Effects of selective attention on the peripheral auditory system: Otoacoustic emission assays of cochlear function during behavior

A THESIS SUBMITTED TO THE FACULTY OF THE UNIVERSITY OF MINNESOTA BY

Jordan A. Beim

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Thesis Advisor

Andrew J. Oxenham

March 2019

Acknowledgments

I want to express profound gratitude for the mentorship of my committee chair, Dr. Magdalena Wojtczak. She taught me many of the foundational skills I need to be a scholar. Her constant support of my academic career has built a lasting relationship that I will continue to treasure.

I would also like to thank my advisor Dr. Andrew Oxenham for his guidance and support. His brilliance as a scholar has left a lasting impression that will continue to shape my career. Without the talented community of students and researchers Andrew has assembled at the Auditory Perception and Cognition lab my education would not have been possible. I'm grateful for committee member Dr. Yuhong Jiang. She provided valuable comments on my research program and provided a set of visual stimuli that were integral to completing my thesis project. I thank Peggy Nelson for her service on my committee, her encouraging mentorship, and for her passion for education that inspires me to mentor students of my own.

Lastly, I am indebted to my family: my parents for their support throughout my education and for helping to instill the inquisitive spirit that drives my scholarship, and my wife for late nights, proof reading, being a test subject, and walking beside me on this journey.

Abstract

Selective attention is an important ability that humans use to navigate a crowded sensory world. In audition, attending to a target while ignoring distractors results in an enhancement of the neural representation of the target and an attenuation of distractors. The neural underpinnings as well as the location of this process within the auditory pathway is not clear. This thesis explores the possibility that the medial olivocochlear reflex, an auditory efferent neural circuit, might aid in selective attention by modulating the gain produced by the cochlea during selective attention. Otoacoustic emissions, small sounds originating in the cochlea, are used to measure changes in cochlear function while participants perform tasks that require attention to either only auditory or only visual stimuli. Experiments assess the replicability of otoacoustic measures made during selective attention, the effect of the perceptual load of attention tasks on these otoacoustic measures, and whether the effects of attention differ with respect to location within the cochlea. Across five experiments results consistently show no evidence for changes in cochlear function during selective attention to auditory or visual stimuli. The results challenge much of the existing literature, which reports small but significant effects of attention, but often in conflicting directions across studies. The results highlight significant variability between individuals in otoacoustic measures of cochlear function, which remains to be explained.

Contents

List of Figures

Chapter 1: A review of the effects of selective attention on peripheral auditory processing

I. Introduction

The study of the effects of selective attention on the auditory system spans several decades, from early electrophysiological work in animals in the middle of the 20th century to complex neuroimaging and otoacoustic emission studies today. While our understanding has advanced considerably, a question that remains unanswered is how selective attention produces the observed modulations in sensory processing. Several studies have attempted to answer this question in audition, by suggesting that efferent pathways extending from the cortex to the cochlea could modulate our auditory system in response to attention. Research from numerous laboratories using otoacoustic emissions and other physiological methods has produced conflicting results and small effects, leaving the debate unanswered. In this review, research on otoacoustic emission measurement methodology, auditory efferent effects, and selective attention is examined to expose a number of flaws in the existing literature on peripheral effects of auditory and cross-modal selective attention. The review proposes new experimental methodologies that incorporate the integration of this research to provide a more definitive answer to the question of how selective attention can modulate peripheral auditory processing. These methodologies are then implemented in the studies described in the remainder of the thesis.

II. Otoacoustic emissions: Mechanisms for generation and measurement methods

Otoacoustic emissions (OAEs) are weak sounds generated within the cochlea. Since their discovery by Kemp (1978), OAEs have become the subject of intense study, as well as a tool for researchers and clinicians to non-invasively probe the human auditory system. The use of otoacoustic emissions as a clinical and research tool is dependent on our understanding of the mechanisms responsible for their generation. The following section provides a brief overview of the current understanding of the mechanisms responsible for generating the various types of OAEs as well as acceptable methods for their extraction and analysis.

Otoacoustic emissions can be classified into two primary categories that relate the emission to the stimulus evoking them: Spontaneous OAEs (SOAEs) and evoked OAEs (EOAEs). Spontaneous OAEs are sounds that can be recorded in the ear canal in the absence of any evoking stimulus.

SOAEs are thought to arise as a result of multiple internal reflections of sound within the cochlea (Shera, 2003). Briefly, a sound reflected from a micromechanical irregularity within the cochlea produces a second reflection when the backward travelling wave reaches the oval window. This secondary reflection travels forward and is amplified by the active process of the cochlea and again reflects backward toward the oval window. When the amplification produced by the cochlea is sufficient, a persistent steady-state sound is generated and escapes as a spontaneous otoacoustic emission. Up to 72% (62% of males, 83% of females) of the normal hearing population has measurable SOAEs (Penner and Zhang, 1997). Although the majority of SOAEs are too weak in intensity to

be audible, research estimates that a small percentage of people (e.g. 6-12%, Norton,

Schmidt, & Stover, 1990) may experience their SOAEs as a form of tinnitus.

In addition to tinnitus, SOAEs have been studied in their relationship to cochlear

pathology (e.g. Ruggero, Rich, & Freyman, 1983)**,** human audiometric threshold

microstructure (e.g. Baiduc, Lee, & Dhar, 2014), frequency selectivity (e.g. Hansen,

Santurette, & Verhulst, 2014), and their interaction with evoked emissions (e.g. Norrix &

Glattke, 1995; van Dijk & Wit, 1990). The results suggest that spontaneous otoacoustic

emissions should always be measured before making measurements of evoked emissions,

so that any possible interactions can be avoided.

Evoked otoacoustic emissions are emissions of acoustic energy that are caused by

the presentation of sound to the cochlea and are named by their properties of generation

or the method used to evoke the emission. A large body of research has focused on

understanding the mechanisms responsible for generating evoked OAEs. The current

consensus about the generation of evoked emissions is that there are two primary sources

of evoked emissions: nonlinear distortion and coherent reflection. The following section

details the types of evoked OAEs discussed in the current literature with respect to their

mechanism of generation within the inner ear.

A. Distortion-product otoacoustic emissions

Distortion-product otoacoustic emissions (DPOAEs) are the primary type of

distortion source OAE. DPOAEs are typically evoked by presenting a combination of

two primary evoking tones. Due to the nonlinearity of cochlear processing, the tones

interact on the basilar membrane to create mechanical distortions. These distortions act as

sources for backward travelling waves at frequencies different from the primary tones. The largest and most commonly analyzed distortion frequency in humans is described by the expression $f_{dp} = 2f_1 - f_2$, where f_1 and f_2 are the frequencies of the primary tones.

While the nonlinear distortion acts as a source for backward travelling f_{dp} waves, it also generates forward travelling f_{dp} waves. These waves travel forwards towards the f_{dp} characteristic frequency place where coherent reflection produces an additional backward travelling wave that adds to the distortion source component. This addition of the two DPOAE components causes fluctuations in DPOAE magnitude as the primary frequencies are varied, which are referred to in the literature as DPOAE fine structure.

DPOAEs are widely used in scientific research as well as in the clinic to assess outer hair cell (OHC) function (Robinette and Glattke, 2007). One reason underlying the popularity of the use of DPOAEs is the ease in the recording and analysis of the emission. This is in large part due to the fact that the emission occurs at a frequency that is not present in the evoking stimulus, and is therefore easy to separate from the evoking stimulus when compared to other types of emission. Depending on the nature of the experiment, DPOAEs can also present several disadvantages. The combination of the two generator components means that changes in DPOAE magnitude over frequency more are difficult to interpret without separating these sources. For instance, phase shifts in one of the components can dramatically change the magnitude of the total emission. Additionally, since the emission comes from a frequency not present in the stimulus, experiments studying the effects of attention or listening on DPOAEs may underestimate changes that occur since the listener is often not attending to f_{dp}, but rather to f_1 or f_2.

B. Stimulus frequency otoacoustic emissions

Stimulus frequency otoacoustic emissions (SFOAEs) are one of three different types of reflection source emissions when measured at low levels (Shera & Guinan, 1999). SFOAEs are generated at the same frequency of the evoking stimulus, and are thought to arise due to coherent reflections of the forward travelling stimulus from micro-mechanical irregularities distributed throughout the basilar membrane (Zweig and Shera, 1995)**.** Unlike DPOAEs, SFOAEs are more difficult to separate from their evoking stimuli because the emission and evoking stimulus the overlap in frequency and in time. There are multiple methods for separating SFOAEs from their evoking stimuli that produce nearly identical estimates of emissions in humans (Kalluri and Shera, 2007). These include using: nonlinear compression (e.g. Keefe, 1998; Kemp & Chum, 1980), suppression (e.g. Puria, Guinan, & Liberman, 1996), and spectral smoothing (e.g. Kalluri & Shera, 2001).

The nonlinear compression method exploits the fact that SFOAEs grow nonlinearly with increasing stimulus amplitude due to cochlear mechanics, such that at high levels the proportion of pressure owing to the emission will be negligible, and the proportion of ear canal pressure due to otoacoustic emission will increase with decreasing stimulus pressure. This measurement method is accomplished by subtracting the response of the double-amplitude stimulus from the sum of two half-amplitude stimuli (Keefe, 1998).

The suppression method utilizes a suppressor tone at a nearby frequency (often within 50 Hz, e.g. Shera & Guinan, 1999) to disrupt the site of emission generation via two-tone suppression, eliminating the contribution of the emission to the total pressure

while the suppressor tone is on. This source pressure is subtracted from the total pressure as recorded in the ear to extract the SFOAE.

Spectral smoothing relies on the use of a smoothing function to remove the fluctuations due to the emission that occur in the total pressure across frequency in order to estimate the stimulus pressure.

Although the majority of the reflected wave is likely to come from a place near the peak of the basilar membrane response to the forward travelling wave, recent OAE data (Charaziak and Siegel, 2015; Moleti et al., 2013) and modelling (Sisto et al., 2015) suggest that some portion of the reflections originate from sources distributed more basally than the peak response. This was at first interpreted as a flaw in the coherent reflection theory of emission generation (Siegel et al., 2005), but subsequent research has suggested that these basally distributed emission sources may represent secondary sources in addition to the coherent reflections near the peak of the forward travelling wave (Shera et al., 2008). These additional sources could be a nonlinear distortion source (as in DPOAEs), additional reflections from sources that are widely basally distributed throughout the cochlea (as argued by Siegel et al., 2005)**,** nonlinear reflections (Talmadge et al., 2000)**,** or additional components that arise near the tail of the travelling wave. Modelling work (Sisto et al., 2015) and some human SFOAE data (Beim and Wojtczak, 2014) support the idea that additional reflections may be coming from an area near the tail region of the travelling wave, while animal OAE data (Charaziak and Siegel, 2015) provide evidence supporting basal emission generators at sites much further away than the tail of the travelling wave.

C. Transient-evoked otoacoustic emissions

Transient evoked otoacoustic emissions (TEOAEs) are stimulus frequency emissions evoked using transient sounds with durations short enough that the emission does not temporally overlap with the stimulus itself. Like SFOAEs, TEOAEs measured at low levels are likely generated through coherent linear reflections of the forward travelling wave from micromechanical irregularities concentrated near the peak excitation of the travelling wave. When measured at high levels, TEOAEs also contain an additional component likely generated by nonlinear distortions.

Evidence for this secondary distortion component can be seen using latency domain approaches similar to those used to separate sources of DPOAEs. (Kalluri and Shera, 2001) TEOAEs are measured using short tone bursts (TBOAEs), clicks (CEOAEs), or chirps (Neumann et al., 1994). Like DPOAEs, TEOAEs are used clinically to assess hearing function due to their ease of measurement. TEOAE measurement typically involves rapidly presenting the transient stimuli (e.g. Kemp & Chum, 1980; Collet et al., 1990) which can result in faster tests than would be obtainable using DPOAEs or SFOAEs. TEOAE magnitudes are often smaller than DPOAEs and SFOAEs and while they are considerably more broadband in frequency, most of the acoustic energy in TEOAEs decays above 4 kHz, making higher frequency measurements difficult.

Based on the development of the theories for otoacoustic emission generation in the normal cochlea, researchers have used evoked emissions to non-invasively study several aspects of cochlear function. One area that has seen much investigation since the

development and refinement of OAE generation theories is the study of the medial

olivocochlear reflex (MOCR).

