product of the data length (N) and half the bandwidth of the main lobe (Δf /2). Lastly, the number of tapers, L, must be slected. The best adapted version of the method (Thomson, 1982) is that L = 2TW – 1, where N = the length of the data, and W = is the frequency resolution. Therefore, the number of tapers applied to the signal is indirectly controlled by time-half-bandwidth product value.

For this experiment, I manipulated the parameter of the time-half-bandwidth product, and thus the overall number of tapers that are applied to the TEOAE data. The reason for this is that an accurate estimate the TEOAE time constant is desired, and the time constants are estimated based on the magnitude of the data points from ΔTEOAE. Depending on the number of tapers applied, there may be more or less power of the frequencies being studied. Therefore, I hypothesized that the number of tapers applied may affect the overall results in the study. Any unnecessary variance of bias may directly affect the ability to fit a time constant as well as the overall reliability of the fit to the data (r-squared). To test this hypothesis, the TEOAE was studied with a time-half-bandwidth product of 1.5 and 5 (2 tapers and 9 tapers respectively). Fig. 14. shows the TEOAE data shown earlier (Fig. 13) for only the baseline and the longest elicitor condition. The panel on the left shows the data when 2 tapers were applied, and the panel on the right is the same data with 9 tapers applied. Note that in either case, the TEOAE fine structure is smoother compared to when no tapers were applied (Fig. 13) As stated above, fewer tapers may provide higher frequency resolution but at the cost of more variance (left panel), whereas more tapers will reduce the variance but at the cost of poorer frequency resolution (right panel).

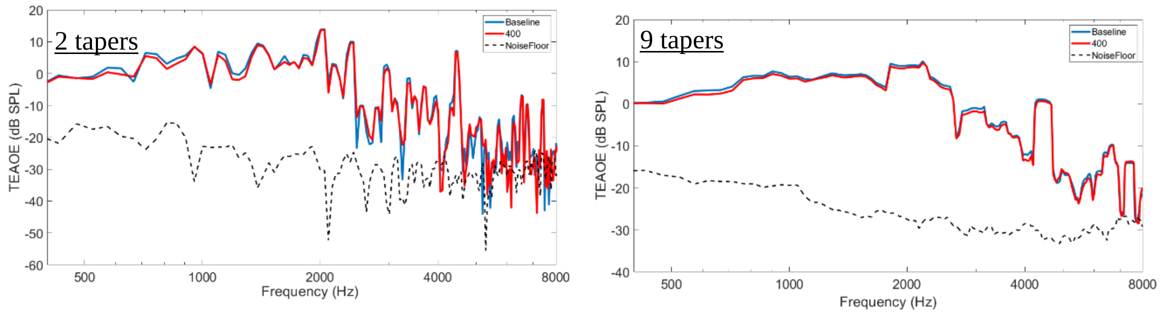

Figure 14. Multitapered TEOAE data for the click alone (baseline) and longest elicitor (400 ms) conditions, with either 2 or 9 tapers applied, shown in the left and right panels respectively. These data are from the same subject as in Fig. 13. The OAE fine structure is more apparent with less tapers applied, and much less apparent with more tapers applied.

3.2.9 Optimal Paramaterization #3: TEOAE Frequency Analysis: Narrowband, Wideband, and Nearest Spectral Peak

One last factor in the TEOAE analysis was to choose a frequency band to measure ΔTEOAE from. Multiple OAE studies have estimated the MOCR shift and stability in different restricted frequency bands and have found that the effect depends on OAE frequency (Goodman et al., 2013; Marshall et al., 2014; Mishra and Abdala, 2015; Mertes and Goodman, 2016). The reported MOCR shifts in the studies using TEOAEs (Goodman et al., 2013; Marshall et al., 2014; Mertes and Goodman, 2016), generally found significant and stable MOCR shifts for the majority of their subjects in the 1-2 kHz frequency region. It has also been shown that MOCR shifts with reflection based emissions (SFOAE and TEOAE) have less reliable MOCR effects for frequencies > 4 kHz compared to lower frequencies (Lilaonitkul and Guinan, 2012; Goodman et al., 2013), which may be partly due to overall lower SNR of the OAE signal.

Studies comparing behavioral to physiological measures of gain reduction have often used OAE measures that vary drastically in the frequency analysis bands used to estimate the MOCR effect. For example, some studies measured the change in the OAE with a narrow frequency range or a single frequency (Kawase et al., 2000; Keefe et al., 2009; Walsh et al., 2010), a wide range of frequencies (Micheyl et al., 1997; Fletcher et al., 2016), or use of both narrow- and wideband analyses (Wicher and Moore, 2014; Marrufo-Perez et al., 2021). Because there are expected differences in the MOCR effect across different OAE frequency bands, as noted above, it may not be surprising that there are mixed results when trying to predict the behavioral responses based on the physiological responses (summarized in section 1.7).

Therefore, to determine which TEOAE frequencies are best to compare to the behavioral results, I used 3 different TEOAE frequency regions to measure MOCR effects from; a wideband analysis (0.001 – 5 kHz), a narrowband analysis (1/3rd octave band) centered at 2 kHz, and a nearest spectral peak analysis relative to 2 kHz. As stated above, 2 kHz is a region with larger expected MOCR effects compared to higher frequencies (i.e., 4 kHz; e.g., Lilaonitkul and Guinan 2012), and provides a comparable condition to compare with the 2 kHz psychoacoustic results (see Chapter 4). Previous studies have used a 1/3rd octave band to approximate human auditory filters (e.g., Goodman et al., 2013). Fig. 15 depicts these conditions in the spectral magnitude domain from a single subject (identical data from Fig. 14). The use of a nearest spectral peak analysis involves estimating the sound pressure level or power of the signal at a nearby peak relative to the frequency region of interest (Shera and Bergevin, 2012; Abdala et al., 2014). This method may be particularly useful in OAE research where there is well known OAE fine structure where dips or troughs may occur at the frequency region of interest eliminating any possibility of a recordable response (≥ 6 dB above the noise floor). The wideband analysis extended from 0.001 up to 5 kHz, where the SNR of the TEOAE rolled off into the noise floor for most individuals.

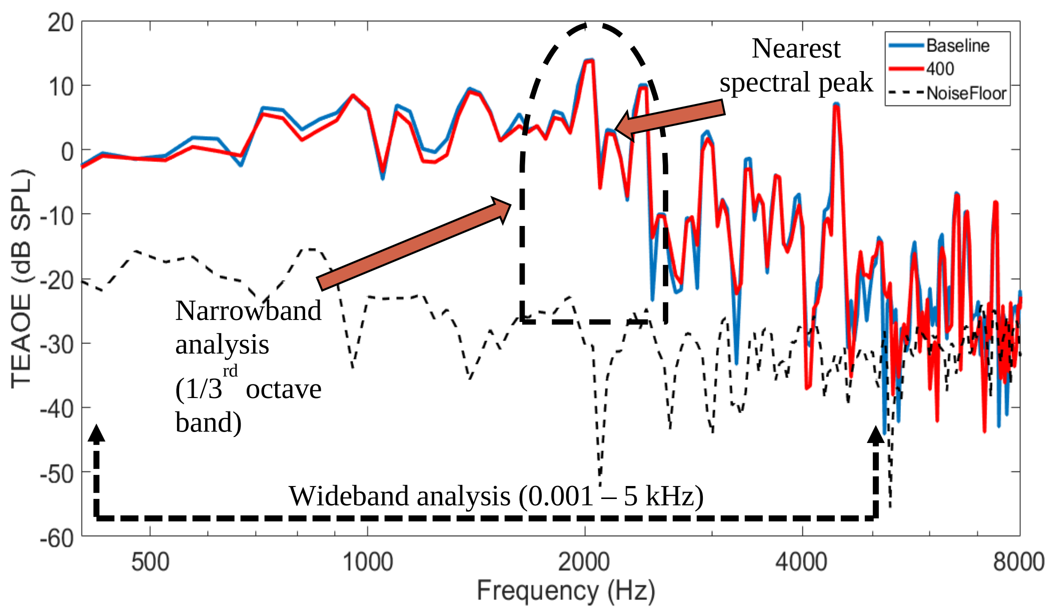

Figure 15. Three different frequency regions used to estimate ΔTEOAE: A wideband analysis (WB: 0.001-5 kHz), a narrowband analysis (NB: 1/3rd octave band), and nearest spectral peak (NP: at one single frequency). TEOAE data is identical the data in Fig. 14.

3.3 Results and Analysis

3.3.1 Clinical, WAI, and TEOAE Measured MEMR Thresholds

Table 4 shows ipsilateral clinical and WAI MEMR thresholds for each subject, as well as the results from the simultaneous MEMR measurement from the TEOAE experiment. These three separate measures are reported to validate that the MEMR was not evoked by the elicitor during the psychoacoustic or TEOAE experiments. In either case, no subject had an MEMR threshold below 50 dB SPL or FPL. Additionally, no subject had a change above 0.1 dB in the click with the presence of the longest elicitor (400 ms). The results from the independent measures indicate that the MEMR was highly unlikely to have been activated by the elicitor during the experiments.

Table 4. Individual MEMR thresholds measured using three techniques. In the first column, MEMR thresholds (dB SPL) measured clinically with a 226-Hz probe and a white broadband noise elicitor are shown. In the middle column, and MEMR thresholds measured using WAI (dB FPL) are shown. All thresholds were measured from right ears except for subject 18 which was made from the left ear.

Subject	Clinical MEMR threshold (dB SPL)	WAI MEMR threshold (dB FPL)	TEOAE measured change in click (dB)
S1	68	67	-0.02
S2	83	60	-0.01
S3	83	57	0.06
S4	83	60	-0.10
S5	83	80	-0.02
S6	68	52	0.02
S7	93	64	0.06
S8	78	65	0.03
S9	78	64	-0.01
S10	88	85	-0.03
S11	88	74	0.01
S12	93	70	0.01
S13	83	66	0.00
S14	73	52	0.03
S15	83	63	0.02
S16	83	60	0.03
S17	93	74	0.04
S18	78	60	0.03
S19	88	56	0.01

3.3.2 Time Constant Estimation

Time constants were estimated for both $\Delta TEOAE_m$ and $\Delta TEOAE_{m+p}$ for individuals and grouped data using the same inverse exponential function that was fit to the psychoacoustic data in Chapter 2. The time constant function estimates the time needed to reach approximately 63% of the maximal effect of the growth function. Each data point in the time constant estimation corresponded to a total duration of the elicitor plus 5 milliseconds to account for the masker or silence between the offset of the precursor and the onset of the signal ([50, 100, 200, 400 ms] -> [55, 105, 205, and 405 ms]). As for the psychoacoustic data, time constants needed at least 60% of the variance to be accounted for in the individual fit in order to be included in the group averaged time constant. Additionally, as reported in Walsh et al., 2010, some individual growth functions end up being flat (i.e., no effect of precursor duration on the signal), and therefore the time constant could not be derived due to having a horizontal slope, which would result in a time constant at or near zero milliseconds. These time constants were also not included in the group averaged time constant.

3.3.3 Optimal Parameters for Estimating Time Constants

As in the psychoacoustic data (section 2.3.3), time constants were fit with an exponential function to estimate their time constant for individual and grouped TEOAE. However, the first step was to test the 'optimal parameters' in the ΔTEOAE, and determine which parameters provided the most reliable data. Reliable here means the number of subjects with fitted data, and highest R^2 values (60% or higher). The parameters tested include: TEOAE magnitude with or without phase included ($\Delta TEOAE_m$ and $\Delta TEOAE_{m+p}$), differing numbers of tapers (2 and 9), and estimating the TEOAE with different frequency bands [narrowband (NB), wideband (WB), and nearest spectral peak (NP)].

Table 5 reports the individual time constants (τ) and the corresponding variance accounted for (R^2) for the $\Delta TEOAE_m$ conditions. The data is grouped by the number of tapers used in the analysis (2 or 9), and by the frequency method (NB, WB, and NP). As in the behavioral section, reasonable fits were not possible for some data. Also, functions were not able to be fitted if the ΔTEOAE was approximately zero, which would put the function near the baseline (i.e., no effect).

For these two scenarios, the '-' symbol is reported. These conditions were marked in yellow, as were any time constant with an R^2 value that was below 60%.

Table 5. Individual time constants (τ) and R^2 values for ΔTEOAE magnitude only. Time constants were estimated for narrowband, nearest-peak, and wideband spectral analyses. Nearest spectral peak had the poorest overall success rate, as the time constants could only be estimated for about 50% of the subject pool. Narrowband and wideband analyses both had higher success rates, as well as very high R^2 values (~80% and ~85% respectively). Yellow boxes indicate a poor fit (R^2 below 60%), or when time constants were unable to be estimated.

TEOAE mag only (1.5 TW) 2 tapers						
	NB (1/3rd oct band)		Nearest-peak		Wideband	
Subj #	τ	R^2	τ	R^2	τ	R^2
1	29.63	.24	83.5	.64	154.25	.96
2	-	-	-	-	52.67	.99
3	-	-	91.55	.23	392.98	.95
4	180.5	.92	129.3	.95	78.55	.98
5	37.91	.81	-	-	-	-
6	135.1	.99	143.2	.99	132.44	.99
7	106.7	.97	58.34	.89	79.47	.97
8	105.9	.80	-	-	98.28	.91
9	173.8	.99	141	.55	711.58	.98
10	136.9	.92	-	-	38.06	.96
11	320.8	.96	95.03	.54	65.38	.89
12	146.8	1.00	-	-	130.89	.87
13	516.5	.84	-	-	77.15	.97
14	36.87	.71	35.31	.38	87.08	.94
15	82.85	.76	-	-	49.67	.93
16	66.54	.97	202.1	.97	49.67	.93
17	-	-	33.46	.24	59.07	.99
18	-	-	-	-	26.05	.85
19	136.7	.98	253.4	.31	33.56	.96
Total N	14	.90	9	.68	18	.95

TEOAE mag only (5 TW) 9 tapers						
	NB (1/3rd oct band)		Nearest-peak		Wideband	
Subj #	τ	R^2	τ	R^2	τ	R^2
1	66.29	0.34	24.06	42.61%	151.2	95.82%
2	68.55	0.97	-	-	55.55	99.25%
3	-	-	-	-	356	94.36%
4	129.11	0.93	155.5	94.47%	94.69	98.26%
5	33.75	0.86	29.5	87.98%	-	-
6	142.15	0.99	147.3	99.21%	115	99.20%
7	114.76	0.97	135.3	96.55%	69.71	93.59%
8	123.29	0.96	137.4	67.53%	116.8	89.40%
9	373.02	0.99	199	98.93%	-	-
10	-	-	884.1	87.80%	28.57	84.36%
11	79.82	0.99	80.71	99.60%	55.95	84.93%
12	-	-	-	-	167.9	91.57%
13	319.07	0.92	-	-	83.21	95.14%
14	68.44	0.65	61.05	53.21%	-	-
15	96.72	0.69	54.5	44.45%	61.7	82.99%
16	43.29	0.87	22.57	81.39%	61.7	82.99%
17	660.73	0.95	350.7	93.30%	64.05	93.28%
18	61.11	0.84	-	-	26.45	85.49%
19	126.18	1.00	137.2	99.64%	33.97	98.20%
Total N	15	.91	11	.91	16	.92

There are two major points about the results of the table. The first point is that the number of tapers had little to no effect on the number of successful fits across subjects. For example, 14 subjects had time constants with the NB analysis with 2 tapers, while 15 subjects had time constants with the NB analysis with 9 tapers, and both overall had similar R^2 in the overall fit (90% and 91%, respectively). The same was true for NP (9 and 11 subjects), and WB analysis (18 and 16), which had similar ratios of subjects with time constants. This was surprising, as I expected there to be large differences between the two tapering conditions given the variance-bias tradeoff and the nature of the small TEOAE signal. Furthermore, while there were sometimes large intraindividual differences (e.g., S9 in the narrowband condition) when comparing a subject's time constants between tapered conditions, the majority of the time there was less than a 10 ms difference between their time constants. The second point is that the NP time constants had the poorest overall fit with only 9 – 11 subjects and 68% - 91% R^2 values respectively, while the NB and WB time constants performed much higher. The fitting with a NB analysis had nearly an 80% (14-15 out of 19 subjects) success rate and a very high R^2. Fitting with the WB analysis also had a very high success rate (16 and 18 out of 19) and a very high R^2, at 92-95%. However, while the WB success rate with the fitting was quite high, it is important to mention that the values from the table only included the time constants for the $\Delta TEOAE_m$.

Time constants were also estimated for $\Delta TEOAE_{m+p}$, however, this could not be done for the WB analysis. The reason for this is because while the change in magnitude (typically a decrease in magnitude) produced very reliable responses across frequency, the change in phase across the frequency range is very large and makes little sense to analyze. For example, if the elicitor is broadening the auditory filters in the cochlea, the TEOAE responses may also have a corresponding phase lead relative to the responses without the elicitor (see Mertes and Goodman, 2015). However, when a large frequency region is analyzed (e.g. WB analysis) very large and somewhat random effects occur because the phase leads and phase lags that occur due to the elicitor add together across channels in the overall response. $\Delta TEOAE_{m+p}$ was sometimes as large as 10-15 dB, and sometimes as small as 1 dB depending on the subject and elicitor duration. Therefore, it only makes sense to analyze changes of phase within a channel (NB) or at a single frequency (NP), because the MOCR should be reducing the gain at the signal frequency place, and in this study the comparison frequency behaviorally is at 2 kHz. The discrepancy between the NB and the NP time constants for the ΔTEOAE with magnitude plus phase was similar to the $\Delta TEOAE_m$, where below ~50 percent of the subjects had a reliable time constant, and this was true for either taper number. In contrast, the NB analysis provided the best responses, with over 80% of the subject's time constants estimated for either taper amount. These data is reported in Table 6 (only the NB is reported for clarity). Overall, the NB analysis was a superior estimate of ΔTEOAE compared to the nearest spectral peak analysis, as it had the highest reliability in the responses and time constants across subjects, and it also was a more robust measure of MOCR effects (magnitude alone and magnitude plus phase) compared to the wideband analysis. The time constants from this section were analyzed in more detail in the next section.

Table 6. Individual time constants (τ) and R^2 values for ΔTEOAE magnitude plus phase in the narrowband analysis. Nearest-peak analysis was not included because less than 50% of the subjects had an estimated time constant. Yellow boxes indicate that time constants could not be estimated.

TEOAE mag + phase Narrowband centered (1/3rd oct band @ 2 kHz)				
	(2 tapers)		(9 tapers)	
	1.5 TW		5 TW	
Subj #	NB (1/3rd oct band)	R^2	NB (1/3rd oct band)	R^2
1	33.8	0.91	24.55	0.94
2	67.11	0.98	89.49	0.97
3	49.99	0.98	44.7	0.99
4	140.4	0.98	141.3	0.98
5	20.07	0.89	21.68	0.91
6	183	1.00	181.9	1.00
7	111.6	0.99	114.6	0.99
8	63.7	0.99	62.07	0.99
9	357.9	1.00	356.6	1.00
10	14.25	0.92	22.14	0.96
11	-	-	-	-
12	-	-	-	-
13	-	-	30.58	0.96
14	13.68	0.85	23.69	0.91
15	37.24	0.80	37.18	0.83
16	62.3	0.99	52.16	0.99
17	27.59	0.94	27.64	0.97
18	29.07	0.91	26.72	0.93
19	89.28	0.98	81.53	0.99
Total N	**16**	**.95**	**17**	**.96**

3.3.4 MOCR Effects from TEOAEs: Magnitude and Time Constants

With the optimal TEOAE parameters found (a NB analysis, TW = 5) in the previous section (3.3.3), the change in TEOAE magnitude and time constants for $\Delta TEOAE_m$ and $\Delta TEOAE_{m+p}$ data could be tested. The larger taper number was selected as the optimal parameter as it had a slightly higher success in the total number of time constants estimated across subjects. Results for $\Delta TEOAE_m$ are discussed first.

Each individual subject's $\Delta TEOAE_m$ responses are shown in Fig. 16, and the averaged ΔTEOAE is shown in Fig. 17. The individual functions in Fig. 16, generally build up over a period of a 100-200 msec and then asymptote. These individual time constants are re-reported in Table 7, to highlight these optimized conditions, and to establish the overall time constants and their r-squared values. An oscillating effect, where there are slight increases and decreases as a function of elicitor duration is apparent for some subjects (S1, S5, S10, S12, S16, S18), which is a pattern seen in the behavioral data as well (Fig. 6). In these cases, it may be that the elicitor is fully active at shorter duration elicitors, and it turned its own effectiveness down as the elicitor's duration increased, as described by Roverud and Strickland (2014). The range of time constants was very large ~34 – 661 msec, mostly due one subject (S17; 660.73 msec), and most of the time constants were within 200 ms. This finding is similar to the wide range of time constants reported in previous work (see Backus and Guinan, 2006; Walsh et al., 2010). In terms of magnitude, the majority of the ΔTEOAE shifts were within 1 dB or smaller.

Figure 17 shows the average $\Delta TEOAE_m$ shifts with the fitted exponential function. The average time constant (N =15) for $\Delta TEOAE_m$ magnitude only was 171.19 ± 45.11 msec (R^2 = 91%). Note that the averaged $\Delta TEOAE_m$ shifts are all smaller than 1 dB.

Next, $\Delta TEOAE_{m+p}$ data will be discussed. Fig. 18 shows each individual function with a fitted curve, and the averaged response is shown in Fig. 19 As noted earlier, fits were not possible for data from two subjects (S11 and S12). In contrast to the $\Delta TEOAE_m$ data, these individual $\Delta TEOAE_{m+p}$ functions generally build up over the course of tens of milliseconds and then asymptote. Table 7 indicates that the range of time constants was ~22-182 msec. Note that the shifts for $\Delta TEOAE_{m+p}$ are larger than the shifts for $\Delta TEOAE_m$. A large portion of the effects in this data set are 1-2 dB, and some had larger effects.

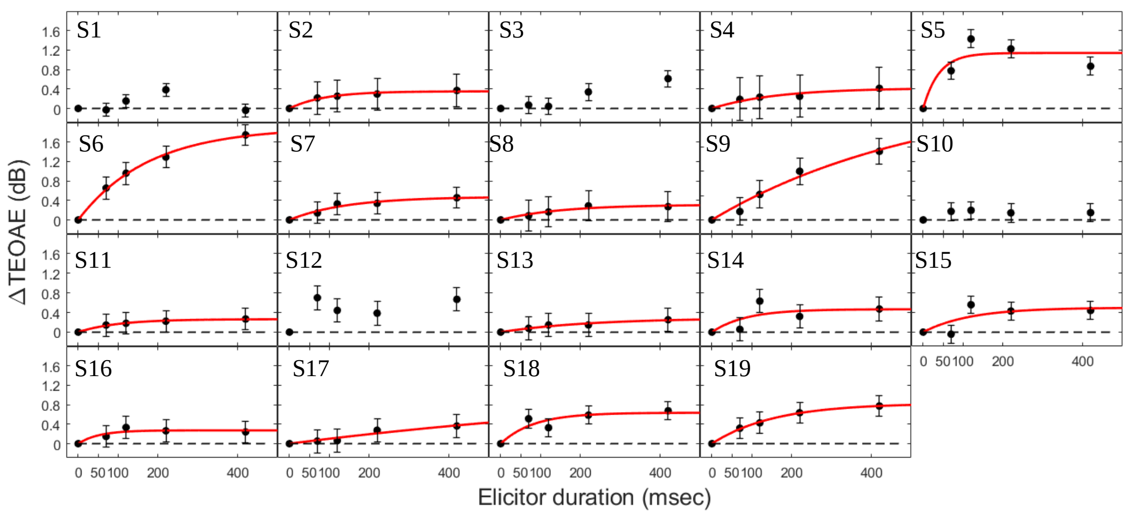

Figure 16. $\Delta TEOAE_m$ as a function of the elicitor duration for each subject. Each cell is data for an individual subject, and each individual data point is indicated by a filled symbol with a corresponding standard deviation indicated by the error bars. The solid curves are the exponential curve fits. The function could not be fit to data for subjects 1, 3, 10, and 12 due to horizontal (and sometimes oscillating) patterns, and thus no time constants were estimated for these subjects.

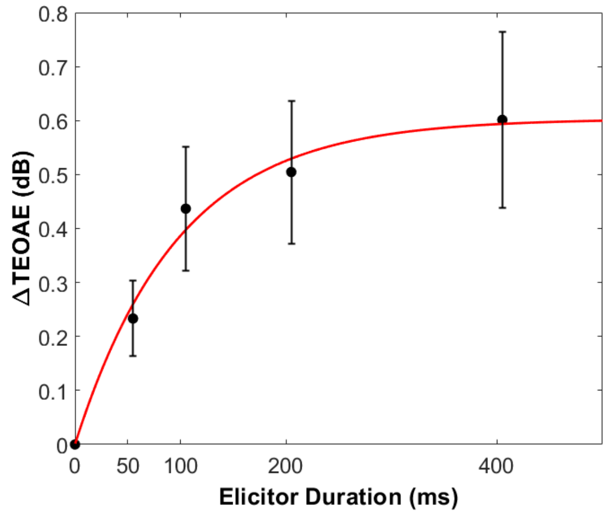

Figure 17. The average change in the TEOAE magnitude as a function of the elicitor duration. The layout and coding are identical to Fig. 16. Individual data that were unable to be fit by the function were not included in the average (N = 15). SEM is indicated by the error bars.

Fig. 19 shows the average $\Delta TEOAE_{m+p}$ shifts with the fitted with exponential function. The average time constant (N =17) for $\Delta TEOAE_{m+p}$ was 78.73 ± 20.75 msec (R^2 = 96%). Note that the averaged ΔTEOAE is considerably larger (0.8-1.5 dB) when phase is accounted for in the response. This large response ultimately shifts upwards, compared to the relatively smaller effects seen in the magnitude only data. Importantly, the average time constant in this condition is similar to the time constants reported in the behavioral section at the 2 kHz signal frequency (~83 - 85 msec), but not similar when the ΔTEOAE included only magnitude (~171 msec). Consistent with the behavioral data, the exponential function had very high R^2 values (Table 7), indicating that it accounted for both the psychoacoustic and physiological data quite well. Chapter 4 will delve directly into comparing these data sets to the behavioral measures of gain reduction that was discussed in chapter 2.

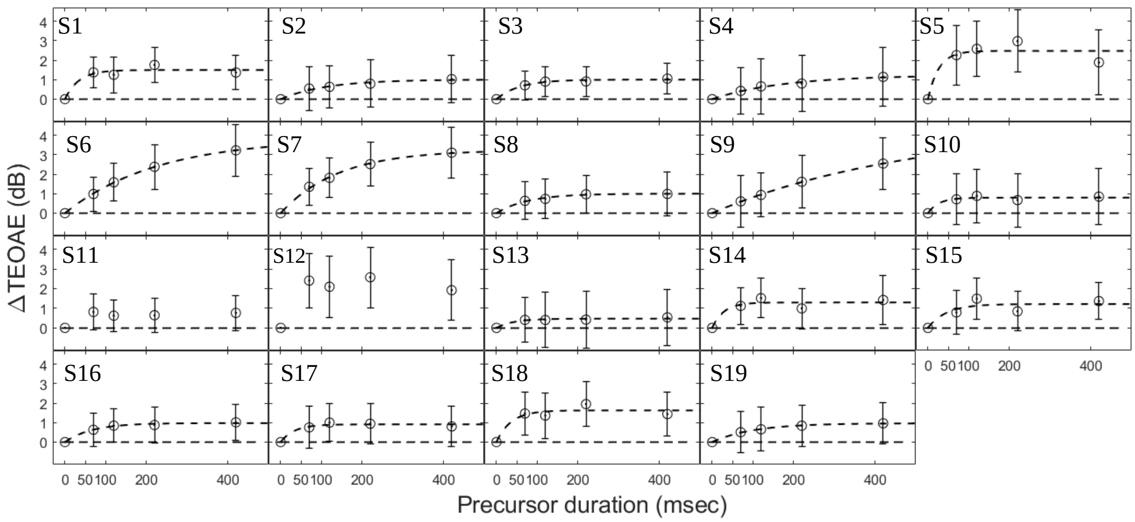

Figure 18. $\Delta TEOAE_{m+p}$ as a function of the elicitor duration for each subject. Each individual data point is indicated by an open circle with a corresponding standard deviation indicated by the error bars. The dashed curves are the exponential curve fits. The function could not be fit to data for subjects 11 and 12.