III. The medial olivocochlear reflex: measurements of efferents and their function

The MOCR is an acoustically elicited efferent reflex comprised of a series of

neuronal projections descending from the medial part of the superior olivary complex and

synapsing directly with the OHCs in the Organ of Corti (Rasmussen, 1953; Warr &

Guinan, 1979). The basic pathway of the MOCR as reviewed by Guinan (2006) is as

illustrated in Fig. 1.1: (1) Auditory nerve fibers innervate MOCR interneurons in the

cochlear nucleus (CN). (2) MOCR interneurons project from the CN across the midline

of the brainstem innervating MOC neurons in the olivocochlear bundle (OCB) on the

contralateral side of the brainstem. There are additional ipsilateral MOC neurons that are

innervated by MOCR interneurons projecting from the contralateral cochlear nucleus. (3)
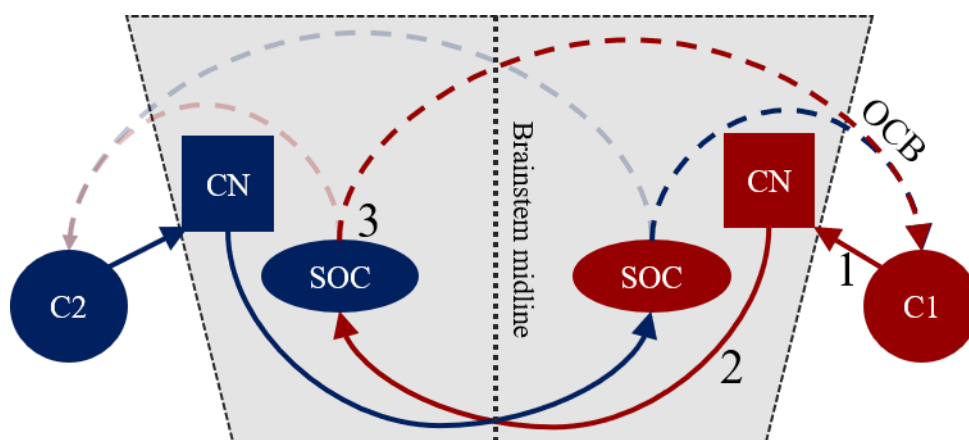
Figure 1.1

Simplified schematized representation of the MOCR. Numbers correspond to major pathway components described with corresponding numbers in the text. Abbreviated shapes denote key auditory system structures: cochlea (C1 & C2), cochlear nucleus (CN), and superior olivary complex (SOC). Shaded trapezoidal region illustrates which structures and pathways are part of the auditory brain stem. Solid lines indicate afferent neural pathways, while dashed lines represent efferent pathways. Red shapes and pathways illustrate the ipsilateral reflex, while blue corresponds to the contralateral pathways.

The OCB neurons cross back over the midline, innervating the OHCs in the ipsilateral cochlea. An important implication from the anatomy of this pathway is that auditory stimulation from one ear has the potential to modulate OHC activity in both ears.

While the MOCR pathway to the ipsilateral cochlea is mediated by both ipsilateral and contralateral input, research suggests that the relative contributions of each pathway are not equal. In small mammals, as reviewed by Warr, (1992), the ipsilateral reflex contains more neurons than the contralateral reflex by a ratio of 2-3:1. This ratio varies by species and also by the frequency region of OHCs innervated by the MOC neurons, where the higher-frequency OHCs are typically innervated by greater numbers of ipsilateral MOC fibers than at lower frequencies (Warr, 1992). These animal anatomical data are also supported by tests of the reflex strength via single unit recordings and changes in OAEs that show that ipsilateral reflexes in these animals are between 2 and 3 times stronger than their contralateral counterparts (Gifford and Guinan, 1987). The difference in strength between ipsilateral and contralateral MOC reflexes may not be as pronounced in humans. Measurements of the MOCR in humans using OAEs shows little if any systematic difference in the strength of ipsilateral and contralateral reflexes on the change in OAE magnitude in response to an acoustic elicitor (Guinan et al., 2003). This is similar to neuronal counts in some new world monkeys, where neuronal populations for each side of the reflex are within 10% of each other (Thompson and Thompson, 1986). Consideration of the anatomy of the MOCR has led to a large body of research devoted to understanding the functional consequences of its activation.

Activation of the MOCR has been demonstrated to inhibit the active amplification process of the OHCs in the cat and other animals through experiments investigating the

effect of electrically activating the reflex on compound action potentials (CAP) and the effect of acoustically activating the reflex on OAEs (reviewed by Guinan, 1996). This inhibition reduces the gain produced by the cochlear amplifier. The effects of activating the reflex have been shown to be frequency specific (Lilaonitkul & Guinan, 2012) with considerably broader tuning than that of the afferent auditory pathway. There two primary effects of stimulating the MOC efferent system: fast effects and slow effects. Fast effects result in inhibition of basilar membrane motion building up over tens of milliseconds, while slow effects build up over tens of seconds (Sridhar et al., 1995). Recent work examining the effect of MOC activation on SFOAEs has provided some evidence that the recovery from MOC activation, shown previously to be in the hundreds of milliseconds at 1-kHz (Backus and Guinan, 2006) is actually frequency dependent with decay time constants increasing with increasing frequency (Walsh and Wojtczak, 2014)**.** The behavioral consequences and utility of this frequency dependence of recovery are not yet understood.

In humans, questions concerning the MOCR are often investigated experimentally with the use of OAEs because they are non-invasive, and their generation relies on the function of the cochlear amplifier (which is modulated by the MOCR). Studies of the MOCR in humans using OAEs have used all of the types of OAE discussed previously, but some types of evoked emission may be less suitable for accurate measurement of MOCR activity due to differences in the evoking stimuli. Guinan et al. (2003) reviewed the caveats associated with the various evoked emissions and argued that SFOAEs may be the most suited to making accurate measurements of the MOCR. This is in part due to the difficulty in interpretation of changes in DPOAEs, as the direction of effects can vary

widely for a given level of MOCR activation, and because conventional measurements of TEOAEs involve stimulus levels and presentation rates that may elicit the MOCR or the middle ear muscle reflex (MEMR) in baseline conditions. SFOAEs require only one low-level tone stimulus to be measured and are the least likely of the evoked emissions to evoke either the MOCR or MEMR.

These studies of the effects of MOCR activation have led to a plethora of hypotheses concerning the functional utility of the reflex. Potential functional roles of the MOCR include: protecting basic hearing function (Liberman, Liberman, & Maison, 2014; Maison, Usubuchi, & Liberman, 2013; Rajan, 1988), enhancing sensitivity to transient sound in background noise (Kawase et al., 1993), and providing a potential pathway for attentional modulation of the peripheral auditory system (Lukas, 1980; Maison, Micheyl, & Collet, 2001). Researchers have also implicated the MOC efferent system as an explanation for various potential psychophysical effects such as: the auditory enhancement effect (Viemeister, 1980), the overshoot effect (e.g. Jennings, Heinz, & Strickland, 2011), and off-frequency phase effects in forward masking by Schroeder phase complexes (Wojtczak and Oxenham, 2009) among others. However, direct tests of these effects has not provided strong support for MOCR involvement in the enhancement effect (Beim et al., 2015) or in the phase effects seen by Wojtczak and Oxenham (2009) as investigated by Wojtczak, Beim, & Oxenham (2014).

Direct evidence of cortical connection to the MOC system (Mulders and Robertson, 2000), confirmed functionally by Dragicevic et al. (2015), has sparked further interest and debate over the role of the efferent system in selective attentional modulation of the auditory periphery. Before reviewing the relevant research investigating the role of

the MOC efferent system in selective attention, it is important to review the current

theories underlying selective attention and its observed effects.

IV. Attention: Selection theories, perceptual load, and auditory selective attention

Attention is a widely studied cognitive process with profound impact on our daily

lives (Broadbent, 1952, 1958; James, 1890; Posner, 1980). As reviewed by Knudsen

(2007), attention is the cognitive process that humans and other animals use to select the

most behaviorally relevant information from their environment. Knudsen describes

attention as the interrelation between four fundamental component processes: working

memory, competitive selection, top-down sensitivity control, and salience filtering. The

end result of the attentional processes of competitive selection, sensitivity control, and

salience filtering is that relevant information achieves entrance to working memory,

where decisions about the information can be made and actions planned. Competitive

selection refers to the selection of relevant information for entry into working memory,

and is mediated by a combination of top-down sensitivity control, and bottom-up salience

filtering (Knudsen, 2007). Top-down sensitivity control is the modulation of the strength

of representation of competing sources of information based on goals, while bottom-up

saliency filtering is defined as the enhancement of highly different, rewarding, or noxious

stimuli in the afferent information pathway. A wide breadth of research focuses on each

of these four component processes, but the following review will focus on research

related to the selection of target information among competing distractors and the current

understanding of the effects of top-down sensitivity control on the representation of

selected stimuli.

All the processes described above operate on perceptual objects. In the auditory

system, a perceptual object can be loosely defined as sound originating from a distinct

source in the environment (e.g. Shinn-Cunningham & Best, 2008). In contrast to vision,

where objects can be thought of in terms of static images, auditory objects from the same

source (e.g. individual words spoken by a single talker) group together across time to

form a stream (Bregman, 1990). In the everyday world, perceptual objects or streams

often overlap and the sensory systems must be able to segregate perceptual objects

correctly for competitive selection to occur. In the auditory system, the formation and

segregation of perceptual objects relies on the spectrotemporal characteristics of the

varying stimuli (Bregman, 1990). Grouping cues include: similar pitches, intensities,

timbres, spatial cues, simultaneous onsets, similar temporal modulations over time,

reverberation, and aspects of signal structure that are learned through experience (e.g.

linguistic properties) (Fritz et al., 2007; Shamma et al., 2011; Shinn-Cunningham, 2008)**.**

The correct formation of auditory objects or streams is itself not sufficient for selection to

take place. Each object or stream must have a property that can differentiate it from other

competing objects and the listener must have knowledge about what this property is in

order for successful selection to proceed. In audition, the process of identifying objects

and segregating streams can be largely automatic, builds up over time, and is based on

the strength of cues used to segregate the objects, such that more distinct objects or

streams will be segregated more quickly (Bregman, 1990; Shinn-Cunningham, 2008;

Shinn-Cunningham and Best, 2008)**.** Once objects or streams are sufficiently segregated,

they can be selected as targets for attention. However, the process should not be

considered as being strictly serial, as attention can affect the process and build-up of segregation (Carlyon et al., 2001, 2003).

When the listener/observer has access to the features of the object they wish to attend to, top-down sensitivity control (Knudsen, 2007) enhances the neural representation of objects that have the desired feature(s) to bias the competition between objects in favor of the targets for selection. This enhancement is often seen as increased neural activity in the cortical areas responsible for processing the desired feature within a modality (e.g. phonetic vs spatial information modulates activity in "what" vs "where" pathway, Ahveninen et al., 2006). Similar effects are seen when selective attention is applied across modalities (e.g. Johnson & Zatorre, 2006). While there is a rapidly growing body of research on the locations and types of changes brought about by top-down sensitivity control, the mechanisms for the enhancement of target features are still not clear. Understanding the locus and mechanisms behind the top-down sensitivity control has great impetus, and debate surrounding potential mechanisms has been persistent in the field of attention for half a century.

Early research on the locus of attentional control focused on the idea that perception itself was a limited process and that selection of relevant stimuli was necessary to avoid overloading our perceptual systems. This idea was first advanced by Broadbent (1958) who found that when presented with multiple sources of auditory information, listeners demonstrated that there was a limit to the amount of information they could actively make use of. This was accomplished by presenting different spoken questions to each ear over headphones. When listeners were asked to answer one of the questions afterward performance dropped compared to when they knew which question

to listen to beforehand. Broadbent proposed that selection acted to filter out unnecessary perceptual input so that it would not be processed beyond the level necessary to identify it as necessary or unnecessary. This work was advanced by Treisman (1969) who showed that rather than being discarded, the representations of unattended stimuli were likely only attenuated compared to that of the attended stimuli. This theory of attenuation of distractors during perception is called early selection, as stimuli to be attended were selected early in the processing stream from sensory neurons to conscious awareness. If selection operates early, enhancements to relevant stimuli should be (and are) seen in early cortical areas. The human auditory pathway contains subcortical sites of information processing (Webster et al., 1992), which could be loci for top-down sensitivity control.

A later group of researchers posited "late-selection" as an alternative theory. Late selection theories centered on the idea that perception was automatic and parallel, without requiring the selection of relevant stimuli until a behavioral response was required. A critical distinction from early selection theory is that all stimuli, rather than only the relevant stimuli, are processed fully. These ideas were advanced by the work of Deutsch and Deutsch (1963) and Norman (1968). The research supporting the idea of late selection comes from studies showing that unattended or irrelevant stimuli produce an influence on behavior towards attended stimuli (e.g. Miller, 1987; Stroop, 1935). Studies supporting this theory typically used a paradigm that required subjects to detect or identify a target stimulus presented alone or amongst a few distractors. In visual attention research, this often took the form of detecting a target letter presented with some kind of noise stimulus (e.g. other letters, Eriksen & Eriksen, 1974). The key findings from these

studies show that the presence of irrelevant distractors increases the time required to make a correct identification of the target.