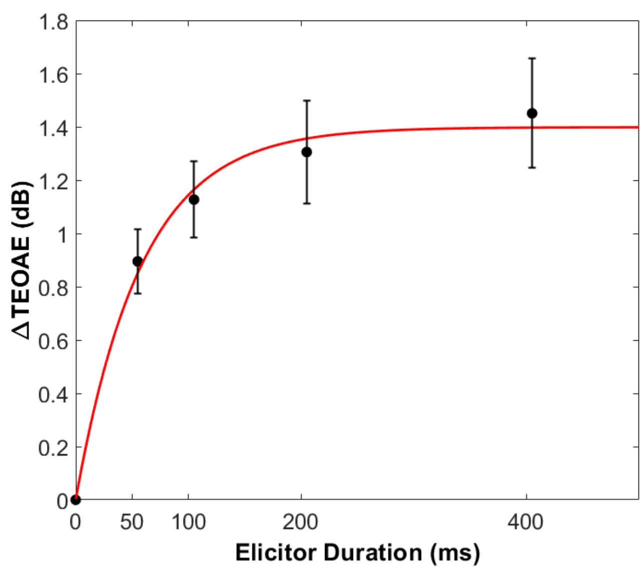

Figure 19. The average change in the TEOAE magnitude plus phase as a function of the elicitor duration. The layout and coding is identical to Fig. 16. Individual data that were unable to be fit by the function were not included in the average (N = 17). SEM is indicated by the error bars.

Table 7. Individual and averaged time constants with corresponding R^2 values when fitting the exponential function for each elicitor duration, when the ΔTEOAE magnitude or magnitude plus phase. Yellow boxes indicate that the response was either a poor fit (R^2 = below 60%), or that a time constant was not able to be estimated due to a horizontal pattern. τ units are in milliseconds, and the ± error of the averaged time constants is the SEM of the individual time constants.

	$\Delta TEOAE_m$ (magnitude only) N = 15		**$\Delta TEOAE_{m+p}$ (magnitude plus phase) N = 17**	
Subj	τ	R^2	τ	R^2
1	***66.29***	***0.34***	***24.55***	***0.94***
2	***68.55***	***0.97***	***89.49***	***0.97***
3	**-**	**-**	***44.7***	***0.99***
4	***129.11***	***0.93***	***141.3***	***0.98***
5	***33.75***	***0.86***	***21.68***	***0.91***
6	***142.15***	***0.99***	***181.9***	***1.00***
7	***114.76***	***0.97***	***114.6***	***0.99***
8	***123.29***	***0.96***	***62.07***	***0.99***
9	***373.02***	***0.99***	***356.6***	***1.00***
10	**-**	**-**	***22.14***	***0.96***
11	***79.82***	***0.99***	**-**	**-**
12	**-**	**-**	**-**	**-**
13	***319.07***	***0.92***	***30.58***	***0.96***
14	***68.44***	***0.65***	***23.69***	***0.91***
15	***96.72***	***0.69***	***37.18***	***0.83***
16	***43.29***	***0.87***	***52.16***	***0.99***
17	***660.73***	***0.95***	***27.64***	***0.97***
18	***61.11***	***0.84***	***26.72***	***0.93***
19	***126.18***	***1.00***	***81.53***	***0.99***
Average	***171.19 ± 45.11***	***0.91***	***78.73 ± 20.75***	***0.96***

3.3.5 A Control Experiment with a Longer Buffer Window

One potential issue that will be addressed is the interstimulus interval between buffers in our TEOAE experiment and its relationship to the decaying effects of the MOCR. As mentioned in the TEOAE Testing section (3.2.6), the buffer window is exactly the same length in duration (approximately 451 msec total; see Fig. 9). This includes 400 milliseconds to account for the longest elicitor duration, 5 milliseconds delay between the offset of the elicitor and the onset of the click, the click itself, and approximately ~43 milliseconds post-click before the onset of the next buffer. Fig 20. depicts how subsequent buffers would occur in time for any elicitor condition. It is known that the MOCR decay has a time constant in the hundreds of milliseconds (~ 100–200 msec; James et al., 2005; Backus & Guinan, 2006). Therefore, there shouldn't be a carry-over effect from a previously presented elicitor onto the following buffer given the time course of the decay of the MOCR, at least for any of the shorter duration elicitor conditions (50, 100, 200 msec; top three panels of Fig. 20), as there would be hundreds of milliseconds for the system to recover to baseline. However, for the longest elicitor duration (400 msec; bottom panel), it is possible that there may have been a carry-over effect of the elicitor onto the following buffer, as there would only be ~43 milliseconds before the onset of the next buffer.

To test for potential carry over effects, two subjects (S2 and S13) completed a control experiment for the 400-ms elicitor condition, either with the original short post-click window (~43 msec), or with a longer post-click window with an additional 200 milliseconds (~243 msec total). The baseline condition was the same as before, which was the click presentation by itself fixed in its temporal position at 405 milliseconds from the onset of either buffer window, but with either a short post click window or long post click window. The response window from which the elicitor effects were recorded stayed the same, as did the duration of the elicitor (400 ms). Therefore, any difference between these measures (ΔTEOAE) should be due to the difference in the duration of the post-click between the two conditions. For these two subjects, ΔTEOAE was estimated with a NB analysis (1/3rd octave band centered at 2 kHz) and a WB analysis (0.001 – 5 kHz). In addition to the short and long post-click window data tested, I also re-reported the original recorded responses with 400 msec elicitors which used a short post-click window to see how reproducible our data was from session to session. The control conditions were carried out approximately 5-6 months from the original recordings reported in the current study. These data are shown in Fig. 21.

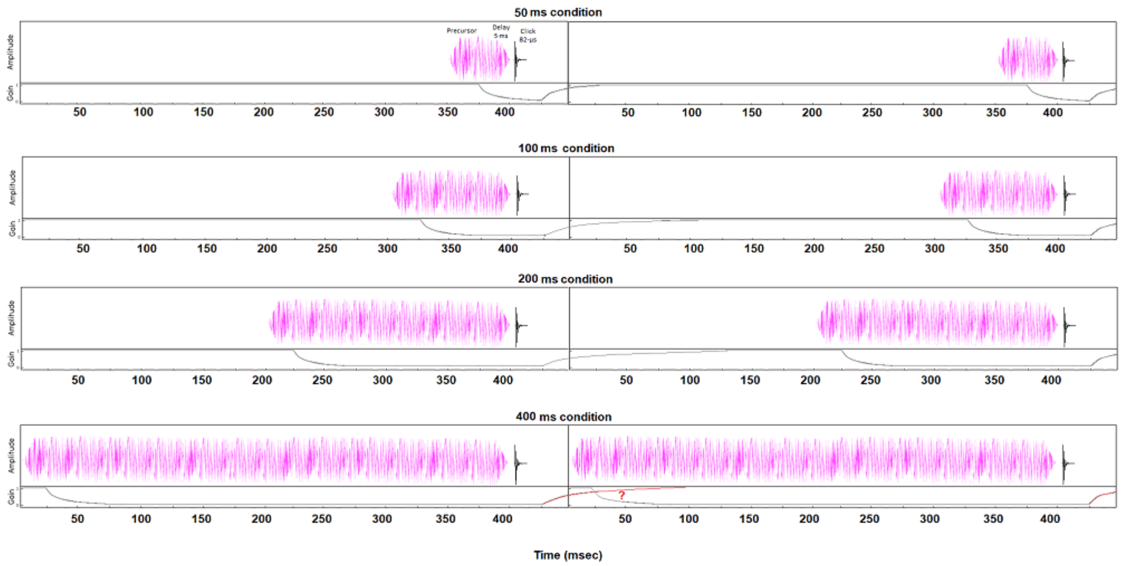

Figure 20. Schematic of multiple buffer presentations for each TEOAE listening condition with an elicitor. For the shortest elicitor conditions [50, 100, 200, 400 msec], the decay of the MOCR can be seen in the bottom portion of each panel, labeled “gain”. Because the time constant of the MOCR decay is roughly 100-200 msec, there should be more than enough time for the MOCR effect to wear off before the onset of the elicitor in a subsequent buffer. However, for the longest condition, it is possible that there is a carry-over effect from the elicitor in the preceding buffer to the following buffer, shown in the red tracing. It is not clear how much of an impact this has on the response at this duration.

It is evident from Fig. 21. that extending the post-click window from 43 to 243 msec had little to no influence on the ΔTEOAE. This is because the shorter conditions (orig-short and re-short) do not show substantially different shifts in the TEAOE compared to the longer condition (long). This is more evident for subject 2, where there is virtually no difference in ΔTEOAE across conditions. There is a bit more variability in subject 13’s data, however, the differences are within ~0.1 dB of one another in the NB analysis, which was also the preferred optimal analysis in the current study. There is also little to no change in dB for the original short window (orig-short) data compared to the newer short window data (re-short), indicating that our data is quite consistent in magnitude even over the course of a multiple months between sessions. Subject 13’s wideband data showed a bit smaller ΔTEOAE (0.3 dB) compared to either of the shorter windowed responses, however, all of the other responses were in the margin of error.

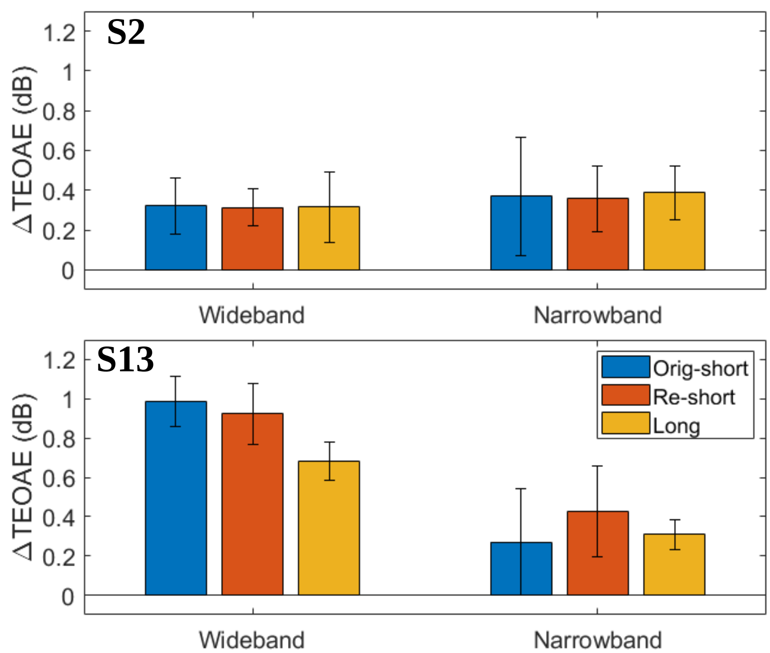

Figure 21. The change in TEOAE (dB) with a 400-ms elicitor, either with a short post-click window or a long post-click window, for two subjects (subject 2 - top panel, subject - 13 bottom panel). ΔTEOAE was measured over a wideband frequency range (0.001 – 5 kHz) or over a narrowband frequency range (1/3rd octave band centered at 2 kHz). Data from the original experiment (orig-short) is presented which used a short post-click window. Then the pilot experimental data was collected either with a short (Re-short) or long (Long) post-click window. Overall, there were very small differences between any of the conditions, indicating that there were likely no significant carry over effects in the longest elicitor duration condition, and that the effects were highly replicable from session to session.

CHAPTER 4: A COMPARISON OF PSYCHOACOUSTIC AND PHYSIOLOGICAL MEASURES OF GAIN REDUCTION

4.1 Introduction

The relationship between the magnitude (strength) of the MOCR from OAE based methods and various psychoacoustic measures of gain reduction or intensity discrimination tasks have been compared in correlational (and observational) studies, with overall mixed results (summarized in section 1.7). Intuitively, a positive predictive relationship between the two measures would be expected due to a shared underlying gain reduction mechanism: gain reduction via the MOCR. While some studies have found that both measures relate well (Micheyl et al., 1997; Kawase et al., 2000; Walsh et al., 2010), multiple studies found that the measures have not related well (Keefe et al., 2009; Wicher and Moore, 2014; Fletcher et al., 2016; Maruffo-Pérez et al., 2021). While there may be many reasons for the overall mixed effects in the literature between psychoacoustic and physiological measures of gain reduction (see Marrufo-Pérez et al., 2021; Jennings, 2021), the comparisons made in the current study will focus on addressing specific concerns outlined in section 1.7: the similarity of methodology and stimuli between gain reduction measures, the laterality of the elicitor, the overall subject size, and subject variability as a factor. The comparisons made will include the psychoacoustic data from Chapter 2 and the physiological data from Chapter 3, where the magnitude and time constants from those sections will be directly compared. Individual subject variability will be accounted for in a linear mixed-effects model.

4.2 Methods and Design

4.2.1 Subjects

All nineteen subjects who completed the psychoacoustic (Chapter 2) and physiological (Chapter 3) gain reduction experiments will have their data directly compared in this chapter.

4.2.2 Psychoacoustic and Physiological Data Selection for Comparison

The psychoacoustic and physiological stimuli and procedure used in this section to make comparisons were outlined in Chapter 2 and Chapter 3, respectively. For the psychoacoustic experiments, cochlear gain reduction was measured with a “masker present” and a “masker absent” condition. Only the 2 kHz psychoacoustic data was used to compare with the TEOAEs, while the 4 kHz psychoacoustic data was not considered for comparison. This is because the SNR of the TEOAE is quite low for most subjects at 4 kHz and higher frequencies compared to lower frequencies. Previous studies have used 2 kHz as a frequency to compare between psychoacoustic and physiological measures of gain reduction (Kawase et al., 2000; Fletcher et al., 2016). For the physiological measures, the MOCR magnitude was measured by the vector difference between the TEOAE with and without the elicitor (ΔTEOAE), either in terms of magnitude alone, $\Delta TEOAE_m$, or magnitude plus phase, $\Delta TEOAE_{m+p}$. With the optimal TEOAE parameters found in Chapter 3, the ΔTEOAE was estimated with a NB analysis (1/3rd octave band, centered at the 2 kHz frequency) with 9 tapers. Data from elicitor durations of [50, 100, 200, and 400 ms] were compared directly. While a larger number of precursor durations were tested in Chapter 2, data were selected for the precursor durations that corresponded to the elicitor durations used in the TEOAE experiment.

4.2.3 Two Comparisons of Psychoacoustic and Physiological Gain Reduction

Two separate comparisons were made to test whether physiological measures of gain reduction (TEOAEs) can predict the changes in psychoacoustic measures of gain reduction, as a function of elicitor duration. The first comparison made is the $\Delta TEOAE_m$ to the masker present gain reduction condition. The second comparison made is the $\Delta TEOAE_{m+p}$ to the masker absent condition. Figure 22 shows a schematic of these comparisons.