The debate between researchers in the early selection and late selection camps has produced clear evidence in support of both models of attentional selection under certain circumstances. Differences between results supporting late or early selection were often attributed to differences in experimental design possibly causing the use of different attentional systems (Kahneman and Treisman, 1984). An attempt to reconcile both viewpoints in a common framework is provided by the perceptual load theory of selective attention and cognitive control. (Lavie & Tsal, 1994; Lavie, 1995, 2005). The perceptual load theory combines ideas from both the late and early selection theories. As in late selection, the perceptual load theory assumes that perception is automatic and parallel, but similar to early selection, assumes that there is a limit to the information that can be simultaneously processed in full detail. The amount of perceptual load a task requires is often defined by either the number of distinct stimuli presented to a subject during an experimental trial (set size), or by the amount of processing required by the task (feature vs conjunctive search) (Lavie, 1995).

When load is high, selection is expected to take place early, and when load is low, irrelevant stimuli are processed to greater depth automatically. This greater depth in processing allows irrelevant stimuli to produce the distraction effects demonstrated by the research supporting late selection. Perceptual load experiments demonstrate behaviorally that as load increases the distraction (typically measured as an increase in reaction time) produced by irrelevant stimuli decreases (Lavie, 1995).

The effects of distractors seen in neuroimaging studies are attenuated or completely eliminated under conditions of high perceptual load. Studies of motion processing show that responses of motion-sensitive cortex to irrelevant motion in the background of a visual task appear only in low-load conditions. (Rees et al., 1997) The elimination of distractor effects has been shown in earlier visual cortical areas (e.g. primary visual cortex Rees et al., 1997) as well as subcortical structures (e.g. Lateral Geniculate Nucleus; see O'Connor, Fukui, Pinsk, & Kastner (2002). A high load task in audition showed that primary auditory cortex contained representations of target, but not distractor speech stimuli (Mesgarani and Chang, 2012). Physiological manifestations of the effects of selective attention have been seen at the level of the auditory nerve in animals by Delano, Elgueda, Hamame, & Robles (2007), who reported that compound action potentials (a gross report of activity from the auditory nerve) elicited by tones were decreased when chinchillas were performing a visual discrimination task compared to when they performed an auditory discrimination task.

While it is now commonly accepted that there exist cortical and possibly subcortical modulations in the representation of attended or unattended stimuli, the direct mechanisms are not yet understood. More research is needed in both vision and audition to carefully examine the precise location where attentional modulations begin to occur before potential mechanisms can be elucidated. In the auditory system, a growing body of literature suggests that these attentional modulations may begin at the earliest stage of sound processing, the cochlea. The evidence supporting this claim is still the subject of debate and more work is necessary to resolve differences in outcomes observed by different research groups. As will be reviewed in the following section, many

experimental investigations of the peripheral auditory system under selective attention have demonstrated small and conflicting results. One of the reasons responsible for this is that many investigators did not consider the perceptual load imposed by their tasks, nor the impacts that perceptual load may have in finding results.

V. Peripheral effects of auditory selective attention: Summary and pitfalls
A. Summary of previous findings

Since the discovery of the olivocochlear bundle (Rasmussen, 1946, 1953), there has been continued interest in whether or not the effects of selective attention can extend as far peripherally as the cochlea. One of the first such studies compared electrical activity in the cochlear nucleus (CN) of cats in response to tone pips when the cats were passive or attending to non-auditory stimuli (Hernández-Peón et al., 1956). The authors found that the response to tones was reduced or eliminated when cats were attending to olfactory or tactile stimuli, when compared to the passive condition. This work was followed by several other investigations demonstrating some support for attentional modulations of peripheral auditory processing. Two other electrophysiological investigations found decreased amplitude and increased latency in the 5th wave of the auditory brainstem response (Lukas, 1980) and a reduction in the compound action potential of the auditory nerve (Delano et al., 2007) in response to visual attention. In contrast, other studies (e.g. Picton, Hillyard, Galambos, & Schiff, 1971) have shown a lack of attentional effect in the auditory periphery. Picton et al. (1971) found that there were no differences in the externally recorded auditory nerve response when human subjects attended to clicks or read a book. This study did, however, show enhancements

in auditory evoked potentials, a cortical measure, when listeners attended to the auditory stimuli.

Some of the only evidence from psychophysical studies in support of the involvement of the OCB in came from case studies of patients who had an OCB section during the removal of a neuroma (Scharf et al., 1994, 1997). In both studies, Scharf and colleagues found few differences in psychophysical performance before and after an olivocochlear bundle (OCB) section. The only task that produced a difference in results before and after surgery was a task involving selective attention. Before surgery, participants demonstrated good detection for a cued target tone, and poor detection for tones at different frequencies from the cue. After surgery, patients detected both cued and unexpected tones with equal performance, evidence for a widening of the attentional filter with respect to frequency, and a potential side effect of the OCB sectioning. The strength of this evidence is limited, as it is not known whether the entire OCB was successfully sectioned by the surgery, and additional effects may be hidden by some residual MOC efferent activity.

In addition to electrophysiological and psychophysical studies, several investigations have used have used otoacoustic emissions as non-invasive probes of cochlear function (Ferber-Viart et al., 1995; Froehlich et al., 1993; Giard et al., 1994; Lukas, 1980; Meric and Collet, 1992; Puel et al., 1988). These studies all measured decreases in TEOAE magnitude in conditions of visual attention compared to a passive condition where listeners were not required to attend to the auditory stimuli. The decreases observed in these studies were between 0.2 and 1.2 dB. The changes observed in these conditions cannot be attributed solely to selective attention because other factors

like general arousal due to the comparison between a task and no task condition. In addition being very small effects, they remain controversial, because other studies investigating attentional changes in OAEs have failed to find effects (Avan and Bonfils, 1992; Michie et al., 1996).

Avan and Bonfils (1992) measured changes in SFOAE and DPOAE when human subjects were performing a visual task as in Lukas (1980) and when they were passive. The authors tried to control for the lack of frequency specificity of the CEOAEs used in Lukas (1980) by using SFOAEs and DPOAEs (which provide stimulation to a more limited frequency range) to probe cochlear function. The authors reported no difference in either SFOAE or DPOAE magnitudes when subjects were performing the visual attention tasks compared to no task. A second study compared OAEs evoked when subjects were performing an auditory task or a visual task and also found no consistent effects of selective attention on OAE magnitudes (Michie et al., 1996). The only significant effects observed were small increases in OAE magnitude in response to ignoring the auditory stimuli in some of the experimental conditions.

The conflicting evidence surrounding a potential role of the MOC efferent system in selective attention has left the issue an open question. More recent studies have showed significant differences in OAEs when subjects are attending to visual vs auditory stimuli (Srinivasan, Keil, Stratis, Woodruff Carr, & Smith, 2012; Srinivasan et al., 2014; Walsh, Pasanen, & McFadden, 2014a, 2014b, 2015). Srinivasan et al. (2012, 2014) saw shifts in DPOAE magnitudes when people were attending to an auditory compared to a visual task. In contrast with most other studies, the authors found increases in DPOAE magnitudes from ignored stimuli. This was the case both when subjects were attending

visual stimuli, or stimuli in the opposite ear. Like in previous reports, the size of these

attentional effects was small (approx. 0.25 dB). Walsh et al. (2014a, 2014b, 2015) found

differences in a nonlinear component of SFOAE and in the physiological noise when

participants were attending to auditory or visual stimuli, relative to a no-attention

condition. Like much of the previous research using OAEs to explore effects of attention,

Walsh et al. noted decreases in OAE magnitude under attentive conditions relative to

inattentive conditions, but no differences between auditory and visual attention. The

studies by Walsh et al. (2014a, 2014b, 2015) show a promising departure from previous

work on the effects of selective attention on the peripheral auditory system. Their well-

controlled experiments demonstrated effects an order of magnitude larger than most of

the previous work, but a careful review of the selective attention literature reveals that

additional improvements may reveal yet even larger modulations of the auditory

periphery under selective attention task with high perceptual load.

 Overall, the majority of studies investigating the effects of attention on the

auditory periphery show that there are small changes in OAE magnitudes in response to

selective attention, but the amount of change is small enough to lead many researchers to

question the perceptual or neural-coding relevance of the effects. The effects may be

small due to the fact that OAEs may underestimate changes in basilar membrane motion

(Puria et al., 1996) or due to flaws in the experimental design of many of the studies, as

outlined below.

B. Pitfalls and limitations of previous research

Many of the findings of studies on effects of selective attention on the peripheral auditory system are limited by flaws in experimental or stimulus design. Four critical flaws that can be found in the earlier studies are reviewed below.

1. MOC efferent activity may be underestimated due to complications caused by OAE stimuli

Most of the early studies providing evidence for or against the effects of selective attention on OAEs used TEOAEs as the primary assay of cochlear function (Ferber-Viart et al., 1995; Froehlich et al., 1993; Giard et al., 1994; Lukas, 1980; Michie et al., 1996; Puel et al., 1988). In a review of OAE measurements of MOC effects, Guinan et al. (2003) found evidence that many types of OAE evoking stimuli can also elicit MOC activity themselves. OAE-evoking stimuli presented at high levels (> 50 dB SPL) or rapid rates (> 40 Hz) often elicited MOC effects and sometimes activated the middle ear muscle reflex, which also produces changes in OAE measured in the ear canal. TEOAEs were found to be the most potent elicitors of MOC efferents among the various OAE stimuli, while SFOAEs were the least likely to elicit MOC effects. Many of the early reports noted increases in MOC activity under attention conditions, and these effects may appear smaller if the MOC system is activated above baseline due to the OAE evoking stimuli. Only one group of researchers used SFOAEs in their experimental design (Walsh et al., 2015). Walsh et al. (2015) also report the largest attention related effects seen

among all of the previous research, but additional methodological differences prevent isolating the use of SFOAEs as the reason larger effects were seen.

Another limitation of the previous work is that many of the studies that acoustically elicited the MOC system used a contralateral elicitor (e.g. Michie et al., 1996). As mentioned above, animal research suggests that the ipsilateral MOCR is between 2 and 3 times stronger than the contralateral reflex due to the pattern of innervation of the OHCs (reviewed by Warr, 1992). Human research shows variability in the relative strengths of the ipsi- and contra-lateral reflexes, but bilateral elicitors evoke the greatest changes in OAE (Guinan et al., 2003).

Lastly, OAE extraction based on nonlinear methods (e.g. Walsh et al., 2015) only preserve MOC-induced changes that are nonlinear. If a significant portion of MOC-induced change on OAEs measured in selective attention is linear then these nonlinear methods may significantly underestimate the size of the effect.

2. DPOAEs are not an optimal stimulus for measuring effects of selective attention

Like SFOAEs, DPOAEs should be less likely than TEOAEs to elicit MOC activity (Guinan et al., 2003). Some of the previous OAE research on selective attention utilized DPOAEs as the assay of cochlear function (Avan and Bonfils, 1992; Srinivasan et al., 2012, 2014). A major limitation of the use of DPOAEs to assess changes in MOC activity is that changes in DPOAE magnitude can result from changes to either the distortion source or reflection source component (Guinan et al., 2003). Changes in the magnitude or phase of one of these components can produce increases or decreases in

DPOAE measured in the ear canal. The results of Srinivasan et al (2012, 2014) are in the opposite direction of results reported by most other studies, and this could be due to differential MOC effects on the reflection and distortion sources of the emission used. Another limitation in using DPOAEs to study selective attention is that the distortion frequency is remote from the primary frequencies that the listener may attend to in order to complete the attentional task. While the tuning of the MOC system has been shown to be relatively broad (Lilaonitkul & Guinan, 2012), measuring off frequency may still produce smaller MOC effects on DPOAE than might be shown with SFOAEs or if attention was directed towards the distortion frequency.

3. *Comparison of attentional state to a passive baseline*

In much of the early work on the peripheral effects of selective attention, researchers compared OAE or electrophysiological recording magnitudes when subjects were performing a visual task to a passive baseline (e.g. Avan & Bonfils, 1992; Lukas, 1980; Puel et al., 1988). The major limitation to the findings of such work is that the effects observed by the researchers could be due to differences in arousal or other factors that change between a passive and task-based condition, rather than attention. Additionally, when subjects are involved in a passive condition, they may choose to attend to at their discretion, which can also influence the size of the effect if it were possible to rule out non-attention factors. Later research corrected this limitation by requiring comparisons to utilize an attentional task in all conditions, (e.g. Giard et al.,

1994; Maison, Micheyl, & Collet, 2001) but avoiding this limitation remains important for future research.