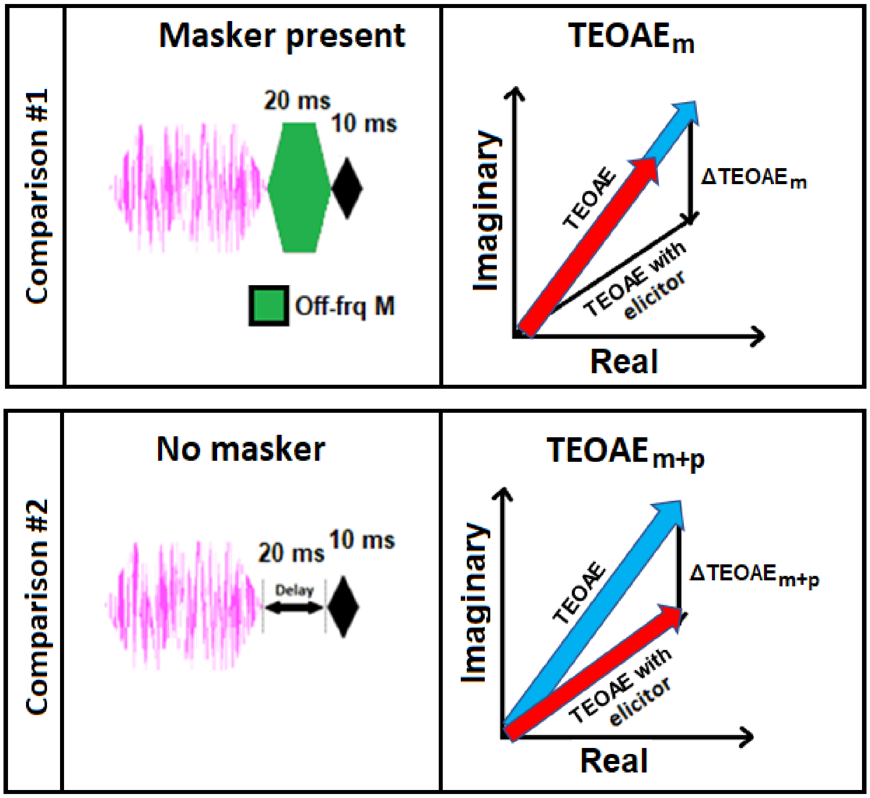

Figure 22. Schematic of the two comparisons made. Comparison #1: Masker present and TEOAE$_m$ (top), and Comparison #2: No masker and TEOAE$_{m+p}$ (bottom). These comparisons were made in terms of magnitude and time constants.

For the first comparison, the ΔTEOAE$_m$ metric is most comparable to what most other studies generally have tried to measure when comparing psychoacoustic and physiological data, which is to measure MOCR strength as a change in the magnitude of the OAE by the elicitor. The masker present condition in this study is consistent with gain reduction and not excitatory masking, and is mostly similar to the PTC method whereby the off-frequency masked condition is basically the tail response of a PTC, and the on-frequency masked condition is like the response at the tip of the PTC. In other words, our psychoacoustic masker present method is like a sparse tuning curve. The difference in masking thresholds from PTCs with and without the presence of an elicitor have been used in previous studies comparing psychoacoustic gain reduction to physiological OAE data (Kawase et al., 2000; Wicher and Moore, 2014).

For the second comparison the ΔTEOAE$_{m+p}$ metric provided larger responses compared to the ΔTEOAE$_m$ metric, and it is not known if including phase in the response would alter the overall relationship between the psychoacoustic and physiological data. The no-masker condition was used because the responses were large (statistically larger than the off-frequency masked conditions), and the stimuli in this condition are more similar to the TEOAE stimuli. Previous studies have shown the no masker estimate of gain reduction to be comparable in magnitude to the off-frequency masked conditions (DeRoy Milvae and Strickland, 2018; Salloom and Strickland, 2021). Therefore, by using the psychoacoustic and physiological conditions that provided the largest responses by the elicitor, I hypothesized that the relationship between the measures would increase (positively) as a function of elicitor duration due to strong activation of the MOCR.

4.2.4 Linear Mixed-Effects Model (LMM)

A linear mixed-effects model (LMM) was used to test the hypothesis that physiological measures of gain reduction (TEOAEs) can predict the changes in psychoacoustic measures of gain reduction, as a function of elicitor duration within the same subjects. LMMs are very useful method for analyzing data that are non-independent, multilevel/hierarchical, longitudinal, or correlated. In a traditional correlational analysis, the relationship between a predictor (an independent variable) and an outcome (dependent variable) can be determined by estimating the goodness of fit (R^2) of the estimated regression line to the observed data. However, a correlational analysis model cannot account for random factors, such as individual subject variability in the data. This is important because MOCR effects seem to vary from subject to subject measured behaviorally (e.g., Jennings et al., 2009) and physiologically (e.g., Backus and Guinan, 2007; Marshall et al., 2014; Mertes and Goodman, 2016). In contrast, LMMs allow both fixed effects (non-varying parameter) and random effects (truly random parameter) to be studied, which is typical in non-independent data sets (e.g., multiple levels of the independent variable) and nested groups. Assumption of normality tests were conducted prior to analysis using the Shapiro-Wilk test of normal distribution and the corresponding normal Q-Q plots. With this test, any subset of data tested for this assumption with a p-value equal to or greater than 0.05 would meet the criterion to assume normal distribution.

4.3 Results and discussion

4.3.1 A comparison of gain reduction: magnitude effects

Analysis was conducted to test whether changes in the physiological responses (ΔTEOAE) by the elicitor could predict the changes in the psychoacoustic responses, as a function of elicitor duration. The first comparison of physiology and psychoacoustic measures of gain reduction made is between the $\Delta TEOAE_m$ and the masker present data. Figure 23 shows the data for this comparison. Here, a single subject's physiological and psychoacoustic responses for an elicitor duration are plotted as a single coordinate, thus each subject will have four points total in the figure. Each coordinate is color-coded to the duration of the elicitor used. Linear regressions were fitted to each elicitor duration, and their slopes and R^2 values are shown in Table 8. These lines are color coded to the corresponding elicitor duration. As noted in Chapter 2, subjects 10 and 16 were removed from the analysis as they were statistical outliers, but their data are indicated in Figure 23 as cross (S10) and asterisk (S16) symbols, respectively. Therefore, N = 17, for this analysis. From observation of the data and corresponding linear fits, the physiology does not positively predict the perceptual data for any of the elicitor durations. In fact, some of the trends are negatively associated with each other. Importantly, there is no simple linear association between the two variables, as seen with the very poor R^2 values.

As mentioned in sections 1.7 and 4.2.4, individual subject variability cannot be determined in a simple correlational analysis. Furthermore, it is not clear from the data in Fig. 23 if there was a significant amount of individual variability because the data are organized in clusters, where the group mean (as a linear fit) is analyzed, and individual trends are unknown. So, from this cluster data, I plotted individual subject data as a function of elicitor duration in Figure 24, to see if there were positive relationships that were being hidden from the correlational analysis. Uncertainty bars are the SD for the psychoacoustic data (vertical bars) and physiological data (horizontal bars). The bottom right cell is the averaged responses, and the uncertainty bars are the SEM for the psychoacoustic (vertical bars) and physiological data (horizontal bars). The data in Figure 24 are color coded identically to Figure 23.

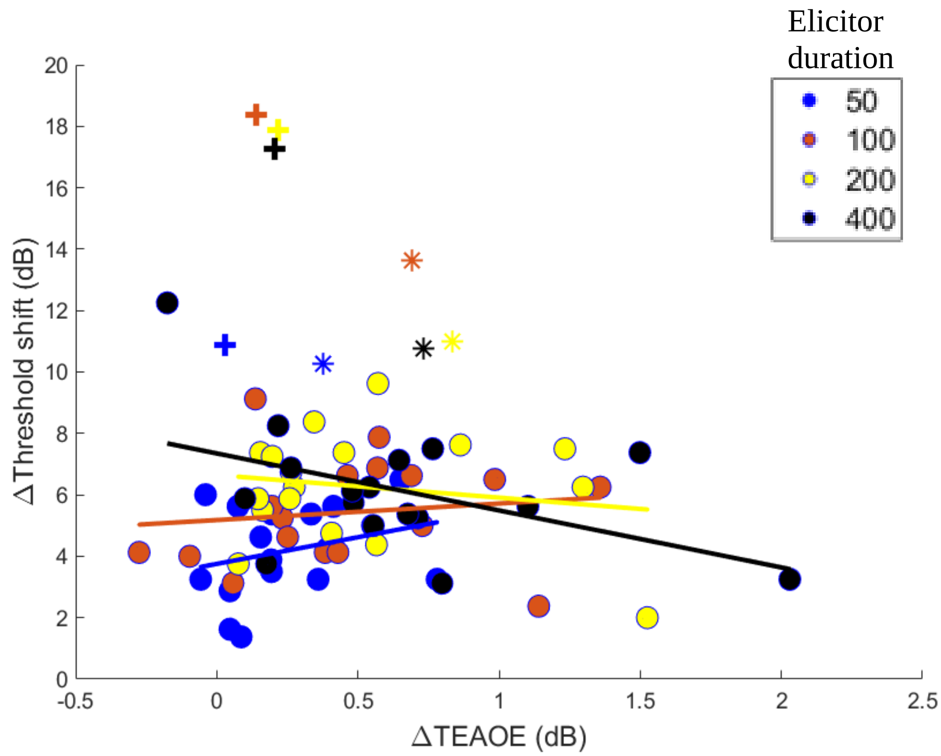

Figure 23. Group data (N = 17) for $\Delta TEOAE_m$ and masker present conditions (Comparison #1). Each point (circles) represents a single subject's corresponding physiological and psychoacoustic response for a given elicitor duration, which is color coded. Two subjects' data were considered outliers, shown by cross (S10) and asterisk (S16) symbols, respectively. While regression lines were fitted to each cluster, the R^2 values were all below 20%.

Individual trends are quite variable. Some subjects do show a positive relationship between the two responses (S2, S9, S17) as do the overall averaged responses (bottom right cell), however, others show a flat (e.g., S5 and S6) or negative relationship (e.g., S19) with duration. In some cases, there is no trend, and the responses do not increase much with duration (e.g., S11) or the pattern is somewhat random with duration (e.g., S14). Another important point is that the magnitude of the responses vary quite a bit from subject to subject, some on the smaller side (e.g., S2 and S13), whereas other subjects are on the larger side (e.g., S9). An additional analysis was included to account for this individual subject variability, by using a linear mixed-effects model (LMM).

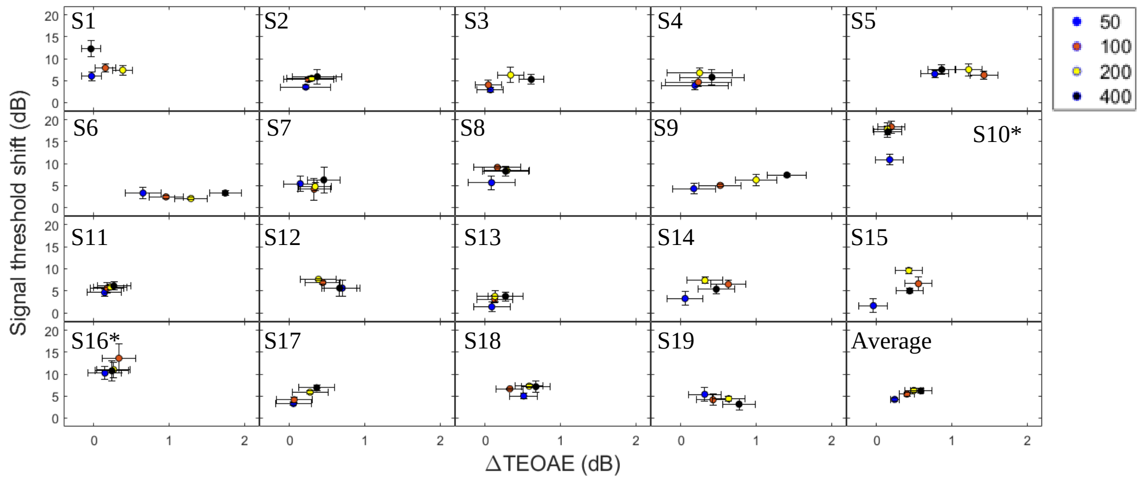

Figure 24. Individual subject trends with elicitor duration for comparison #1 ($\Delta TEOAE_m$ and masker present conditions). Elicitor duration is coded as in Fig. 23. Asterisks on the subject number indicate that their data were statistical outliers and not included in the average responses or statistical analysis. Uncertainty bars from the behavioral data are placed vertically and physiological data are placed horizontally, and correspond to the SD for individual data and SEM for the average data.

Table 8. Reported slopes and R^2 values for the first comparison of physiology to psychoacoustic measures of gain reduction, for each elicitor duration.

Elicitor duration (ms)	Slope	R^2
50	2.52	0.19
100	-0.1128	0.0005
200	-1.236	0.0553
400	-1.992	0.1641

A LMM approach was conducted to test whether the physiological changes could predict the psychoacoustic changes, as a function of elicitor duration. The dependent variable was the signal threshold shift (dB) from the psychoacoustic data (i.e., gain reduction). Fixed effects included $\Delta TEOAE_m$ (a continuous variable) and its interaction with elicitor duration (a categorical variable: 4 levels). A significant interaction between $\Delta TEOAE_m$ and elicitor duration would indicate a difference in slope of the clusters with elicitor duration. This is what I had originally hypothesized, that the relationship between the two measures would become increasingly positive with increasing elicitor duration. Individual subject responses were accounted for by making them a random effects variable in the model. All model coefficients were estimated using the restricted maximum likelihood procedure in the lme4 (version 1.4.1) library in *R* 3.6.3 (Bates et al., 2007)

(R Core Team, 2020). Homoscedasticity and normality assumptions tests were conducted on individual subject data, and both tests successfully passed (both could be assumed) based their residual versus fitted plots. Statistical inferences about the experimental fixed effects were made with the F approximation for the scaled type-II Wald statistic (Kenward and Roger, 1997). Approximation with this statistic is a conservative approach in estimating false-alarm rates (type 1 errors) than the Chi-squared approximation of the log-likelihood ratios and has shown to be effective with smaller and complex datasets (Schaalje et al., 2002).

The model estimations are tabulated in Table 9. Both fixed effects in the model, $\Delta TEOAE_m$ and the interaction of $\Delta TEOAE_m$ and elicitor duration, failed to meet statistical significance ($p > 0.05$). That is, the physiological data ($\Delta TEOAE_m$) did not predict the psychoacoustic (masker present) data, as a function of elicitor duration. This is opposite of what I hypothesized before the experiment, where I hypothesized that the $\Delta TEOAE_m$ would have a positive relationship with signal threshold shifts from the psychoacoustic experiment. This means, generally, there is no overall significant predictive relationship between the two measures in this comparison. Furthermore, the non-significant interaction between $\Delta TEOAE_m$ and elicitor duration means that there was no change in this relationship with increased elicitor duration (i.e., no significant change in the responses with elicitor duration). One last point is that the random effect variable, individual subject variance (by-subject intercepts), was actually lower than the residual error in the model (random error), indicating that the overall results were not dependent on the variability across subjects. This was also surprising to me, given the relatively large array of patterns from Figure 24.

Table 9. Summary of the linear mixed-effects model (LMM) on the effects of $\Delta TEOAE_m$ and the interaction of $\Delta TEOAE_m$ and elicitor duration (independent variables) on signal threshold shift (independent variable). Type II Wald approximation was used to generate a corresponding F-statistic and p-value for the fixed effects. Neither fixed effect was statistically significant, $p > 0.05$. Individual subject variability (random effect) was relatively similar to the residual error (uncontrolled variance) in the model, indicating that the random effect played little to no role in the inferential statistics.

	Signal threshold shift			Type II Wald test	
Fixed effects	Estimate	Std. Error	t-value	F-statistic	P value
Intercept	5.1654	0.5143	10.043		
$\Delta TEOAE_m$	0.6513	0.8733	0.746	$F(1, 61.932) = 2.5052$	$p = 0.1186$
$\Delta TEOAE_m$*Elicitor duration	0.9838	0.9029	1.090	$F(3, 50.976) = 0.6101$	$p = 0.6115$
Random effects	Variance		SD		
By-subject intercepts	1.902		1.379		
Residual	2.233		1.494		

The next comparison included the $\Delta TEOAE_{m+p}$ and no-masker conditions, and is shown in Fig 25. The coding is identical to Figure 23. Here, N = 19, as the residuals of the data met the assumption of normality and homoscedasticity. Linear regressions were fitted to each elicitor duration, and their slopes and R^2 values are shown in Table 10. In contrast to the first comparison, the clusters appear to stratify noticeably more with duration in Fig. 25, and the overall responses are larger, particularly for the $\Delta TEOAE_{m+p}$ data. As noted in Chapter 3, the $\Delta TEOAE_{m+p}$ is a larger (and possibly more sensitive) measure of MOCR activation, and seems to shift the clusters to the right in the correlational plot. Despite this, the trends are negatively associated with each other as a function of elicitor duration, and the R^2 values are still low (< 20%) (see Table 10).

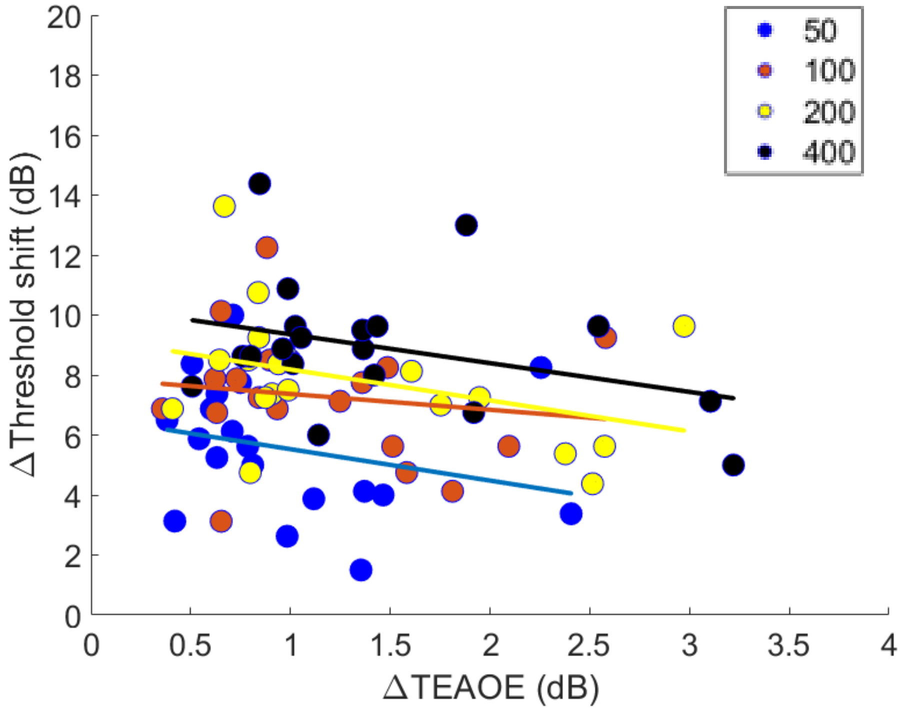

Figure 25. Group data (N = 19) for $\Delta TEOAE_{m+p}$ and no masker conditions (Comparison #2). This figure is coded identically as Fig. 23. The overall shifts in $\Delta TEOAE_{m+p}$ are substantially larger than $\Delta TEOAE_{m}$. The R^2 values were all below 20%, and negatively associated between the independent and dependent variables.

Table 10. Reported slopes and R^2 values for the first comparison of the physiology ($\Delta TEOAE_{m+p}$) to psychoacoustic measures of gain reduction (no-masker), for each elicitor duration. Similar to comparison #1, the R2 values are low, and there is a negative relationship between the physiological responses and psychoacoustic responses.

Elicitor duration (ms)	Slope	R^2
50	-1.048	0.0726
100	-0.5361	0.021
200	-1.034	0.1388
400	-0.9667	0.1145

Individual subject trends are shown in Figure 26, coded as in Figure 24. Observationally, one thing that stands out is that multiple subjects show a positive trend with elicitor duration (S2, S3, S4, S6, S7, S8, S9, and S16), as did the overall average. However, others showed somewhat random effects with duration (e.g., S12), or little to no effect with duration (e.g., S13 and S17). Another noticeable thing from the individual data is that the SD of the $\Delta TEOAE_{m+p}$ is higher than the $\Delta TEOAE_m$. SD was calculated as the SD of 1024 averaged responses for any listening condition, and the average of these SDs for the 4096 responses was taken as the SD of the TEOAE listening condition. There are other ways to calculate the SD from these responses and this can be done in the future.

Next, following the exact same procedure as the first comparison, a LMM was used to test if the $\Delta TEOAE_{m+p}$ responses could predict the change in signal threshold in the no-masker responses with elicitor duration. In this test, the Fixed effects included $\Delta TEOAE_{m+p}$ (a continuous variable) and its interaction with elicitor duration (a categorical variable: 4 levels), and the individual subject responses were set as a random effect in the model. These parameters are virtually identical to those in the first comparison. The model estimations are tabulated in Table 11. Both fixed effects in the model, $\Delta TEOAE_{m+p}$ and the interaction of $\Delta TEOAE_{m+p}$ and elicitor duration, were highly statistically significant ($p < 0.0001$). That is, the physiological data ($\Delta TEOAE_{m+p}$) was successful at predicting the psychoacoustic (no-masker) data, as a function of elicitor duration. However, the relationship between the measures was negative, an effect that was true for all elicitor durations. My original hypothesis was that there would be a positive relationship between the measures. Therefore, with this comparison, there was an overall significant predictive relationship between the two measures, but a negative relationship was found. Furthermore, there was a significant interaction between $\Delta TEOAE_{m+p}$ and elicitor duration, indicating that the overall responses increased with elicitor duration. This was apparent in both the group and individual data, and was probably due to the very large $\Delta TEOAE_{m+p}$ responses when phase was included in the response compared to the $\Delta TEOAE_m$. That is, for both physiological and behavioral measures, the responses tended to increase with elicitor duration. Importantly, the random effects variable, individual subject variance, was considerably larger than the residual error in the model (over 4 to 1 ratio), indicating that the overall results were highly dependent on the variability across subjects. As stated before, individual subject variability is not accounted for in traditional correlational analysis, and it had a large impact on the results of this comparison.

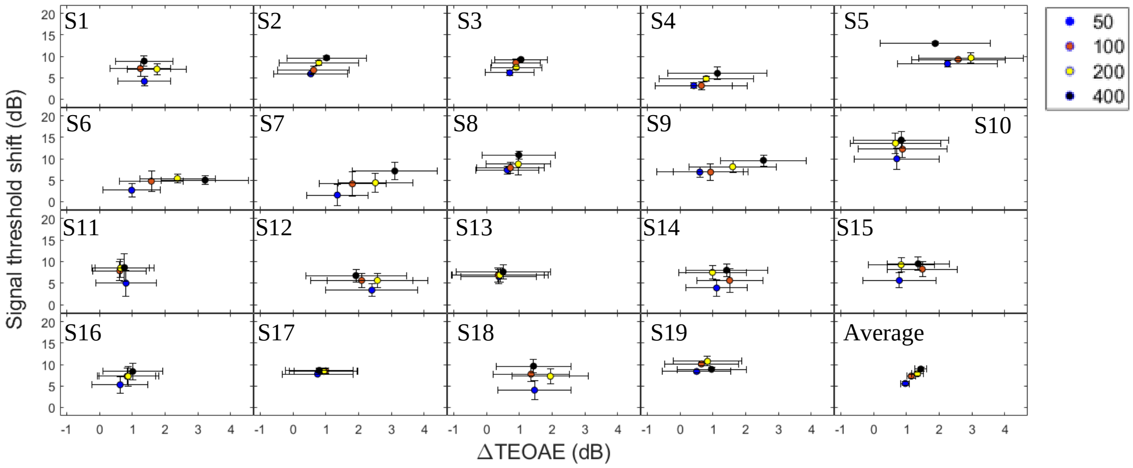

Figure 26. Individual subject trends with elicitor duration for comparison #2 ($\Delta TEOAE_{m+p}$ and no-masker conditions). Elicitor duration are coded as in Fig. 24.

Table 11. Summary of the linear mixed-effects model (LMM) on the effects of $\Delta TEOAE_{m+p}$, and the interaction of $\Delta TEOAE_{m+p}$ and elicitor duration on signal threshold shift (N = 19). Type II Wald approximation was used to generate a corresponding F-statistic and p-value for the fixed effects. Both fixed effect were highly statistically significant. Individual subject variability very large compared to the residual error, indicating that it played a large in the inferential statistics.

	Signal threshold shift			Type II Wald test	
Fixed effects	Estimate	Std. Error	t-value	F-statistic	P value
Intercept	8.0888	0.6212	13.020		
$\Delta TEOAE_{m+p}$	-0.7003	0.3310	-2.115	F(1, 64.159) = 19.687	P < 0.0001
ΔTEOAE*Elicitor duration	1.9359	0.1856	10.429	F(3, 54.725) = 37.895	P < 0.0001
Random effects	Variance		SD		
By-subject intercepts	4.1446		2.0358		
Residual	0.7978		0.8932		

4.3.2 A comparison of gain reduction: time constants

The next analysis that was conducted between the psychoacoustic and physiological measures of gain reduction was to compare the time constants directly, which were estimated from Chapter 2 (psychoacoustic) and Chapter 3 (physiology), respectively. The hypothesis tested was that the time constants from the physiological measure can predict the time constants from the psychoacoustic measures. A strong positive correlation would indicate that the measures share an underlying mechanism, gain reduction via the MOCR. The first comparison was the $\Delta TEOAE_m$ and the masker present conditions, and the second comparison was $\Delta TEOAE_{m+p}$ and the no-masker conditions. To be included in the analysis, the time constants needed to be well fit (R^2 above 60%), as indicated in Chapters 2 and 3.

Figure 27 shows the results of these data. Data for the first comparison are on the left (N = 12), and data for the second comparison is on the right (N = 17). Schematics of these conditions are shown directly beneath each correlational plot for clarity. It should be noted that there are differences in the number of subjects in each comparison. Data was not included unless a time constant could be fit for both the psychoacoustic and physiological data. It should also be noted that the x and y axes are different in comparison 1 for visual clarity, while the x and y axes are the same for comparison 2. Dashed linear references are shown in both comparisons to indicate a hypothetical 1:1 ratio in ms between the two measures. The solid line is the linear regression fit to the data.

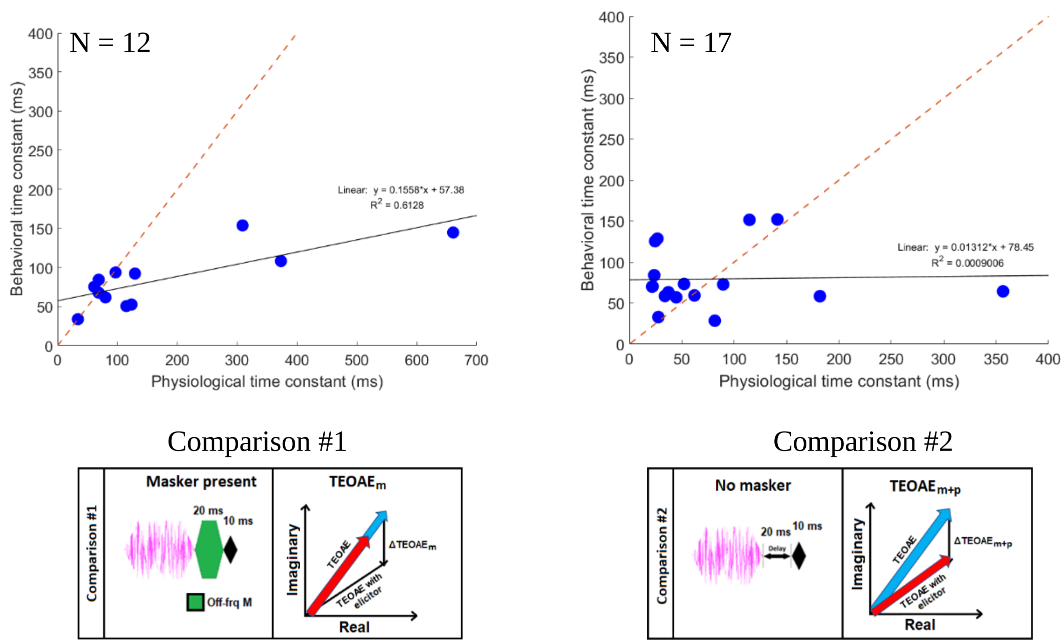

Figure 27. Individual time constants, measured physiologically and psychoacoustically, for comparison #1 (N = 12; left side) and comparison #2 (N = 17; right side). Dashed line in each plot indicates a reference 1 ms/ 1 ms slope.

There was a strong predictive relationship found for the first comparison, with over 60% of the variance explained by the regression line. It is somewhat surprising that such a strong relationship was found when the time constants in these conditions were compared, but this relationship was non-significant when their magnitudes were compared with the linear mixed model. A few points to note, there were a smaller amount of subjects compared in this analysis (N = 12) compared to the comparison of magnitudes (N = 17). Another point to note is that most of the time constants lie on the bottom left portion of the plot, indicating that the time constants were fairly short overall for either measure, but a few points lie much further outside this cluster (see the subject with a ~660 ms TEOAE time constant). A "leave one out" method was used to take data that strayed away from the main cluster, and to determine how much of an impact some of these data had on the overall correlation. Interestingly, taking out the data with the 660 ms TEOAE time constant only reduced the overall correlation to R^2 = ~58%. I followed up by doing a "leave two out" method, and removed the data furthest away from the cluster, just as before. Again, the overall correlation was still relatively large, close to ~60%. When all three larger values, with TEOAE time constants of 300 ms or larger, were removed, the linear trend flattened and R^2 went to 9% of the variance accounted for. It seems that while there were not one or two influential points, there were longer time constants, particularly the physiologically measured time constants that were driving the relationship. It should also be noted that 1/4 of the subjects in this analysis had these longer time constants in the $\Delta TEOAE_m$, which implies that they were not artifacts. Variability in time constants across human subjects been found in multiple studies (e.g., Backus and Guinan, 2006; Walsh et al., 2010).