4. The perceptual load in most studies was low

While many studies compared attentional performance in two or more task-based conditions, it is important to carefully examine the nature of the attentional tasks. In studies that examined visual attention, the visual tasks included: reading a book (e.g. Avan & Bonfils, 1992), monitoring a stream of letters to detect a target (e.g. count "Q" among "O", Lukas, 1980), monitoring a series of visual gratings for a change (e.g. Srinivasan et al., 2012), or memorizing a series of visually presented digits (e.g. Walsh et al., 2015). Auditory attentional tasks involved either monitoring clicks or tone pips for changes in intensity (e.g. Giard et al., 1994) or memorizing a series of digits presented as speech (Walsh et al., 2015). In most of these cases, the perceptual load of the attentional tasks was low. As discussed previously and reviewed by Lavie (2005), attentional tasks result in early selection (and potentially subcortical modulation of processing) only when the perceptual load of the task is high.

Perceptual load is distinct from task difficulty, so attempts to make attentional tasks more difficult by reducing the fidelity of sensory information (e.g. reducing the difference between stimuli to be discriminated, Michie et al., 1996) is unlikely to result in early selection. Increasing perceptual load means placing additional demands on the attentional system, and this is often accomplished by presenting targets simultaneously

with larger groups of distractor stimuli or by requiring targets to be selected based on

conjunction of features (e.g. red "t" instead of any red letter, Lavie, 2005).

Only one of the previous studies presented more complex distractor stimuli as part

of the attentional task (Walsh et al., 2015). Walsh and colleagues presented two

competing auditory streams, requiring subjects to select and attend one of the two streams

to complete the auditory task. This auditory attention condition is most similar to some of

the early work showing evidence for early selection (e.g. Broadbent, 1952, 1958). The

visual task used by Walsh et al. (2015) was somewhat analogous, presenting two streams

of visual digits in different locations on the screen. The visual task has somewhat lower

load than the auditory task since the visual stimuli were more easily segregated than the

auditory stimuli and there was only 1 distractor series. A complication in this study is that

the behavioral task required memorization of the digit strings, an imposition of cognitive

load, which is distinct from perceptual load and has been shown to have the opposite

effects as high perceptual load (Lavie, 2005).

VI. Thesis Overview

The work presented in this dissertation represents improvements in methodology

compared to previous research in two key areas: (1) Optimizations to otoacoustic

emission evoking stimuli to ensure robust, high SNR measurements of cochlear function

that are not contaminated by MEMR activity; (2) Incorporates findings from selective

attention literature to produce a behavioral task with high perceptual load that requires

sustained attention throughout a trial. These improvements in methodology are

implemented in a series of experiments presented in the following chapters in an attempt

to provide comprehensive evidence for or against the attentional modulation of cochlear function. Chapter 2 introduces a novel experimental paradigm in which SFOAE evoking stimuli are the target of attention in a behavioral task and the effects of selective attention are compared both across sensory modalities and within the auditory modality. Chapter 3 modifies the experimental design used in experiment 2 to include a visual task with high perceptual load and examines whether or not attentional effects depend on the presence of the MOCR elicitor. Chapter 4 examines whether or not changes in cochlear processing are more evident at higher frequency regions where higher cochlear gain could produce a larger dynamic range for MOCR effects. Finally, chapter 5 summarizes the findings of the thesis work, provides a brief meta-analysis of the results, and provides an outlook on prospects for further research.

Chapter 2: Examining replicability of an otoacoustic measure of cochlear function during selective attention[1]

I. Introduction

In auditory perception, solving the cocktail party problem (Cherry, 1953) relies on our ability to selectively attend to one target speaker among many distracting background sounds, including other speakers. Recent work has revealed that cortical representations of unattended speech are attenuated relative to attended speech (Mesgarani and Chang, 2012). These changes are robust enough to allow computational decoding of the attended speaker using cortical activity in a closed-set environment, even within a single experimental trial (e.g. Ding and Simon, 2012; Choi et al., 2013; O'Sullivan et al., 2015).

Although these changes in cortical activity are robust and well documented, it is unclear at what stage of auditory processing attentional modulation is first observed. From an anatomical standpoint, attentional effects could extend to the most peripheral part of the auditory nervous system, the cochlea, via the medial olivocochlear (MOC) efferent system. The MOC system is a network of neurons in the medial part of the superior olivary complex that receives input from afferent peripheral auditory neurons and from corticofugal projections. MOC efferent neurons innervate the outer hair cells (OHCs) in the cochlea (for review, see Guinan, 2006). The primary functional consequence of MOC efferent activity is a reduction of the cochlear gain produced by the OHCs.

A long line of research has investigated the role of the MOC system in human selective attention, mostly using otoacoustic emissions (OAEs; Kemp, 1978) as a

---

[1] The contents of this chapter are published as Beim et al. (2018)

noninvasive window into cochlear function (Avan and Bonfils, 1992; de Boer and

Thornton, 2007; Giard et al., 1994; Harkrider and Bowers, 2009; Maison et al., 2001;

Michie et al., 1996; Puel et al., 1988; Srinivasan et al., 2012; Walsh et al., 2015). These

studies have generally revealed small and inconsistent effects of selective attention, with

effects (if any) of only a fraction of a decibel. While small effects observed through

OAEs can translate to larger changes in the afferent neural signal (e.g. Puria et al., 1996),

the inconsistent direction of these effects makes it difficult to ascribe to them any

important functional role in attention and perception.

   Although these small or null results may indicate no role of the MOC system in

attention, it is also possible that the failure to observe large effects has been due to the

suboptimal choice of stimuli and procedures in previous studies. For instance, earlier

studies have often employed relatively high sound levels, which themselves could

activate the MOC system even in the absence of attentional manipulations. In addition,

the attention of the participants was not always well controlled, leading to uncertainty and

potential individual differences related to the degree to which attention was focused on or

away from the acoustic stimuli. Specifically, conditions in which listeners performed an

auditory or a visual task, typically referred to as attention conditions, were often

compared with conditions in which listeners were not performing a task, referred to as

inattention conditions. However, there is no guarantee that listeners did not pay attention

to the stimuli when no specific task was assigned to them.

   In the current study, we leveraged recent improvements in the measurement

techniques used to probe MOC effects, and implemented rigorous manipulations of

attention, to revisit the question of attentional modulation of human cochlear responses.

Stimulus-frequency otoacoustic emissions (SFOAEs) were recorded from participants while they were engaged in behavioral tasks requiring selective attention to either auditory or visual stimuli. Care was taken in the stimulus design to produce high signal-to-noise ratios (SNR) when measuring SFOAEs and to avoid stimuli that might inadvertently activate the MOC efferent system or the middle ear reflex, for reasons discussed below. Our new combination of stimulus and attentional manipulations revealed large and significant changes in human cochlear responses in response to changes in selective attention in the first group of participants (N=15). However, a subsequent attempt at replicating this experiment in a second group of participants (N=15) failed to produce the same results despite using improved instrumentation, although there remained a small but significant effect when the results from both groups were combined. Careful analysis of differences in SFOAE levels and SNRs resulting from the use of different measurement systems for the two groups of participants showed that the different outcomes could not be attributed to the differences in instrumentation. The large differences between individuals, despite relatively stable within-individual measurements, reveal inter-individual variability that remains unexplained. The outcomes highlight the need for large samples to aid in examining the large inter-individual variability observed in our study and across previous studies of attentional effects on cochlear responses.

II. **Methods**

A. Overview

Participants in this study were presented with a complex set of audiovisual stimuli consisting of sequences of low- and high-frequency tones, as well as series of visually presented letters on a computer monitor. Participants were required to complete three different tasks. The same stimuli (both auditory and visual) were present in each condition so that our critical manipulation was to vary the stimulus to which attention was directed in the behavioral task. The first two tasks required participants to attend to either a low- or high-frequency sequence of tones and to report at the end of each run how many of the tones in the target sequence were modulated in amplitude. Amplitude modulation (AM) detection was chosen for the behavioral task because the MOC efferent activation has been shown to affect responses to transient or fluctuating stimuli presented in noise (Guinan, 2010). A counting task was used instead of a simple detection task with immediate responses to reduce extraneous movement and noise, which could contaminate the recording of SFOAEs. The tones were embedded in a spectrally notched noise designed to activate the MOC efferent system, as previous studies that investigated attentional modulation of acoustically elicited efferent effects (Froehlich et al., 1990, 1993; Maison et al., 2001; Michie et al., 1996; Veuillet et al., 1991; Walsh et al., 2014a, 2014b) have reported significant attention effects more consistently than studies in which attention effects on cochlear responses were measured without acoustic efferent activation (Picton et al., 1971; Picton and Hillyard, 1974). We also believe this creates a more ecologically valid design, as real-world environments such as a noisy restaurant could easily contain noise levels which would acoustically elicit MOC activity. Our

working hypothesis was that attention modulates the efferent effect produced by the noise (an acoustic elicitor) rather than triggering efferent activation on its own (see Fig. 6 in Walsh et al., 2014a for support of this hypothesis). The third task required participants to direct their attention to the visual portion of the stimuli. Participants monitored a stream of random serially presented letters and were instructed to detect consecutive repetitions of the same letter within the series. The visual task was chosen to compare the auditory attention conditions with a non-auditory task that also required sustained attention.

To examine the effect of attention we compared the magnitudes of the SFOAEs evoked by low-frequency tones across the three attentional conditions. A low- rather than high-frequency tone was chosen to monitor attentional modulation of efferent effects on SFOAEs because larger acoustically elicited efferent effects are observed with SFOAEs evoked by tones with frequencies in the range of 500-1000 Hz than tones with higher frequencies (Lilaonitkul and Guinan, 2012). Comparisons across attentional conditions in this study provided a window into cochlear function when attention was directed towards the SFOAE-evoking tone, towards an auditory object remote in frequency, and away from the auditory stimuli entirely.

B. Participants

Initially 18 participants (7 male, 11 female) ages 20-46 (median: 24) were recruited to participate in the experiment. Three (all male) of the 18 participants had their data discarded due to a lack of robust SFOAEs, described in the screening procedure below, leaving a total of 15 participants (4 male, 11 female) in the first sample. A power analysis based on the variability of the data from initial participants suggested that a

sample size of 15 would be required to reliably detect a 5-dB change in SFOAE

magnitude between attention conditions. A second sample of 24 participants (5 male, 19

female) ages 18-34 (median: 21) was recruited based on the power analyses to assess the

reproducibility of the data from the first sample. A total of 15 participants (3 male, 12

female) in the second sample passed the screening procedures. Eight of the excluded

participants did not have robust SFOAEs in the region around the nominal probe-tone

frequency that was used to monitor effects of selective attention on cochlear responses.

One participant was excluded due to an inability to pass the behavioral-training

condition, described below. All participants were tested for normal hearing (air

conduction thresholds of 20 dB HL or better at octave frequencies between 0.25 and 8

kHz) using a calibrated Madsen Conera audiometer (GN Otometrics, Schaumberg, IL).

Participants provided written informed consent before taking part in the study and were

compensated for their time. All experimental procedures were approved by the

Institutional Review Board at the University of Minnesota.

C. Auditory Stimuli

Auditory stimuli were used both for evoking SFOAEs and for the auditory behavioral task. Ten blocks of task-relevant stimuli were grouped to form a single run (a single row of stimuli in Fig. 2.1). Every run began with a baseline SFOAE measurement utilizing the suppression method (Shera and Guinan, 1999). A 10-s probe tone at the frequency selected by the screening procedure (approximately 0.75 kHz) was presented at 35 dB SPL with a suppressor tone 50 Hz higher in frequency and presented at 55 dB SPL.
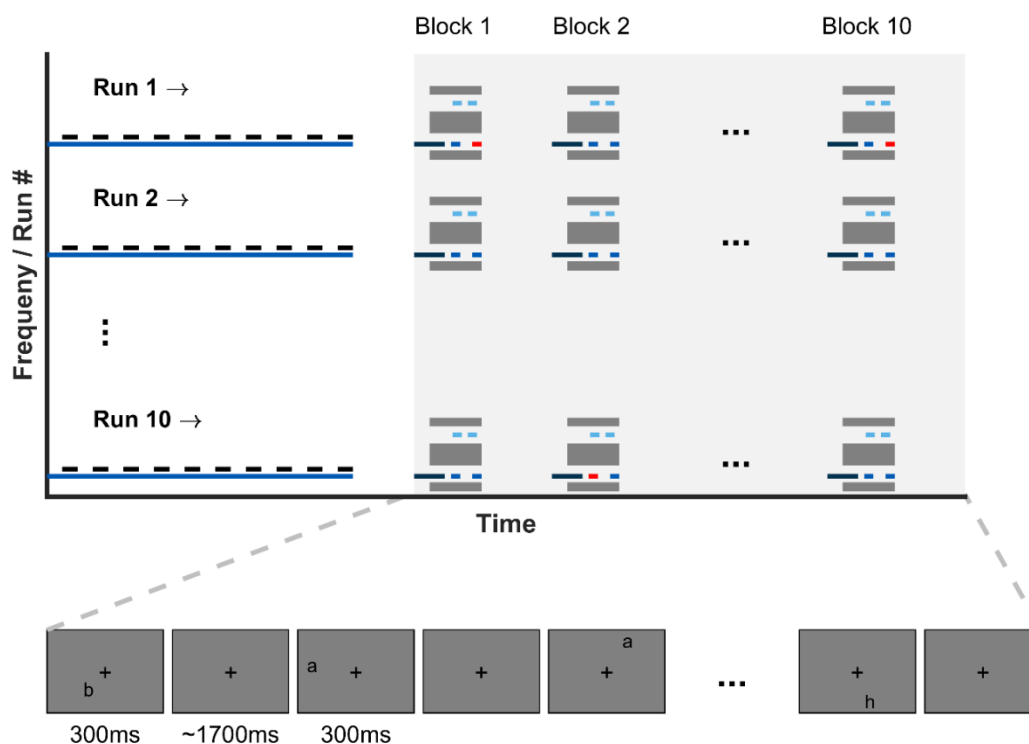

Figure 2.1

A schematic illustration of the stimuli used for evoking SFOAE and for the auditory and visual behavioral tasks. Time-frequency representations of auditory stimuli are shown across multiple runs (10 total runs per attention condition). The baseline SFOAE measured by a continuous probe is illustrated in dark blue. Successive presentations of the suppressor (in black) were presented in alternating phase. Task relevant auditory stimuli, depicted in the lightly shaded region, were presented simultaneously with the visual stimuli. A repeating sequence (block) of sound follows the baseline measurement. Each block contains a cue tone and 2 target tone pips (dark blue), 2 distractor tone pips (light blue) and notched noise (dark grey). Modulated target tones that the listener counted while participating in the AL condition are shown in red. The inset at the bottom is series of panels depicting the location and timing of visual stimuli on the computer screen. Each panel represents the display at one moment during visual stimulus presentation. The duration of each frame is presented on the screen is shown below the frame.