Comparison 2 is interesting in that there was virtually no predictive relationship between the two measures when comparing their time constants. This is the opposite of the findings from the previous section where a significant relationship was found when comparing the magnitudes in these conditions. Most of the time constants in this comparison were on the shorter side (within 120 ms) for both measures, and more subjects were included in the overall analysis. To determine if there were influential points, I did a 'leave one out' and 'leave two out', and the R^2 values never increased substantially and were always close to zero. It is interesting to point out that the overall results for both comparisons flip-flopped depending on if the comparison was made with the corresponding magnitude or time constant data.

4.3.3. Qualitative Comparison Between Psychoacoustic and Physiological Time Constants

In Chapter 2, the time constants from psychoacoustic measures of gain reduction were found, where masker present conditions produced an average time constant of~85 ms, and the no masker condition produced an average time constant of ~83 ms. These effects were consistent with gain reduction, and not excitatory masking. In Chapter 3, the time constants from the $\Delta TEOAE_m$ produced average time constants of ~171 ms, and the $\Delta TEOAE_{m+p}$ produced average time constants of ~79 ms. That is, when phase was included in the TEOAE data, the average time constant were approximately equal to the psychoacoustic experiments (both ~80 ms). There was quite a bit of variability in the individual time constants. For example, time constants in either of the psychoacoustic experiments ranged from ~27 – 152-ms, while the time constants from the $\Delta TEOAE_m$ conditions ranged ~34 – 661-ms, and the $\Delta TEOAE_{m+p}$ ranged from ~22 - 182-ms. This is consistent with results from other studies showing that MOCR effects are variable across subjects (e.g., MOCR magnitude: Lilaonitkul and Guinan, 2012; MOCR time constants: Backus and Guinan, 2006; Walsh et al., 2010). This was one of the key reasons to use a LMM in the magnitude analysis, in section 4.3.1, and was a significant factor in one of the two comparisons made.

Time constants for the psychoacoustic data ranged from 29 to 154 ms, with an average of ~83-85 ms. In Roverud and Strickland (2014), which measured gain reduction by on- and off-frequency tonal precursors as elicitors, the maximal effect for the on-frequency tonal elicitor on signal threshold occurred for the 50-ms precursor, and then plateaued or oscillated with increased precursor duration. In contrast, their off-frequency precursor continued to increase in effectiveness with precursor duration up to 150 ms. In that paper, time constants were fit to the data as part of a model that also included a temporal integration window and a delay in onset of gain reduction. The time constants ranged from approximately 28 to 76 ms. The model was able to predict the oscillation in gain reduction for on-frequency maskers for durations longer than 50 ms. There are several factors that make it difficult to directly compare results with the current study to the Roverud and Strickland study. Precursor durations in Roverud and Strickland (2014) ranged from 10 to 150 ms, while in the present study precursor durations ranged from 50 to 800 ms. For example, the range of precursors used between studies were different and may be a factor in the time constant estimation, and off-frequency data in the Roverud and Strickland study did not always reach a plateau. An important distinction between the two studies is that the current study estimated time constants (63% of Y_{max}), while the Roverud and Strickland study reported the maximal effect of the on-frequency tonal precursor (the 50-ms precursor). To find the maximal effect with broadband elicitors from our estimates, we can extrapolate to find the duration needed to reach Y_{max}, which is ~132-ms. While this varied across subjects, it appears that the duration needed to produce maximum gain reduction in the present study is between the time needed for on- and off-frequency precursors in the Roverud and Strickland study. This is consistent with a recent study (DeRoy Milvae and Strickland, 2021), which found that broadband precursors had only slightly larger effects on signal threshold compared to on-frequency precursors as a function of precursor level when the precursor duration was fixed in duration (50-ms) and the broadband precursor level was calculated as the energy that should pass through a filter centered at the signal frequency. Since filter bandwidth and suppression may be changing with gain reduction, it might be expected that the effect of duration would be longer for a noise than for a tone elicitor. Overall, the data in the current study are consistent with gain reduction and show that the maximal effects of gain reduction occur within 200 ms of precursor onset for ipsilateral broadband stimulation. This is relatively short compared to the continuous or long duration elicitors that have typically been used to activate the MOCR in the psychoacoustic and physiological literature. Future studies using broadband elicitors may fully activate the MOCR with 150 - 200-ms durations, which could potentially save substantial data acquisition time.

The averaged psychoacoustic time constants were very similar to the averaged physiological time constants when phase was included in the response ($\Delta TEOAE_{m+p}$; ~79 ms), but not when phase was omitted ($\Delta TEOAE_{m}$; ~171-ms). As explained in Chapter 3, the $\Delta TEOAE_{m+p}$ produced substantially larger shifts compared to $\Delta TEOAE_{m}$. A roughly equivalent time constant between the two highly-controlled experiments (behavior and $\Delta TEOAE_{m+p}$) is consistent with a common underlying mechanism in gain reduction via the MOCR. While comparing their time constants or magnitudes directly, such as I did in sections 4.3.1 and 4.3.2, resulted in mixed results, there is still evidence from both measures that the effects are consistent with MOCR gain reduction, and not the by-products of excitatory masking or MEMR activation.

While the majority of studies have focused on measuring MOCR effects from OAEs by estimating the change in OAE magnitude only (see Section 1.7), the changes in phase have been less studied, and may provide additional details on MOCR effects in the cochlea (Lilaonitkul and Guinan, 2012). For example, elicitor induced MOCR activation can produce a reduction in the

OAE amplitude and/or a phase lead in the OAE, which may be consistent with broadened tuning of the auditory filter (Francis and Guinan, 2010). Changes in OAE phase from MOCR activation are still a topic of debate (Francis and Guinan, 2010, Lilaonitkul and Guinan, 2012), as are the relevance for auditory perception. Further analysis on OAE phase changes with elicitor duration are logical future directions of the study.

CHAPTER 5. GENERAL DISCUSSION AND CONCLUSIONS

5.1 General Summary and Results

The current study investigated the temporal properties of psychoacoustic and physiologically measured gain reduction with broadband elicitors. The goal of the study was to investigate three research questions: 1) How does broadband elicitor duration affect auditory perception, 2) Are time constants measured psychoacoustically and physiologically similar to one another, indicating a common mechanism, 3) Can the changes in physiological responses predict the changes in behavioral responses in the same subjects, as a function of elicitor duration?

Question 1 was explored in Chapter 2, where two forward masking gain reduction paradigms were used, either with or without a masker present, and using a large range of broadband elicitors. This experiment was an extension of previous work by Roverud and Strickland (2010), and Roverud and Strickland (2014), where they used a shorter range of tonal elicitors, either at (on-frequency) or below (off-frequency) the signal frequency. The results from Chapter 2 indicated that time constants were fairly short, for 2 kHz (~83-85 ms) and 4 kHz (62 – 104 ms). The pattern of growth and corresponding time constants needed to produce maximal gain reduction in the current study is somewhere in between the time needed for on- and off-frequency elicitors in the Roverud and Strickland study. Overall, near-maximal MOCR effects can be measured with ipsilateral broadband elicitor durations of 200 ms or less in a forward masking paradigm.

For question 2, physiologically measured time constants were measured (Chapter 3) and discussed in section 4.3.3. A great deal of this analysis was to find the "optimal parameters" to study MOCR effects from, as the reliability of MOCR effects from OAEs have been an area of study and debate (e.g., Marshall et al., 2014; Mertes and Goodman, 2016; Marrufo-Pérez et al., 2021). Detailed in Chapter 3, optimal parameters included a multitaper method to calculate the power of the TEOAE signal, a comparison of frequency bandwidth to analyse MOCR effects from, and analyzing TEOAE magnitude and phase. The results showed that a narrowband analysis (1/3rd octave band) with a higher number of spectral tapers (9) provided the most reliable responses time constant fits. Physiological time constants were measured using two TEOAE metrics, magnitude only ($\Delta TEOAE_{m}$) or magnitude with phase ($\Delta TEOAE_{m+p}$). When phase was accounted for in the TEOAEs, the time constants were approximately equal to the psychoacoustic time constants and were relatively short (~80 ms). When only changes in TEOAE magnitude were measured, and phase was omitted, the average time constant was longer (~172-ms). Overall, the psychoacoustic and physiological data in the study were consistent with the timecourse of gain reduction by the MOCR, and consistent with the fast time constant measured in previous studies (Bassim et al., 2003; Backus and Guinan, 2006).

Question 3 was explored in Chapter 4, where the magnitude and time constants of the psychoacoustic and physiological measures of gain reduction were directly compared to one another. To improve the design from previous studies comparing behavior to physiology (outlined in section 1.7), I specifically addressed four different areas of concern, including: the similarity of methodology and stimuli between gain reduction measures, the laterality of the elicitor, the overall subject cohort size, and subject variability as a factors. By using a common forward masking configuration between psychoacoustic and physiological measures, I sought to streamline the study design, and reduce the possibility of inter-methodological variability. Forward masking also eliminates the possibility of two-tone suppression in the effect. I used an ipsilateral elicitor in both experiments, which has been shown to be stronger than a contralateral elicitor of the MOCR in humans, at least measured behaviorally (Salloom and Strickland, 2021). I also recruited and tested a larger subject size (N = 19) relative than many of the previous comparative studies, which is a very large number of subjects for the psychoacoustic experiments. Lastly, I took individual subject variability into account by using a linear mixed effects model (LMM).

When comparing magnitudes (see section 4.3.1), neither comparison had a positive predictive relationship as a function of elicitor duration. In fact, both comparisons showed a mostly negative correlation between the two measures of gain reduction, one non-significant and one significant (comparison 2). This was not expected given that I tested for optimal TEOAE parameters in Chapter 3, and that I addressed the four concerns (see last paragraph or section 1.7) in the research design to improve the probability of a significant correlation if one existed (i.e., a common generation mechanism). Despite an optimized study design, with many areas controlled for, the results are similar to other studies that found a non-relationship or negative association when comparing behavior to physiology (Keefe et al., 2009; Wicher and Moore, 2014; Fletcher et al., 2016; Maruffo-Pérez et al., 2021). Physiological and psychoacoustic time constants were also directly compared in Chapter 4, where comparison #1 found a strong positive correlation for the masker present and $\Delta TEOAE_m$ data (R^2 = ~62%), but may have been driven by some of the subjects who had much longer $\Delta TEOAE_m$ responses. There was a non-existent relationship (i.e., R^2 = ~0%) in comparison #2, masker absent and $\Delta TEOAE_{m+p}$. The results from the individual comparisons of time constants are in contrast to the results from question 2 (above), where the average time constants were similar to one another, both approximately ~80 ms.

5.2 Individual Subject Variability

From the results in Chapters 2, 3, and 4, it is clear that there is a lot of individual subject variability when measuring gain reduction psychoacoustically and physiologically. This is true for the magnitude effects with the elicitor, the range of individual time constants, and the direct comparisons made with the LMM. By using the LMM to analyse the individual subject variation as a random factor, it was shown that it should be accounted for in future studies. Traditional correlational analysis cannot account for this variability, and it has been completely overlooked when studies compare physiology to behavior (see section 1.7 for review). Multiple studies have documented the individual variability when studying MOCR effects, behaviorally (e.g., Jennings et al., 2009) and physiologically (Backus and Guinan, 2007; Lilaontikul and Guinan, 2012; Marshall et al., 2014). This has practical implications on the interpretation of the effects found in the literature, where many studies use low subject sizes, and statistical design that does not account for this variance, and large conclusions are drawn from the results. To improve the power of the statistical inferences made, future studies on the MOCR should use parametric designs that allow for subject variability to be accounted for and recruit enough subjects for their comparisons between physiology and behavior.

5.3 Other Comparisons of MOCR

The current study showed that individual time constants were similar when measured behaviorally and phyisologically, (both approximately ~80 ms) when phase was accounted for in the response (section 4.3.3). However, the relationship between the two measures when comparing their magnitude (section 4.3.1) or time constants (section 4.3.2) directly either did not relate well (i.e., no relationship) or were negatively correlated with one another despite the measures to improve the study design. This is not the only study where a similar conclusion has been made. For example, Lichtenhan et al., (2016) measured compound action potentials (CAP) and DPOAEs in humans, and measured the effective attenuation of the respective signal with and without a contralateral elicitor. While there were only 5 subjects in that study, it can be seen in their figure 4, that the overall change in the DPOAE by the elicitor was not predictive of the change in the CAP. They note that their responses were quite variable overall, similar to individual and group results found in this study. Puria et al., (1996) also studied the affective attenuation of the CAP and the DPOAE, with and without the presence of a contralateral elicitor in 5 cats (2 anesthetized during recording, and 3 were decerebrated and unanesthetized). Very large differences were found between the two measures, and they were non-predictive of one another in terms of strength, similar to the results Lichtenhan et al., study. Despite the differences between the methods used to measure the MOCR in those studies, the mechanisms in each of those measures appear to be consistent with MOCR, and not other effects such as MEMR or excitatory masking. Therefore, there may be other factors that need to be accounted for, depending on the stimuli used, attentional state of the subject, physiological differences across subjects, and others factors that lead to inconsistency or non-predictive results when using independent measures of the MOCR effects. Reliability of MOCR measures is a growing area of interest (e.g., Backus and Guinan, 2007; Mertes and Goodman, 2016; Marrufo-Pérez et al., 2021), and seems to be gaining traction, especially when comparing physiological and behavioral data. In conclusion, the fact that two independent measures of the MOCR don't relate well in a single study doesn't imply that one or both of the measures is invalid or incorrect in its measurement of MOCR effects.