The suppressor tone was gated on and off every 0.5 s, starting 0.5 s after the onset of the 10-s tone. The baseline SFOAE magnitude was used to ensure that an SFOAE at the probe frequency has a sufficiently high SNR to observe robust efferent effects (Goodman et al., 2013).The baseline was also needed to ensure that SFOAE magnitude due to the acoustically elicited MOC reflex (MOCR) allowed room for attentional modulation of efferent effects (i.e., the cochlear gain reduction by the MOCR did not obliterate the probe-evoked SFOAE).

The task-relevant stimuli began after a 2-s silent gap following the offset of the probe tone. After the silent gap, a 1-s cue tone was presented at either the probe-tone frequency or at 4 kHz to indicate which frequency participants should attend to complete the perceptual auditory task (AM detection). A spectrally notched Gaussian white noise with 1-octave-wide notches centered on both the probe-tone frequency and on 4 kHz began 0.5 s after the onset of the cue tone and was presented for 2 s. The overall rms level of the noise was set to 55 dB SPL. The function of the noise was to activate the MOC reflex; according to our working hypothesis, attention should modulate the change in cochlear gain produced by MOCR activation. A total of four 0.3-s tone pips (two at each frequency) were presented within the duration of the noise in each block. Tone pips of the same frequency had randomly selected inter-stimulus intervals (ISIs) to ensure sustained attention by preventing listeners from attending only at the expected times. The ISIs were constrained to have a minimum duration of 100 ms and to ensure that the final-tone offset occurred no later than the offset of the noise. The pips were presented at 35 dB SPL. This low level reduced the possibility of the pips activating MOC efferents. It was important to avoid efferent activation by the pips because the efferent effect on

subsequent pips in a run would vary depending on the ISI and could introduce unwanted

variability and thus contaminate effects of attentional modulation. The tone pips and

noise were gated with 5-ms onset and offset cosine-squared ramps to reduce spectral

splatter. The high- and low-frequency tones could overlap in their presentation time so

that performance in the task would benefit from efferent inhibition of cochlear responses

in the unattended frequency region. The task-relevant stimuli ended with a 2-s silent

interval.

The target tone pips for the behavioral task (between one and five within each

run) were amplitude modulated for their full duration at a rate of 10 Hz with a modulation

depth of -5 dB expressed in units of $20\log(m)$, where m is the modulation index between

0 (no modulation) and 1 (100% modulation). A total of 10 recordings (from 10 runs)

were made for each experimental condition. A schematic representation of the auditory

stimuli is shown in the top section of Fig. 1. All auditory stimuli were presented to the

right ear of each participant. Visual stimuli shown at the bottom of Fig. 1 were also

presented, as described in detail below.

All stimuli were generated digitally with a sampling frequency of 48 kHz and a

24-bit depth using custom MATLAB software (The Mathworks, Natick, MA), converted

into analog signals using a LynxTwo-B sound card (Lynx Studio Technology, Costa

Mesa, CA), and delivered to participants via transducers in either an ER10C or ER10X

system (Etymotic Research, Elk Grove, IL). The ER10C system was used to present and

record stimuli for the first group of participants, while the ER10X system was used for

the replication sample. Recordings were made using the microphones of the ER10C or

ER10X systems. Each sound system provided a 20-dB gain to the microphone signal

before it was digitized by the LynxTwo-B sound card and saved on the computer hard

drive for offline analysis. Both sound delivery systems were calibrated with a Type 4153

artificial ear with a 2cc coupler (Bruel & Kjaer NA, Duluth, GA) to verify output sound

pressure in an average ear canal.

D. Visual Stimuli

The visual stimuli were presented concurrently with the auditory stimuli (i.e.

beginning after the 2-s gap following the baseline SFOAE measurement as shown in the

bottom row of Fig. 2.1) and consisted of 35 letters presented serially. The letters were

drawn randomly from the first 11 letters of the English alphabet. This subset of letters

was chosen because random sampling with replacement from this set produced on

average 3 pairs of consecutive letter repetitions (the same number as average target tone

modulations in auditory stimuli). Letters were displayed on the screen for 0.3 s with

random ISIs (mean 1.7 s) but the intervals were constrained so that the total visual

stimulus presentation time exactly matched the total duration of the auditory stimuli in a

run. The location of each letter on the screen was randomly selected. Letters were

presented at 80% contrast and subtended approximately 2 degrees of visual angle. The

horizontal and vertical dimensions of the screen (and hence possible locations of the

letters) subtended approximately 35 and 26.25 degrees, respectively. A schematic

illustration of the visual stimuli is shown in the bottom row of Fig. 1.

E. Procedure

All procedures described below were conducted in a double-walled sound attenuating chamber (Industrial Acoustics Company, Bronx, NY) with the participant seated comfortably in a semi-reclined position in a chair with a head rest to reduce movement-related artifacts. Participants kept a computer keyboard on their lap for responses at the end of each run to minimize movements between recordings and were instructed to remain as still and relaxed as possible without closing their eyes or falling asleep.

Before the testing began, participants underwent a two-part screening procedure to ensure that: 1) spontaneous otoacoustic emissions (SOAEs), if present, were remote from the experimental frequency range and 2) robust SFOAE measurements were obtainable within the range of the experimental stimuli. Spontaneous emissions were measured using a procedure described by Penner et al. (1993). To screen for robust emissions, SFOAEs were measured using the suppression method across a frequency range from 675-825 Hz in 15-Hz steps. In order to pass the screening, SFOAEs needed to be at least 20 dB greater than the measurement noise floor for at least one of the frequencies used and the phase coherence of SFOAE needed to be greater than 0.95. These parameters were chosen in order to maximize the SNR of the measurements, as high SNR measurements are necessary to reveal significant MOC efferent effects (Goodman et al., 2013). For any given participant, the frequency with the largest SFOAE magnitude in this screening was used as the SFOAE probe frequency in the experiment. Three participants were rejected from the total sample because they did not show

SFOAEs with SNR > 20 dB. No participants were rejected due to SOAEs near the experimental frequency range.

After passing screening, participants were initially exposed to a less complex set of the experimental stimuli to orient them to the task. First, audio examples with tone pips of just one frequency were played to participants, until they could correctly identify modulated tone pips to the experimenter. The distractor tone pips at the other frequency (with and without modulation) were then added, followed by the addition of the notched noise. After it was clear that the participants understood the behavioral task and were familiarized with all the stimuli they were supposed to attend to or ignore in the actual experiment, the next step was to complete a set of 10 training runs for each behavioral condition before beginning the experiment. Training runs utilized the same stimuli as in the actual experiment, except modulation never occurred at the distractor frequency and no MOC-reflex eliciting noise was presented. If task performance during these training runs was too poor, the acoustic stimulus was reduced in complexity (in reverse of the procedure described above) until performance was above 70% correct. This percent correct means that participants reported a correct number of AM tones on at least 7/10 runs (the actual percent of correct AM detections was likely higher). Participants typically completed the training in the first 10 runs and training lasted 30-60 minutes. The training and screening procedures were typically completed during one two-hour experimental session.

In a second experimental session after completing the training, participants completed runs of experimental trials under three different attentional manipulations. Participants were required to attend either to the sequence of low-frequency (hereafter

AL) or high-frequency (AH) tone pips, or to attend to the sequence of visually presented letters (V) on the computer screen and to perform the relevant behavioral task for the attended stimulus (i.e. counting AM tones in the AL and AH conditions, and counting consecutive letter repetitions in the V condition). Responses were made at the end of the run to ensure as little movement as possible during SFOAE recording. Ten experimental runs were recorded for each condition, providing a total of 100 presentations of the SFOAE stimuli for analysis. Experimental runs were grouped by attention condition, and the order in which attention conditions were presented was counterbalanced across participants to attempt to control for any effects that may slowly build up over the course of the experiment. Participants completed all three attention conditions in a single session, during which the probe assembly was not removed from their ears.

F. SFOAE analysis

Before emission extraction, audio recordings were manually scanned for noise and movement-related artifacts. Recorded runs containing visible artifacts were excluded from analysis. Additional recordings were made if the magnitude of SFOAEs was less than 20 dB above the noise floor. On average 7 recordings (containing approximately 70 probe-suppressor pairs for the baseline measurement) per attention condition per participant were used in the following analyses. The numbers of artifact-rejected recordings were very similar across the different attention conditions.

To estimate the SFOAE magnitude evoked by the task-relevant stimuli, the SFOAE from the baseline measurement was extracted utilizing the suppression technique (Shera and Guinan, 1999). First, the 1-s segments of SFOAE baseline measurement stimuli consisting of repetitions of 0.5-s SFOAE probe and 0.5-s SFOAE probe plus suppressor were averaged together. Next, a heterodyne procedure was used to extract the complex-valued sound pressure in the ear canal at the stimulus frequency (Guinan et al., 2003). A phasor diagram depicting the vector subtraction resulting in the extracted emission is shown in Fig. 2.2. The magnitude and phase of this sound pressure from an example listener is shown in panels A and C of Fig. 2.3. The suppression method of
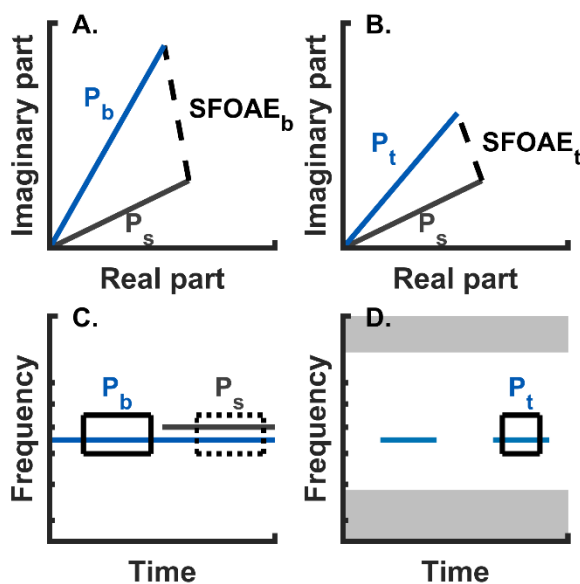

Figure 2.2.

Phasor diagrams and time-frequency schematic illustrations depicting the emission extraction procedure from the averaged temporal waveform of recorded audio. A). A phasor diagram depicting the extraction of baseline SFOAE as the resultant of vector subtraction of baseline pressure Pb and suppressed baseline Ps. B). Phasor diagram depicting the extraction of SFOAE from the task relevant stimuli as the vector difference between pressure during the behavioral task Pt and Ps. Panels C and D are time-frequency schematics of the stimuli used to extract SFOAEs depicted in panels A and B, respectively. Shaded grey regions represent the notched noise elicitor. Dashed boxes highlight analyses windows where averaging occurs to perform the vector subtraction.