5.4 Implications of MOCR Time Constants in the Real World

As stated in section 1.6, the MOCR time course may be important for certain aspects of speech perception in noisy backgrounds. In the current study, both psychoacoustic (Chapter 2) and physiological (Chapter 3) measures of gain reduction produced relatively short time constants with a broadband elicitor. In terms of speech intelligibility, the MOCR may reduce the neural response to the background noise, thereby increasing the overall SNR of the target speech. The current study used short duration tonal stimuli, and broadband elicitors. It would be interesting to use more realistic stimuli in testing, and show that speech reception thresholds improve with MOCR activation. As stated in section 1.6, modelling studies that implemented MOCR-like effects into their model have reported that different time constants are associated with better speech intelligibility for different SNRs. Some parts of speech benefit from a longer time constant (Yasin et al., 2018), while other parts of speech benefit from a shorter time constant (Yasin et al., 2020). Hearing aid and cochlear implant users have a reduced dynamic range, and may not be able to adjust to background noise in the same way that someone without hearing loss can. Understanding the temporal dynamics of the MOCR with basic stimuli in typical hearing people (such as the current study) is a good starting point, as different time constants could be implemented in hearing devices depending on the input to the hearing aid or cochlear implant. Perhaps a ‘one size fits all’ efferent inspired time constant wouldn’t be beneficial for all SNRs in these devices, and more psychoacoustic and modelling data would be needed to determine this. It is likely that the effect of the MOCR is severely weakened in subjects with SNHL, as more afferent input is needed to achieve MOCR stimulation. This is in addition to the fact that individuals who suffer from SNHL already have a reduced dynamic range due to loss of gain and compression from OHC loss/damage, making the chance for adjustment by the MOCR even less likely. In conclusion, understanding gain adjustment by the MOCR may have the potential to benefit speech perception in noise (Brown et al., 2010; Clark et al., 2012), and depending how fast (or slow) this activation occurs may have a direct impact on if the speech is understandable (Yasin et al., 2020).

REFERENCES

Abdala, C., Guérit, F., Luo, P., & Shera, C. A. (2014). Distortion-product otoacoustic emission reflection-component delays and cochlear tuning: Estimates from across the human lifespan. *The Journal of the Acoustical Society of America, 135*, 1950-1958.

Aguilar, E., Eustaquio-Martin, A., and Lopez-Poveda, E. A. (2013). Contralateral efferent reflex effects on threshold and suprathreshold psychoacoustical tuning curves at low and high frequencies. *Journal of the Association for Research in Otolaryngology*, 14, 341-357.

Allen, J. B. (1986). Measurement of eardrum acoustic impedance. In *Peripheral auditory mechanisms* (pp. 44-51). Springer, Berlin, Heidelberg.

Almishaal, A., Bidelman, G. M., and Jennings, S. G. (2017). Notched noise precursors improve detection of low-frequency amplitude modulation. *The Journal of the Acoustical Society of America, 141*, 324-333.

Babadi, B., & Brown, E. N. (2014). A review of multitaper spectral analysis. *IEEE Transactions on Biomedical Engineering*, *61*, 1555-1564.

Backus, B. C., and Guinan, J. J., Jr. (2006). Time-course of the human medial olivocochlear reflex. *The Journal of the Acoustical Society of America, 119*, 2889-2904.

Backus, B. C., & Guinan, J. J. (2007). Measurement of the distribution of medial olivocochlear acoustic reflex strengths across normal-hearing individuals via otoacoustic emissions. *Journal of the Association for Research in Otolaryngology*, *8*, 484-496.

Bacon, S. P., and Healy, E. W. (2000). Effects of ipsilateral and contralateral precursors on the temporal effect in simultaneous masking with pure tones. *The Journal of the Acoustical Society of America, 107*, 1589-1597.

Bacon, S. P., and Savel, S. (2004). Temporal effects in simultaneous masking with on- and off-frequency noise maskers: Effects of signal frequency and masker level. *The Journal of the Acoustical Society of America, 115*, 1674-1683.

Bacon, S. P., & Takahashi, G. A. (1992). Overshoot in normal-hearing and hearing-impaired subjects. *The Journal of the Acoustical Society of America*, *91*, 2865-2871.

Bassim, M. K., Miller, R. L., Buss, E., & Smith, D. W. (2003). Rapid adaptation of the 2f1–f2 DPOAE in humans: binaural and contralateral stimulation effects. *Hearing research*, *182*, 140-152.

Bates, D., Sarkar, D., Bates, M. D., & Matrix, L. (2007). The lme4 package. *R package version*, *2*, 74.

Bidelman, G. M., Jennings, S. G., & Strickland, E. A. (2015). PsyAcoustX: A flexible MATLAB® package for psychoacoustics research. *Frontiers in psychology*, *6*, 1498.

Bharadwaj, H. M., Hustedt-Mai, A. R., Ginsberg, H. M., Dougherty, K. M., Muthaiah, V. P. K., Hagedorn, A., & Heinz, M. G. (2021). Cross-Species Experiments Reveal Widespread Cochlear Neural Damage in Normal Hearing. *bioRxiv*.

Boothalingam, S., & Purcell, D. W. (2015). Influence of the stimulus presentation rate on medial olivocochlear system assays. *The Journal of the Acoustical Society of America*, *137*, 724-732.

Boothalingam, S., Kurke, J., & Dhar, S. (2018). Click-evoked auditory efferent activity: Rate and level effects. *Journal of the Association for Research in Otolaryngology*, *19*, 421-434.

Brown, G. J., Ferry, R. T., & Meddis, R. (2010). A computer model of auditory efferent suppression: implications for the recognition of speech in noise. *The Journal of the Acoustical Society of America*, *127*, 943-954.

Carney, L. H. (2018). Supra-threshold hearing and fluctuation profiles: Implications for sensorineural and hidden hearing loss. *Journal of the Association for Research in Otolaryngology, 19*, 331–352.

Champlin, C. A., and McFadden, D. (1989). Reductions in overshoot following intense sound exposures. *The Journal of the Acoustical Society of America, 85*, 2005–2011.

Clark, N. R., Brown, G. J., Jürgens, T., and Meddis, R. (2012). A frequency-selective feedback model of auditory efferent susppression and its implications for the recognition of speech in noise. *The Journal of the Acoustical Society of America*, *132*, 1535-1541.

Collet, L., Kemp, D. T., Veuillet, E., Duclaux, R., Moulin, A., and Morgon, A. (1990). Effect of contralateral auditory stimuli on active cochlear micro-mechanical properties in human subjects. *Hearing Research, 43*, 251-261.

Cooper, N. P., and Guinan, J. J., Jr. (2006). Efferent-mediated control of basilar membrane motion. *Journal of Physiology*, *576*, 49-54.

Dean, I., Harper, N. S., and McAlpine, D. (2005). Neural population coding of sound level adapts to stimulus statistics. *Nature Neuroscience, 8,* 1684-1689.

DeRoy Milvae, K., and Strickland, E. A. (2018). Psychoacoustic measurements of ipsilateral cochlear gain reduction as a function of signal frequency. *The Journal of the Acoustical Society of America, 143*, 3114-3125.

DeRoy Milvae, K., and Strickland, E. A. (2021). Behavioral Measures of Cochlear Gain Reduction Depend on Precursor Frequency, Bandwidth, and Level. *Frontiers in Neuroscience*, *15*.

Feeney, M. P., & Keefe, D. H. (2001). Estimating the acoustic reflex threshold from wideband measures of reflectance, admittance, and power. *Ear and hearing*, *22*, 316-332.

Feeney, M. P., Keefe, D. H., Hunter, L. L., Fitzpatrick, D. F., Garinis, A. C., Putterman, D. B., and McMillan, G. P. (2017). Normative wideband reflectance, equivalent admittance at the tympanic membrane, and acoustic stapedius reflex threshold in adults. *Ear and Hearing, 38*, e142–e160.

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

Fletcher, M. D., Krumbholz, K., and de Boer, J. (2016). Effect of contralateral medial olivocochlear feedback on perceptual estimates of cochlear gain and compression. *Journal of the Association for Research in Otolaryngology*, *17*, 559-575.

Francis, N. A., & Guinan Jr, J. J. (2010). Acoustic stimulation of human medial olivocochlear efferents reduces stimulus-frequency and click-evoked otoacoustic emission delays: Implications for cochlear filter bandwidths. *Hearing Research*, *267*, 36-45.

Goodman, S. S., Mertes, I. B., Lewis, J. D., & Weissbeck, D. K. (2013). Medial olivocochlear-induced transient-evoked otoacoustic emission amplitude shifts in individual subjects. *Journal of the Association for Research in Otolaryngology*, *14*, 829-842.

Groon, K. A., Rasetshwane, D. M., Kopun, J. G., Gorga, M. P., & Neely, S. T. (2015). Air-leak effects on ear-canal acoustic absorbance. *Ear and Hearing*, *36*, 155.

Guinan, J. J., Jr. (1996). Physiology of olivocochlear efferents, in The Cochlea (Springer, New York), pp. 435–502.

Guinan, J. J., Jr. (2006). Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and Hearing*, *27*, 589-607.

Guinan, J. J., Backus, B. C., Lilaonitkul, W., & Aharonson, V. (2003). Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. *Journal of the Association for Research in Otolaryngology*, *4*, 521-540.

Hicks, M. L., and Bacon, S. P. (1992). Factors influencing temporal effects with notched-noise maskers. *Hearing Research*, *64*, 123–132.

James, A. L., Harrison, R. V., Pienkowski, M., Dajani, H. R., and Mount, R. J. (2005). Dynamics of real time DPOAE contralateral suppression in chinchillas and humans. *International Journal of Audiology*, *44*, 118-129.

Jennings, S. G. (2021). The role of the medial olivocochlear reflex in psychophysical masking and intensity resolution in humans: A review. *Journal of Neurophysiology*, *125*, 2279-2308.

Jennings, S. G., Strickland, E. A., and Heinz, M. G. (2009). Precursor effects on behavioral estimates of frequency selectivity and gain in forward masking. *The Journal of the Acoustical Society of America*, *125*, 2172-2181.

Jennings, S. G., and Strickland, E. A. (2012). Evaluating the effects of olivocochlear feedback on psychophysical measures of frequency selectivity. *The Journal of the Acoustical Society of America, 132*, 2483-2496.

Jürgens, T., Clark, N. R., Lecluyse, W., & Meddis, R. (2016). Exploration of a physiologically inspired hearing-aid algorithm using a computer model mimicking impaired hearing. *International Journal of Audiology, 55*, 346-357.

Kawase, T., Delgutte, B., and Liberman, M. C. (1993). Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. *Journal of Neurophysiology, 70*, 2533-2549.

Kawase, T., Ogura, M., Hidaka, H., Sasaki, N., Suzuki, Y., and Takasaka, T. (2000). Effects of contralateral noise on measurement of the psychophysical tuning curve. *Hearing Research, 142*, 63-70.

Keefe, D. H., Ling, R., & Bulen, J. C. (1992). Method to measure acoustic impedance and reflection coefficient. *The Journal of the Acoustical Society of America*, *91*, 470-485.

Keefe, D. H., Schairer, K. S., Ellison, J. C., Fitzpatrick, D. F., & Jesteadt, W. (2009). Use of stimulus-frequency otoacoustic emissions to investigate efferent and cochlear contributions to temporal overshoot. *The Journal of the Acoustical Society of America*, *125*, 1595-1604.

Keefe, D. H., Fitzpatrick, D., Liu, Y. W., Sanford, C. A., and Gorga, M. P. (2010). Wideband acoustic-reflex test in a test battery to predict middle-ear dysfunction. *Hearing Research, 263*, 52-65.

Keefe, D. H., Feeney, M. P., Hunter, L. L., & Fitzpatrick, D. F. (2017). Aural acoustic stapedius-muscle reflex threshold procedures to test human infants and adults. *Journal of the Association for Research in Otolaryngology, 18*, 65-88.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. *The Journal of the Acoustical Society of America*, *64*, 1386-1391.

Kenward, M. G., & Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, 983-997.

Kiang, N. Y. S., Watanabe, T., Thomas, E. C., and Clark, L. F. (1965). Discharge Patterns of Single Fibers in the Cat’s Auditory Nerve, Research Monograph No. 35 (Massachusetts Institute of Technology, Cambridge, MA).

Kim, D. O., Dorn, P. A., Neely, S. T., & Gorga, M. P. (2001). Adaptation of distortion product otoacoustic emission in humans. *Journal of the Association for Research in Otolaryngology*, *2*, 31.

Kochkin, S. (2000). MarkeTrak V: "Why my hearing aids are in the drawer" The consumers' perspective. *The Hearing Journal*, *53*, 34-36.

Krull, V., and Strickland, E. A. (2008). The effect of a precursor on growth of forward masking. *The Journal of the Acoustical Society of America, 123*, 4352-4357.

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

Lilaonitkul, W., and Guinan, J. J. (2009a). Human medial olivocochlear reflex: Effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. *Journal of the Association for Research in Otolaryngology, 10*, 459-470.

Lilaonitkul, W., and Guinan, J. J., Jr. (2009b). Reflex control of the human inner ear: A half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking. *Journal of Neurophysiology, 101*, 1394-1406.

Lilaonitkul, W., and Guinan, J. J., Jr. (2012). Frequency tuning of medial olivocochlear-efferent acoustic reflexes in humans as functions of probe frequency. *Journal of Neurophysiology, 107,* 1598-1611.

Liu, F., & Demosthenous, A. (2020). Effect of Time Constant on Speech Enhancement in Hearing Aids Based on Auditory Neural Feedback. In *2020 IEEE International Symposium on Circuits and Systems (ISCAS)* (pp. 1-5). IEEE.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., & Wilson, B. S. (2016). A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear and Hearing*, *37*, e138.

Maison, S., Micheyl, C., Andeol, G., Gallego, S., and Collet, L. (2000). Activation of medial olivocochlear efferent system in humans: Influence of stimulus bandwidth. *Hearing Research, 140*, 111-125.

Marrufo-Pérez, M. I., Johannesen, P. T., & Lopez-Poveda, E. A. (2021). Correlation and reliability of behavioral and otoacoustic-emission estimates of contralateral medial olivocochlear reflex strength in humans. *Frontiers in Neuroscience*, *15*, 122.

Marshall, L., Lapsley Miller, J. A., Guinan, J. J., Shera, C. A., Reed, C. M., Perez, Z. D., ... & Boege, P. (2014). Otoacoustic-emission-based medial-olivocochlear reflex assays for humans. *The Journal of the Acoustical Society of America*, *136*, 2697-2713.

McFadden, D., and Champlin, C. A. (1990). Reductions in overshoot during aspirin use. *The Journal of the Acoustical Society of America, 87*, 2634–2642.

Mertes, I. B., & Goodman, S. S. (2016). Within-and across-subject variability of repeated measurements of medial olivocochlear-induced changes in transient-evoked otoacoustic emissions. *Ear and Hearing*, *37*, e72-e84.

Mertes, I. B., Johnson, K. M., & Dinger, Z. A. (2019). Olivocochlear efferent contributions to speech-in-noise recognition across signal-to-noise ratios. *The Journal of the Acoustical Society of America*, *145*, 1529-1540.

Mertes, I. B., Wilbanks, E. C., & Leek, M. R. (2018). Olivocochlear efferent activity is associated with the slope of the psychometric function of speech recognition in noise. *Ear and Hearing*, *39,* 583.