SFOAE extraction uses a vector difference between ear-canal sound pressure recorded with the SFOAE probe tone alone and the sound pressure recorded with the SFOAE probe tone and suppressor to estimate the SFOAE residual. The resultant is termed the residual since the suppressor tone may not be completely effective in suppressing the SFOAE. The mean complex-valued sound pressure during a 300-ms window temporally centered in the duration of the suppressor tone was subtracted from each point of the heterodyned waveform. This analysis yields the magnitude and phase of the SFOAE residual during the first 500 ms of the waveform when no suppressor was present and the noise floor during the final 500 ms, as shown in panels B and D of Fig. 2.3. The same mean complex-valued sound pressure was also subtracted from the heterodyned average sound pressure of the low-frequency tone pips used during the behavioral task. Ten tone pips (one per behavioral stimulus block) were averaged together across blocks and runs to yield a single average tone pip. Tone pips that contained amplitude modulation (behavioral targets) were always excluded from the averaging procedure. If neither of the two tones within a block contained amplitude modulation, the first tone was always chosen. The level of SFOAE residual obtained from the tone pips during a 200-ms window in the temporal center of the tone-pip is shown in panel E of Fig. 2.3 (in red) with the level of SFOAE residual during the baseline measurement (in grey). Predictably, the level of the residual during the task is lower than the baseline, due to MOC efferent activation by the notched noise presented with the test tones but it is well above the noise floor (compare Fig. 2.3 B and E), leaving room for attention-related changes in MOC efferent effect on SFOAE magnitude.
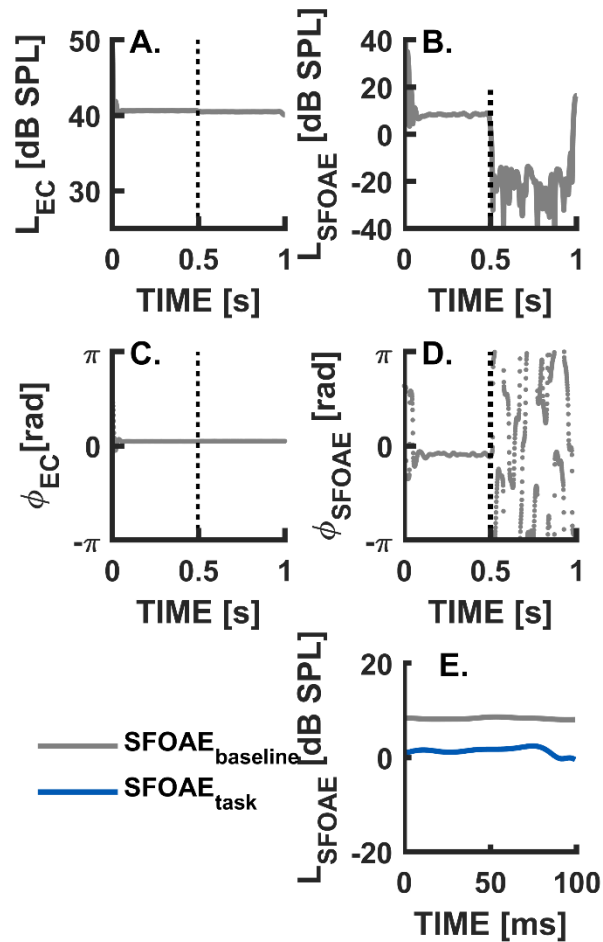
Figure 2.3.

Example extracted SFOAE data from a single listener (S12) during the attend low experimental condition. Dashed lines in panels A-D depict the onset time of the tone used to suppress the emission. Panels A and B show extracted levels of ear-canal sound pressure at the stimulus frequency and SFOAE, respectively. Panels C and D depict phase of the same signals. Panel E shows the level of SFOAE during the baseline (grey) and behavioral (blue) task segments during the 200-ms window used to compute average magnitudes for each listener and condition.

G. Bootstrap analysis procedure

In order to determine whether differences in emission level during the behavioral tasks across attentional conditions were significant within an individual, a bootstrapping procedure was used to estimate the variability in emission magnitudes. Artifact-free pairs of recording segments for both the baseline and task measurements were randomly sampled with replacement, such that the total number of resampled segments was always

equal to the total number of original segments. The SFOAE extraction procedure detailed

above was repeated on each resampling and the mean emission magnitudes were saved.

The resampling and analysis procedure was repeated 10,000 times per participant to

estimate distributions of emission magnitudes across the attentional conditions.

III. Results

A. Behavioral results

Since participants responded with counts of detected targets at the end of a run

instead of at the moment of target detection it is not possible to perform many detailed

analyses on the behavioral responses. Although this limits our ability to correlate our

behavioral and OAE metrics we still provide evidence that our listeners could complete

the behavioral tasks with reasonable accuracy and that our two samples did not differ in

their performance of the task.

1. Original group

Participants responded with the exact count of targets in 51.7% of runs across all

three conditions. Because responses in single runs were aggregate counts of the number

of detected modulations over 10 stimulus blocks, this metric is not the same as the typical

percentage of correct detections as even a single miss would cause a block of nine other

successful trials to be counted as incorrect. To provide a secondary account of accuracy

in the behavioral task, we calculated the participants' root mean square error (RMSE)

from the correct number of presented targets. This measure better compares counts of

responses to targets across runs. On average participants' RMSE was 0.84 targets across

all three attention conditions. Based on this we are confident that participants were able to attend the required stimuli, considering that a total of 20 (10 at each frequency) tone pips are presented within a single run and up to 10 (5 at each frequency) could be modulated and listeners could respond with any number. A repeated-measures analysis of variance (ANOVA) revealed no significant effect of task on behavioral performance after applying a Greenhouse-Geisser correction for a violation of sphericity [$F(1.32,18.8) = 3.81$, $p = .056$]. The nearly significant effect of task on performance was driven by lower performance in the AL (mean RMSE: 1.21) condition relative to AH (mean:0.79) and AV (mean: 0.52). Because modulation depth was set to make AM detection challenging (recall that not all instances of AM at the cued frequency were detected for the stimuli without the distractor AM and the notched noise during the training), the less-than-perfect performance does not necessarily indicate slips in attention to the cued frequency in the auditory tasks.

2. Replication group

The second (replication) group of participants responded with correct target counts in 52% of runs. Their RMSE was 0.81 across all three conditions. A repeated-measures ANOVA using just the participants from the second group revealed no significant effect of task on performance [$F(1.46,20.4) = 2.97$, $p = 0.09$]. While the effect of task on performance was not significant, there was a trend for lower performance in the AL (mean: 1.08) relative to the AH and V conditions (means: .67 and .68, respectively), as found in the original group.

3. Group comparison

A mixed repeated-measures ANOVA including a between-subjects factor of experimental group showed no significant effect of experimental group on performance [$F(1,28) = .049$, $p = .827$]. There also was no significant interaction between experimental groups and task [$F(1.45,40.5) = .536$, $p = .532$]. There was a significant effect of task on performance in the pooled results from the two groups [$F(1.45,40.5) = 6.55$, $p = .008$], based on the lower performance in AL condition than in the AH and V conditions.

B. SFOAE results

1. Original group

First, a within-subjects analysis was undertaken, using the bootstrap procedure described above. The bootstrap distributions were used to estimate 95% confidence intervals around the mean emission magnitudes; selective-attention effects within a participant were considered significant if the median emission magnitude in the AH or V conditions (when attention was directed away from the SFOAE-evoking tone) fell outside the 95% confidence interval for the magnitude of emission in the AL condition (when attention was directed towards the SFOAE-evoking tone). Nine out of 15 participants demonstrated a significant change in emission level between the AL and V conditions, all with the emission levels lower (indicating more efferent inhibition of cochlear responses to the low-frequency tone) in the V condition than in the AL condition. Seven participants had significantly different emission levels in the AH condition compared to the V condition with the AH condition having higher emission levels in 6/7 cases. The

two auditory conditions (AL and AH) had significantly different emission levels in 9/15

cases with the AH condition showing greater efferent inhibition than the AL condition in
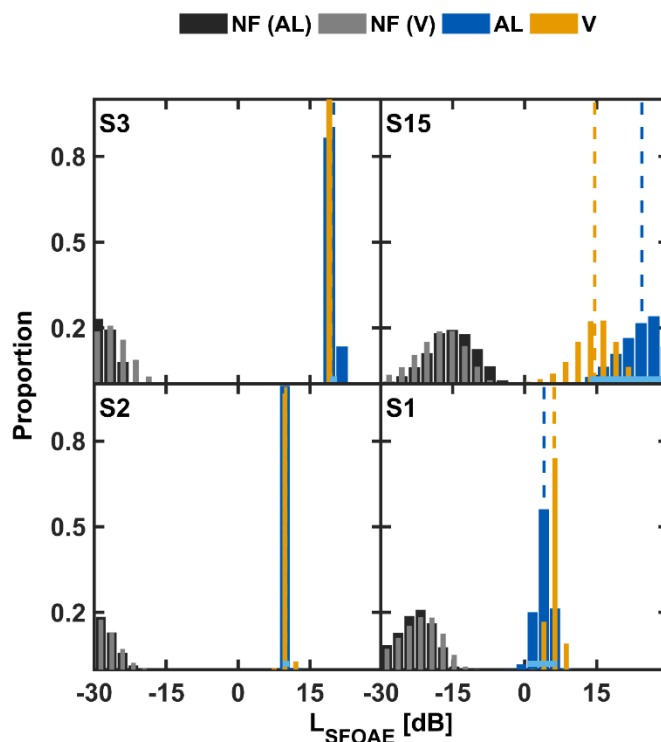
5 of those cases.

Figure 2.4.

Results from the bootstrapping procedure used to quantify measurement variability within participants. Histogram bars depict the mean level of the noise floor NF and SFOAE in both the AL and V conditions. Top panels show distributions for representative participants from the original sample, bottom panels show distributions for participants from the replication sample. Left panels show examples of data where no significant change existed across attention conditions, while the right panels are typical of participants who displayed an individually significant difference between the AL and V conditions. Vertical dashed lines show the mean of each SFOAE distribution. Light blue horizontal bars under the AL distributions indicate a 95% confidence interval around the mean of the distribution.

The top panels in Fig. 2.4 show results of the bootstrapping procedure from a

representative participant with a significant effect and a representative participant with no

significant effect. Each panel contains histogram distributions of SFOAE and noise floor

magnitude in the AL and V conditions. Six participants did not show significant

differences in emission magnitude across conditions, but the results often exhibited the

same trend of lower emission levels during the V than during the AL condition. We

emphasize the AL and V conditions here for plot clarity since we observed the largest

differences between these two conditions. Note that in all cases shown in Fig. 2.4 there is

an almost complete overlap of the distributions representing the noise floor, indicating
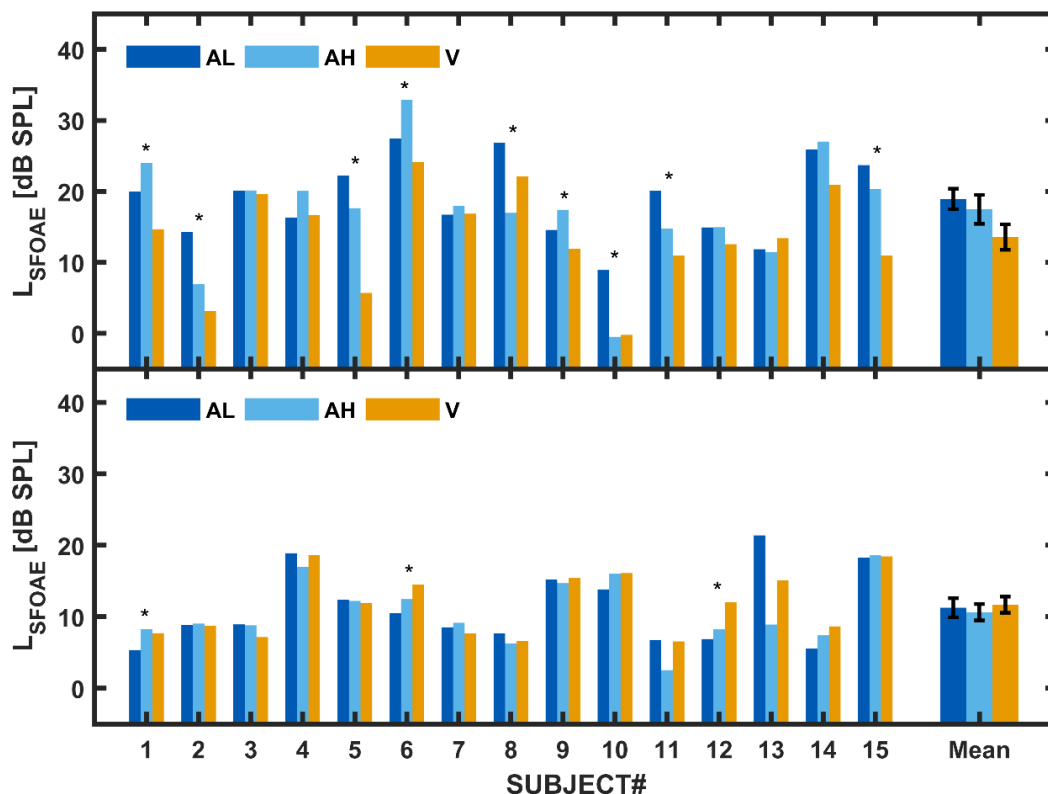

Figure 2.5.

Individual SFOAE levels across attention conditions obtained from the standard heterodyne procedure (not the bootstrap procedure). Asterisks denote participants for whom the bootstrap procedure found significant differences in SFOAE level between the AL and V conditions. Mean values across participants with standard error are shown by the rightmost set of bars.

that a change in the level of the noise floor cannot account for the differences between the

conditions.

Next, effects of condition were evaluated across participants. The individual

SFOAE magnitudes for each attention condition as well as the mean across participants

are shown in the top panel of Fig. 2.5. The patterns of results depicted in the bootstrap

analysis can be seen across attention conditions. Participants for whom the bootstrap

procedure found significant differences in emission between AL and V are denoted with

asterisks above the bars. As shown by the rightmost set of bars, the SFOAE magnitude

was on average smallest for the V condition, highest in the AL condition, and

intermediate in AH condition.

A Bayesian analysis procedure (Kruschke, 2013) was used to determine whether

differences in emission magnitudes between attention conditions were credibly different

from zero. Briefly, SFOAE magnitudes from the task stimuli were subtracted across pairs

of conditions (e.g. AL – V) for each participant. These differences were modeled using a
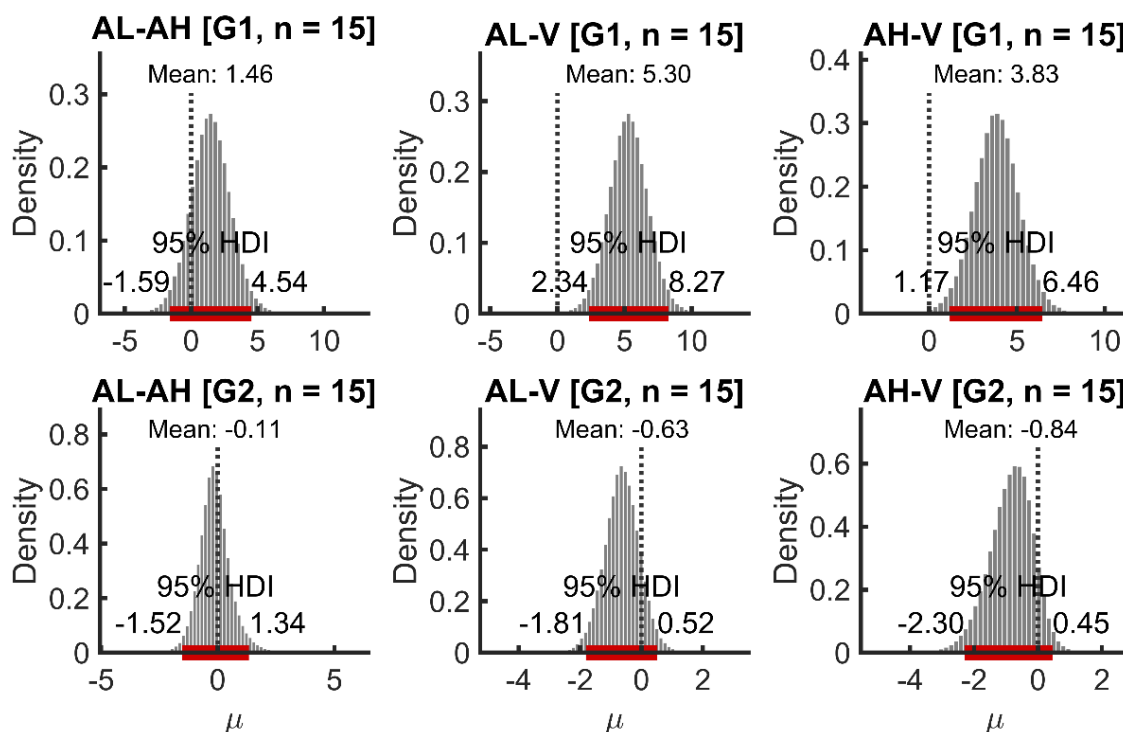
Figure 2.6.

Posterior distributions of the mean difference in OAE magnitude between pairs of attention conditions from the Bayesian analysis of group data. The shaded horizontal bar on the x-axis denotes the 95% highest density interval (HDI) used to assess the credibility of the differences in each pair. Vertical dashed lines at x = 0 help to visualize how each distribution relates to a 0-dB effect. If the bar illustrating the 95% HDI does not overlap with the dashed line the effect can be considered credibly different than zero.

t-distribution and posterior distributions of the mean, standard deviation, and degrees of

freedom of the t-distribution were simulated using custom MATLAB software to perform

the Markov-Chain Monte Carlo simulation via Gibbs sampling. The 95% highest density

interval (HDI) of the posterior distribution of the mean difference between each pair of

attention conditions was compared to 0 to determine the credibility of the difference in

emission magnitudes. The analysis found a credible difference in emission magnitudes

between AL and V (mean = 5.3 dB) conditions as well as a credible difference between

AH and V (mean = 3.84 dB) conditions, but not between AL and AH conditions. The

posterior distributions of the mean difference in SFOAE magnitude for each pair of

conditions and the 95% HDI are shown in the top panels of Fig. 2.6. A repeated-measures

ANOVA also confirmed a significant effect of attention condition on emission magnitude

$[F(2,28) = 9.37, p = .001]$, and paired-samples t-tests confirmed the same significant

pairwise differences between the V condition and both the AL $[t(14) = 4.160, p = 0.001]$

and the AH $[t(14) = 3.199, p = .006)]$ conditions.

2. Replication group

The results of the SFOAE analyses for the same within- and between-subjects

comparisons as described in the previous section are shown in the bottom halves of each

of the Figs. 2.4-2.6. As shown in the figures (e.g., top vs. bottom panel of Fig. 2.5), there

was a difference in SFOAE levels between the original and replication groups. However,

the SFOAE in the replication group remained sufficiently high to detect potential

attentional modulation. For instance, S2 and S10 in the original group show large

differences in emission level between AL and V but also have considerably weaker

emissions than S4 and S15 in the replication group. This issue is addressed further in sections III.F and IV.B.2.

The bootstrap analysis revealed significant changes in emission level between the AL and V conditions in only three out of the 15 participants in the replication group. In contrast to the original group, all three participants showed small increases in emission levels in the V condition relative to the AL condition. Bootstrapping did not reveal any individually significant differences in emission levels between the AH and V nor the AL and AH conditions. The Bayesian analysis revealed that there were no credible differences in emission level between any of the three attention conditions across participants. This was also confirmed by a repeated-measures ANOVA [$F(2,28) = .932$, $p = .408$].

3. Group comparison

A mixed ANOVA was used to compare the effects of the within-subjects factor of attention condition and between-subjects factor of group to examine whether attention effects were similar between the original and replication groups. The ANOVA revealed a significant main effect of condition in the combined sample [$F(2,56) = 5.492$, $p = .007$]. This effect was driven primarily by the difference in SFOAE magnitude between the AL and V conditions (mean AL-V = 2.48 dB). There was also a significant effect of group on SFOAE magnitudes [$F(1,28) = 6.69$, $p = 0.010$]. Participants in the replication group had smaller SFOAE magnitudes (mean difference between the two groups: 5.46 dB). There was a significant interaction between experimental group and attention condition [$F(2,56) = 8.67$, $p = 0.001$]. The interaction confirms that the effect of attention differed between

groups (i.e., there was a main effect of attention condition in the original, but not the replication, group).

C. Relationship between behavioral and SFOAE data

To determine whether the strength of attentional modulation of the SFOAE was related to behavioral task performance, the difference in SFOAE magnitude between each auditory condition and the visual condition (AL-V and AH-V) was compared to performance in each condition, in terms of the RMS error of responses to the correct number of targets for each condition. We focus on the difference between these two pairs of conditions as they produced significant differences in SFOAE magnitude in the original experimental group.

There was no significant correlation between SFOAE AL-V magnitude difference and the behavioral performance in the AL and V tasks [$r(28) = 0.08$, $p = 0.68$ for the auditory task, and $r(28) = -0.23$, $p = 0.23$ for the visual task]. There were no significant correlations between performance in any condition and AL-V difference in SFOAE magnitude in either subgroup, although there was a trend towards better performance in the visual task being related to larger differences between SFOAEs during V vs. AL attentional conditions in the original sample [$r(13) = -0.74$, $p = 0.08$]. There were also no significant correlations between SFOAE AH-V magnitude difference and the behavioral performance in the AH [$r(28) = -0.15$, $p = 0.428$] and V tasks [$r(28) = -0.26$, $p = 0.165$]. These correlations were also not significant within each subgroup.

D. Noise floor and baseline emission measurements

To ensure that changes in emission were related to differences evoked by task demands, we recorded the level of the baseline SFOAE as well as the noise floor of the measurements from a time when the suppressor tone was present. These metrics should remain constant across attention conditions and changes in either baseline SFOAE or noise floor measurements could indicate contamination of our selective attention effect by artifacts. A mixed repeated-measures ANOVA confirmed that there was no significant effect of attention condition on the baseline SFOAE magnitudes [$F(2,43.6) = .621$, $p = 0.50$] nor the noise floor [$F(2,56) = 0.29$, $p = 0.75$] across both experimental groups. There were no interactions between the selective attention effect and experimental group on these measures, ([$F(1.56,43.55) = 2.76$, $p = .072$] and [$F(2,56) = 0.266$, $p = 0.767$] for baseline SFOAE magnitudes and noise floor respectively) but there was a group difference in baseline SFOAE magnitudes [$F(1,28) = 8.57$, $p = 0.007$].

E. Changes in emission magnitude over time

Each participant completed the three experimental conditions in a randomized order. Since all conditions were run within a single session without removing the probe, slow systematic variations in SFOAE magnitude due to drifts, for example, could confound selective attention effects. To address this possibility, we investigated changes in the baseline SFOAE measurement across condition order, rather than attention condition, to test for potential slow changes in emission magnitude over time spanning all attention conditions. A repeated-measures ANOVA with a between-subjects factor of experimental group showed that measurement order (or the time elapsed since the start of

data collection) did not significantly influence the noise floor, baseline SFOAE, or task

SFOAE [across all three tests, $F(2,56) < 2.75$, $p > 0.07$]. There were no significant

interactions between group and order [$F(2,56) < 2.25$, $p > 0.12$]. Mean magnitudes across

participants are shown in panel A of Fig. 2.7. We also examined stimulus pressure in the

ear canal within each condition in the participants who showed the largest AL-V

differences and found no readily apparent systematic change in ear-canal sound pressure

over time. Panel B of Fig. 2.7 shows ear-canal sound pressure level as a function of run

number for a representative participant for each of the three attention conditions (S1 from

the original group).
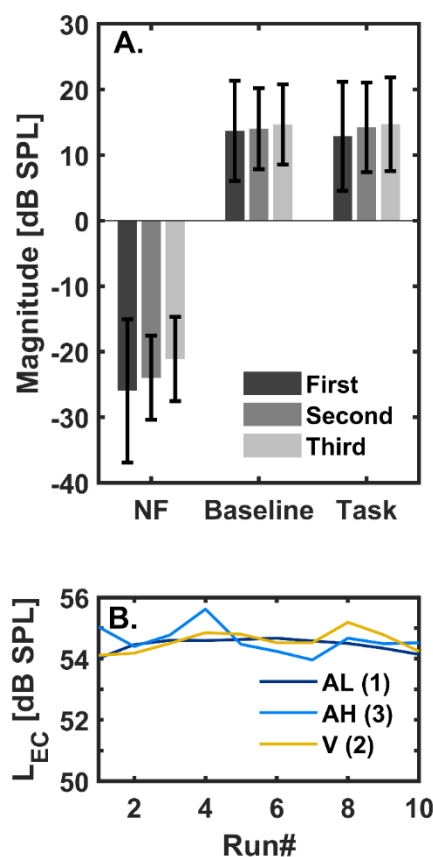
Figure 2.7.

Effects of time on emission measures. Panel A shows bar plots of the average magnitude of the noise floor, baseline SFOAE, and task SFOAE, for the first, second, and third measurements respectively. Error bars denote standard deviation. Panel B plots the stimulus ear canal pressure in a representative participant (S1, original group) as a function of run number and attention condition. Numbers next to the attention condition labels in the legend denote the order in which conditions were run.

F. Group Differences

Due to the large difference in the observed selective attention effect between the two groups, we compared demographic data collected from the groups during intake. The groups differed slightly in age, with the original group being slightly older than the replication group (difference in median age = 3 years). Age was not significantly correlated with the AL-V SFOAE magnitude difference. The groups also differed slightly in racial makeup, with the original group containing a larger proportion of participants who self-identified as Asian/Asian-Pacific Islander (hereafter: Asian) on our participant intake questionnaire (11/15 in the original sample vs. 6/15 in the replication). When grouping all of the participants by race, those who identified as Asian did have larger differences in SFOAE between the AL and V conditions than those who identified with other racial descriptors (3.55 dB and 1.07 dB respectively), but this difference between groups was not statistically significant [$t(28) = -1.34$, $p = 0.19$]. There were no significant differences in SNR for baseline SFOAEs between Asian and non-Asian Participants [$t(28) = -0.92$, $p = .37$].