Micheyl, C., Perrot, X., & Collet, L. (1997). Relationship between auditory intensity discrimination in noise and olivocochlear efferent system activity in humans. *Behavioral Neuroscience*, *111*, 801.

Mishra, S. K., & Abdala, C. (2015). Stability of the medial olivocochlear reflex as measured by distortion product otoacoustic emissions. *Journal of Speech, Language, and Hearing Research*, *58,* 122-134.

Murugasu, E., and Russell, I. J. (1996). The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. *Journal of Neuroscience, 16*, 325–332.

Nelson, D. A., Schroder, A. C., and Wojtczak, M. (2001). A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *110*, 2045-2064.

Neely, S. T., and Zhiqiang Liu. EMAV: Otoacoustic emission averager. *Boys Town National Research Hospital Tech Memo* 17 (1994).

Oghalai, J. S. (2004). The cochlear amplifier: augmentation of the traveling wave within the inner ear. *Current Opinion in Otolaryngology & Head and Neck Surgery*, *12*, 431.

Oxenham, A. J. (2001). Forward masking: Adaptation or integration? *The Journal of the Acoustical Society of America*, *109*, 732-741.

Oxenham, A. J., and Moore, B. C. (1994). Modeling the additivity of non-simultaneous masking. *Hearing Research, 80*, 105–118.

Oxenham, A. J., and 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*, 3666-3675.

Penner, M. J., and Shiffrin, R. M. (1980). Nonlinearities in the coding of intensity within the context of a temporal summation model. *The Journal of the Acoustical Society of America, 67*, 617-627.

Plack, C. J., and O'Hanlon, C. G. (2003). Forward masking additivity and auditory compression at low and high frequencies. J. Assoc. Res. Otolaryngol. 4, 405–415.

Prerau, M. J., Brown, R. E., Bianchi, M. T., Ellenbogen, J. M., & Purdon, P. L. (2017). Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. *Physiology*, *32*, 60-92.

Puria, S., Guinan Jr, J. J., & Liberman, M. C. (1996). Olivocochlear reflex assays: effects of contralateral sound on compound action potentials versus ear-canal distortion products. *The Journal of the Acoustical Society of America*, *99*, 500-507.

R Core Team (2020). R: A language and environment for statistical computing," R Foundation for Statstical Computing, Vienna, Austria. *https://www.R-project.org/*

Rhode, W. T. (1977). Some observations on two-tone interaction measured with the Mössbauer effect. *Psychophysics and Physiology of Hearing*, 27-41.

Robles, L., & Ruggero, M. A. (2001). Mechanics of the mammalian cochlea. *Physiological Reviews*, *81*(3), 1305-1352.

Roverud, E., and Strickland, E. A. (2010). The time course of cochlear gain reduction measured using a more efficient psychophysical technique. *The Journal of the Acoustical Society of America*, *128*, 1203–1214.

Roverud, E., and Strickland, E. A. (2014). Accounting for nonmonotonic precursor duration effects with gain reduction in the temporal window model. *The Journal of the Acoustical Society of America*, *135*, 1321-1334.

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*, 2151-2163.

Salloom, W. B., & Strickland, E. A. (2021). The effect of broadband elicitor laterality on psychoacoustic gain reduction across signal frequency. *The Journal of the Acoustical Society of America*, *150*, 2817-2835.

Schaalje, G. B., McBride, J. B., & Fellingham, G. W. (2002). Adequacy of approximations to distributions of test statistics in complex mixed linear models. *Journal of Agricultural, Biological, and Environmental Statistics*, *7*, 512-524.

Schmidt, S., & Zwicker, E. (1991). The effect of masker spectral asymmetry on overshoot in simultaneous masking. *The Journal of the Acoustical Society of America*, *89*, 1324-1330.

Shera, C. A., & Guinan Jr, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. *The Journal of the Acoustical Society of America*, *105*, 782-s798.

Shera, C. A., & Bergevin, C. (2012). Obtaining reliable phase-gradient delays from otoacoustic emission data. *The Journal of the Acoustical Society of America*, *132,* 927-943.

Strickland, E. A., Salloom, W. B., and Hegland, E. L. (2018). Evidence for gain reduction by a precursor in an on-frequency forward masking paradigm. *Acta Acustica United with Acustica, 104*, 809–812.

Strickland, E. A. (2001). The relationship between frequency selectivity and overshoot. *The Journal of the Acoustical Society of America, 109,* 2062-2073.

Strickland, E. A. (2004). The temporal effect with notched-noise maskers: Analysis in terms of input–output functions. *The Journal of the Acoustical Society of America, 115,* 2234-2245.

Strickland, E. A. (2008). The relationship between precursor level and the temporal effect. *The Journal of the Acoustical Society of America, 123,* 946–954.

Todd, R. M., & Cruz, J. R. (1996). Frequency Estimation and the QD Method. *Control and Dynamic Systems*, *75*, 1-78.

Turner, C. W., and Doherty, K. A. (1997). Temporal masking and the 'active process' in normal and hearing-impaired listeners. *Modeling Sensorineural Hearing Loss* (Taylor & Francis, London), pp. 387-396.

Viemeister, N. F. (1983). Auditory intensity discrimination at high frequencies in the presence of noise. *Science, 221,* 1206-1208.

Vinay, and Moore, B. C. J. (2008). Effects of activation of the efferent system on psychophysical tuning curves as a function of signal frequency. *Hearing Research, 240*, 93-101.

Walsh, K. P., Pasanen, E. G., & McFadden, D. (2010). Overshoot measured physiologically and psychophysically in the same human ears. *Hearing Research*, *268,* 22-37.

Warr, W. B. (1992). Organization of olivocochlear efferent systems in mammals, in The Mammalian Auditory Pathway: Neuroanatomy (Springer, New York), pp. 410–448.

Wicher, A., and Moore, B. C. (2014). Effect of broadband and narrowband contralateral noise on psychophysical tuning curves and otoacoustic emissions. *The Journal of the Acoustical Society of America*, *135,* 2931-2941.

Winslow, R. L., and Sachs, M. B. (1988). Single-tone intensity discrimination based on auditory-nerve rate responses in backgrounds of quiet, noise, and with stimulation of the crossed olivocochlear bundle. *Hearing Research, 35*, 165-189.

Yasin, I., Drga, V., and Plack, C. J. (2014). Effect of human auditory efferent feedback on cochlear gain and compression. *Journal of Neuroscience, 34*, 15319-15326.

Yasin, I., Liu, F., Drga, V., Demosthenous, A., & Meddis, R. (2018). Effect of auditory efferent time-constant duration on speech recognition in noise. *The Journal of the Acoustical Society of America*, *143*, EL112-EL115.

Yasin, I., Drga, V., Liu, F., Demosthenous, A., & Meddis, R. (2020). Optimizing speech recognition using a computational model of human hearing: effect of noise type and efferent time constants. *IEEE Access*, *8*, 56711-56719.

Zwicker, E. (1965). Temporal effects in simultaneous masking by white-noise bursts. *The Journal of the Acoustical Society of America*, *37*, 653-663.

VITA

William Salloom

wsalloom@purdue.edu

Education	**PhD Candidate, Integrative Neuroscience**, PULSe Interdisciplinary Program, Purdue University - Speech Language & Hearing Sciences ***2015 –* 2022**, *West Lafayette*, IN **B.A. Psychology**, California State University Northridge, ***2014, Northridge, CA. Cum Laude*** **A.A. Criminal Justice**, Los Angeles Pierce College, ***2012, Woodland Hills, CA***
Research Experience	**Graduate Research Assistant - Dr. Elizabeth Strickland (Advisor, Psychoacoustics Lab)** **Dr. Hari Bharadwaj (Co-Advisor, SNAP Lab)** *Purdue University, 2016 - 2022* Psychoacoustics Otoacoustic Emissions Wideband Accoustic Immittance tympanometry MATLAB Auditory Neuroscience **Undergrad Research Assistant – Dr. Aïda Metzenberg (Genetic Laboratory – Undergrad research assistant)** *Cal State Northridge, 2012-2014* Worked on a vector that can be transfected to *E. coli* to Neurofibromatosis-type 1 gene Experience with molecular genetic techniques Professional lab meetings, presentations, colloquia
Publications	**Salloom, W.B.,** Strickland, E.A. (**2021**). The effect of broadband elicitor laterality on psychoacoustic gain reduction across signal frequency. *The Journal of the Acoustical Society of America*, *150*, 2817-2835. **https://doi.org/10.1121/10.0006662**

Strickland, E.A., **Salloom, W.B.,** Hegland, E.L. (2018). "Evidence for gain reduction by a precursor in an on-frequency forward masking paradigm," in *Special Issue on Hearing: Psychophysics, Physiology, and Models,* Acta Acustica united with Acustica, Volume 104, pp. 809-812. **DOI10.3813/AAA.919229**

Research Presentations

Salloom, W.B., "A Comparison of Gain Reduction Estimated from Behavioral Measures and Various Transient Evoked Otoacoustic Emission Measures as a function of Broadband Elicitor Duration**"** Departmental Hearing Seminar in Speech Language and Hearing Sciences (SLHS) at Purdue University, March 3rd, 2022.

Salloom, W.B., Wade. K, Bharadwaj. H.B., Strickland, E.A., A Comparison of Gain Reduction Estimated from Behavioral Measures and Various Transient Evoked Otoacoustic Emission Measures as a function of Broadband Elicitor Duration" Poster session presented at: 45rd Annual Midwinter Meeting of the Association of Research in Otolaryngology; 2022 Feb 5-9; (Virtual).

Salloom, W.B., "The effect of broadband elicitor duration on **presentations** transient-evoked otoacoustic emissions and a behavioral measure of gain reduction" Departmental Hearing Seminar in Speech Language and Hearing Sciences (SLHS) at Purdue University, Sept 14th, 2021.

Strickland, E. A. (Presenter), Skaggs, M., Hopkins, A., Mielnicki, N., **Salloom, W. B.**, Morris, H., Holt, A., "A Summary of Behavioral Measures of Cochlear Gain and Gain Reduction in Listeners & with Normal Hearing or Minimal Cochlear Hearing Loss" Acoustical Society of America (virtual), May 2021.

Salloom, W.B., Wade, K., Bharadwaj, H.B., Strickland, E.A., "The effect of broadband elicitor duration on transient-evoked otoacoustic emissions and a behavioral measure of gain reduction" Departmental "Brown bag" Seminar in Speech, Language, and Hearing Sciences (SLHS) at Purdue University, 2020, Mar 2nd.

Salloom, W.B., Wade, K., Bharadwaj, H.B., Strickland, E.A., "The effect of broadband elicitor duration on transient-evoked otoacoustic emissions and a behavioral measure of gain reduction" Departmental Seminar in Speech, Language, and Hearing Sciences (SLHS) at Indiana University, 2020, Feb 28th.

Salloom, W.B., Wade, K., Bharadwaj, H.B., Strickland, E.A., "The effect of broadband elicitor duration on transient-evoked otoacoustic emissions and a behavioral measure of gain reduction" Departmental Hearing Seminar in Speech Language and Hearing Sciences (SLHS) at Purdue University, 2020, Feb 27th.

Salloom, W.B., Bharadwaj, H.B., Strickland, E.A., (2020). "The effect of broadband elicitor duration on transient-evoked otoacoustic emissions and a behavioral measure of gain reduction". Poster session presented at: 43rd Annual Midwinter Meeting of the Association of Research in Otolaryngology; 2020 Jan 25-29; San Jose, CA.

Salloom, W.B., Oral presentation on Jennings, Strickland, & Heinz (2009) paper, "Precursor effects on behavioral estiates on frequency selectivity and gain in forward masking". Auditory Neuroscience Journal Club (ANJC) at Purdue University, 2019, Nov 4th.

Salloom, W.B., Bharadwaj, H.M., Strickland, E.A., "The Effect of Broadband Elicitor Duration on Transient-Evoked Otoacoustic Emissions and Psychoacoustic Measure of Gain Reduction". *Acoustical Society of America, Louisville, KY, 2019.*

Salloom, W.B., Bharadwaj, H.M., Strickland, E.A., "Physiological and Psychoacoustic Measures of Two Different Auditory Efferent Systems". Departmental Hearing Seminar in Speech Language and Hearing Sciences (SLHS) at Purdue University, 2019, Mar 7th.

Salloom, W.B., Bharadwaj, H.M., Strickland, E.A., (2019). **"**Physiological and Psychoacoustic Measures of Two Different Auditory Efferent Systems**".** Poster session presented at: 42nd Annual Midwinter Meeting of the Association of Research in Otolaryngology; 2020 Feb 9-13; Baltimore, MD.

Salloom, W.B., Bharadwaj, H.M., Strickland, E.A., "Physiological and Psychoacoustic Measures of Two Different Auditory Efferent Systems". Departmental "Brown bag" Seminar in Speech Language and Hearing Sciences (SLHS) at Purdue University, 2019, Feb 4th.

Strickland, E.A. (Presenter), **Salloom, W.B.,** Hegland, E.L. (2018). "Evidence for gain reduction by a precursor in an on-frequency forward masking paradigm," Talk presented at the International Symposium on Hearing, July, Snekkersten, Denmark.

Strickland, E. A., Morris, H., Skaggs, M., **Salloom, W. B.**, Holt, A. "Behavioral Measures of Cochlear Gain and Gain Reduction in Listeners with Normal Hearing or Minimal Cochlear Hearing Loss". *Acoustical Society of America, Minneapolis, MN, May 2018.*

Salloom, W.B., Strickland, E.A., "The Effect of an Ipsilateral, Contralateral, and Bilateral & Precursor on Gain Reduction Across the Frequency Range". *Association of Research in Otolaryngology*, San Diego, CA, *Feb 10-15, 2018.*

Teaching Experience

Fall 2019 – Teaching Assistant
Course: Psychoacoustics
Graduate level course
Responsibilities include teaching portions of the course and grading exams

Fall 2018 – Teaching Assistant
Course: Psychoacoustics
Graduate level course
Responsibilities include teaching portions of the course and grading assignments/exams

Awards

2020 Ruth L. Kirschstein Predoctoral Individual National Research Service Award – NIH F31 Fellowship (**Submitted & Scored, not funded**)

2019 Ruth L. Kirschstein Predoctoral Individual National Research Service Award – NIH F31 Fellowship (**Submitted & Scored, not funded**)

2019 Interdisciplinary Training Program in Auditory Neuroscience (TPAN) – NIH T32 Fellowship, Purdue University

2018 Purdue Institute for Integrative Neuroscience (PIIN) Travel Grant

2016 Most outstanding first-year poster award, Purdue University

2015 Lynn Fellowship, Purdue University