Although handedness has been shown to influence the strength of MOC function (Khalfa et al., 1998), it was not assessed in our participants. It is possible that a large number of left handed individuals would possess weaker MOC effects in their right ears (potentially leading to a group difference), but it seems unlikely that either group would contain a large number of left handed participants given the relatively low incidence of left handedness in the general population (Gilbert and Wysocki, 1992).

As mentioned in the Methods, participants in the original and replication groups were tested using different instruments (ER10C vs. ER10X). A consequence of using

different instrumentation was that the groups differed in the measured stimulus-frequency

ear-canal pressure (mean: original-replication = 7.49 dB). Consequently, the original

group had significantly larger SFOAEs (both baseline measurement and during the task)

than those in the replication group (mean: original-replication = 5.62 dB) and larger

SNRs (mean: original-replication = 7.56 dB). Larger SFOAE levels and SNRs may have

increased the chance of observing selective attention effects and thus may have

contributed to the different outcomes for the two groups. To examine this possibility, the

magnitude of differences in SFOAE levels between the two auditory conditions and the

visual condition (AL-V and AH-V) were correlated with the measured ear canal sound

pressure level at the stimulus frequency and the result is shown for the participants

pooled across the two groups in Panel A of Fig. 2.8. If the lower level stimuli limit the

ability to see MOC effects, then the differences in SFOAE between attention conditions

should be smaller at lower sound levels. There was no significant correlation between ear

canal pressure and the size of the difference in SFOAE between conditions for neither

AL-V (Fig. 2.8A, all dark blue symbols) [$r(28) = .255$, $p = .173$] nor AH-V (all light blue

symbols) [$r(28) = .325$, $p = .08$]. Within each group, there was also no significant

relationship between ear canal sound pressure and AL-V, but the data trended towards a

negative relationship. In the original group there was no correlation between ear canal

pressure and AH-V (Fig. 2.8A, light blue open filled) [$r(13) = -0.119$, $p = .672$]. There

was a significant correlation between ear canal pressure and AH-V (Fig. 8A, light blue

open) [$r(13) = -0.825$, $p = .0002$] in the replication group, but this would suggest that

smaller effects were seen with *increasing* ear canal pressure, which is opposite to the

overall trend in the data and would not explain why greater attention effects were seen in the original group.

Although the noise floor did not differ between the two groups, the difference in stimulus pressure in the ear led to a difference in SNR for the SFOAE measurements between groups, as confirmed by an ANOVA [$F(1,28) = 7.77, p = 0.01$]. The ANOVA results did not indicate that SNR varied across conditions [$F(2,56) = 0.17, p = 0.84$]. There was no interaction between SNR and experimental group [$F(2,56) = 0.04, p = 0.96$]. A scatterplot of the individual SNRs for the and the corresponding selective attention effects are shown in Panel B of Fig. 2.8. There were no significant correlations between the SNR and either of the selective attention effects in the pooled [$r(28) < 0.23$, $p > 0.228$] nor subgroup data [$-0.33 < r(13) < 0.034, p > 0.229$].

Figure 2.8.
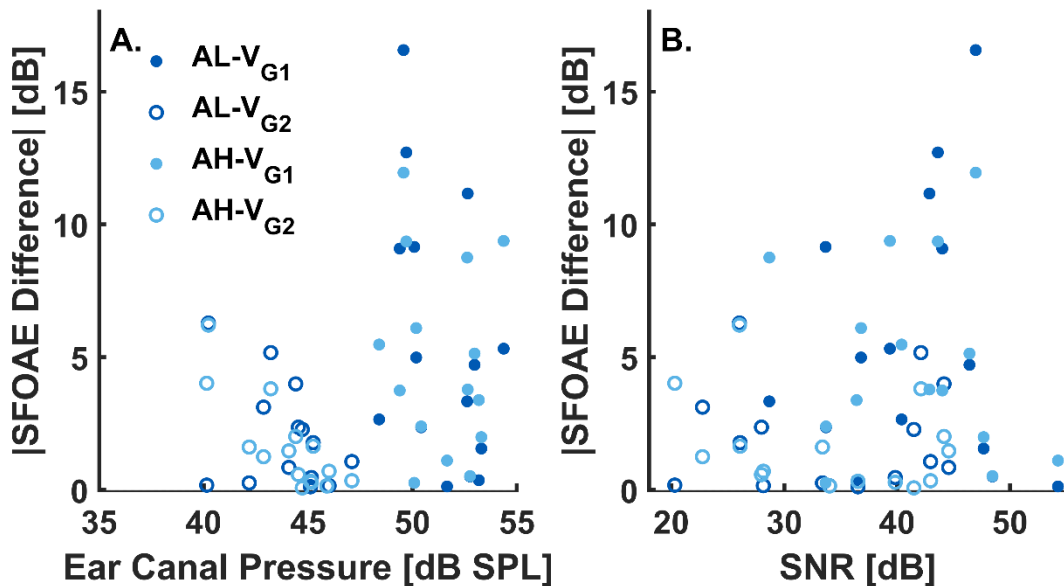
Panel A shows relationships between measured stimulus-frequency ear-canal pressure and the magnitude of the attention effect. Panel B relates signal-to-noise ratio to the attention effect. Dark blue symbols show the AL-V effect, while light blue show the AH-V effect. Closed symbols denote participants in the original group, while open symbols denote participants in the replication group.

IV. Discussion

A. Summary of findings

The primary finding of this study is that cochlear gain is significantly reduced when participants attend to visual stimuli compared with when they attend to auditory stimuli, as indicated by the difference in SFOAE magnitudes between the AL and V attention conditions. It appears that the pattern in the overall data is driven by a subset of individuals (40%) who show large (~5 dB) differences in OAE magnitude between the attention conditions. The results suggest that, for these participants, changes in cochlear gain may play an important role in modulating sensory input based on attentional factors.

Interestingly, no significant difference in SFOAE magnitude was observed between conditions where attention was directed at an audio frequency remote from the SFOAE (i.e. there was no difference in emission magnitude between the AL and AH conditions). The lack of difference in emission magnitude between the two auditory conditions contrasts with the frequency specific attentional effect observed by Maison et al. (2001) and the clear evidence for cortical amplification of attended vs. unattended auditory targets (e.g., Gutschalk et al., 2008; Mesgarani and Chang, 2012). Frequency-specific effects might have been expected, based on the demonstration of frequency tuning of MOC efferent effects (Lilaonitkul and Guinan, 2012), but there also exists evidence that attention can produce broadband effects (Walsh et al., 2014a, 2014b, 2015). These broadband effects were noted in the noise within the ear canal opposite the ear where OAE evoking stimuli were presented (Walsh et al., 2014a, 2014b) or by shifting the OAE analysis filter off the SFOAE frequency to examine changes in energy not evoked by the stimulus tone. Based on a recent report by Francis et al. (2018), it is

unclear whether these broadband effects have cochlear origin and if they reflect MOC activity.

It remains to be determined whether the lack of specificity within the auditory domain is an inherent property of the attentional MOC system or whether it reflects a property of the specific task we used in this study.

B. Failure to replicate strong findings between groups

1. Group differences

Perhaps the most striking finding from this study is that the results differ strongly between two groups of participants. In the original group, more than half (9/15) of the participants showed individually significant changes in SFOAE magnitudes across attention conditions compared to only 3/15 in the replication. The groups did not differ in their performance on the task but were not matched during recruiting, so differences between groups could help explain the differences in effect.

Demographic data indicated a difference in the racial makeup of each group, with the original experimental group containing a significantly higher number of participants who identified as Asian. This difference could be important, considering some work demonstrating racial differences in OAE magnitude (McFadden et al., 2012, 2018a; Shahnaz, 2008; Whitehead et al., 1993). If racial groups differ in SFOAE magnitude then it is possible that the SNR may also differ across racial groups, leading to differential ability to detect MOC effects on SFOAEs. We found in our sample that Asians tended to have a larger difference in SFOAE magnitude between the AL and V conditions than non-Asians, and Asians had slightly higher SNR than non-Asians. These differences were

not statistically significant and therefore unlikely to be responsible for driving the differences observed between the original and replication groups.

2. Effect of ear canal pressure

There was a significant difference in overall SFOAE magnitude between the original and replication groups. We traced this difference to differing sound pressure levels at the stimulus frequency in the ear canal while the stimulus was being delivered. Although the task, stimuli, and procedures were identical between the two groups, the equipment used for measuring OAEs was different. The original group of participants had stimuli delivered and recorded using the ER10C acoustic assembly, while the replication group was tested using the ER10X system. While these systems were both calibrated using the same equipment, the disposable ear tips used with each system were different and may have contributed to differing sound pressure levels in the ear canal.

To determine what effect the difference in ear canal sound pressure may have had on the overall findings we examined the relationship between the ear canal sound pressure at the stimulus frequency and the effect of attention condition. Our correlation analysis showed no significant relationship between stimulus pressure in the ear canal and the condition effect. The lack of correlation was not due to a lack of variability in measured ear canal sound pressures in each group (see Fig. 2.8). We also investigated relationships between the ear canal sound pressure and other factors that are important to robust measures of MOC effects on SFOAEs. For a constant noise floor, an increase in stimulus level will result in an increase in SFOAE magnitude, and thus an increase in SNR for SFOAE measurements. Significant differences in SNR could lead to a

differential ability to detect MOC-induced changes in SFOAE (Goodman et al., 2013) or attentional modulations thereof. Even though SNR was significantly different between the two groups, the lack of effect seen in the replication group is unlikely to be due to a change in SNR, because there was no significant correlation between SNR and the selective attention effect (see Fig. 2.8). The lack of correlation was likely due to the fact that all of our participants had high SNR (> 20 dB) for SFOAE measurements, as necessitated by our screening procedure, and thus attentional modulation of efferent effects, if present, should be reliably detected in all participants.

Despite the observed differences between groups in demographic and experimental data, we do not have compelling evidence that any of the factors discussed above can explain the difference between the groups in overall outcome. Below, we will focus on the combined result for all 30 participants and treat them as one (larger) sample.

C. Comparison to previous studies

1. Methodological Improvements

Although some evidence for peripheral auditory effects of selective attention have been reported in an animal study, based on recordings of the compound action potential and cochlear microphonic (Delano et al., 2007), the many previous human studies of attentional effects on peripheral auditory responses have generally failed to find robust and consistent effects of attention on cochlear responses via OAEs (Giard et al., 1994; Maison et al., 2001; Michie et al., 1996; Srinivasan et al., 2012, 2014; Walsh et al., 2015). We designed the current study with several methodological manipulations with the purpose of improving our ability to detect an effect of attention on the OAE signal. The

main methodological improvements included the use of SFOAEs at low stimulus levels to avoid activation of the MOCR by the test stimulus. This is important because activating the MOCR by the test stimuli may introduce interactions between two acoustically elicited efferent effects and possible attentional effects that would make the interpretation of results difficult. This is because changes in SFOAE phase due to MOCR elicited by different acoustic stimuli can be different, causing complex phase interactions that could affect the resulting SFOAE (Lilaonitkul and Guinan, 2012). An additional improvement is the use of attentional tasks in all conditions, rather than the more common "passive" comparison, for which attention cannot be properly controlled. The use of a low-level notched-noise elicitor also helps to avoid inadvertent activation of the middle ear muscle reflex (MEMR). MEMR activation produces changes in middle ear impedance that would contaminate the SFOAE measurement due to an increased or decreased energy in ear-canal sound pressure at the probe frequency that could be wrongly attributed to the MOC efferent effect of interest in this study. MOCR measurements are only valid under the assumption that the middle-ear impedance remains the same in the presence and absence of the MOCR elicitor. Moreover, efferent projections from higher-level auditory nuclei to the MEMR neural circuit make it possible for attention to affect the MEMR strength, and thus modulate the ear-canal sound level at the OAE-evoking frequency (Mukerji et al., 2010).

Previous studies used high-level, fast-rate clicks to evoke OAEs (Ferber-Viart et al., 1995; Harkrider and Bowers, 2009); these have been shown to risk activation of the MOCR by probe stimuli themselves (Guinan et al., 2003), potentially reducing attention-related effects. Attentional changes in distortion-product OAEs (DPOAEs; Smith et al.,

2012; Srinivasan et al., 2012, 2014; Wittekindt et al., 2014) are difficult to interpret, due to phase interactions between DPOAE components that can result in both increases and decreases in DPOAE magnitude. We also used the suppression method for estimating SFOAE, rather than a non-linear stimulus subtraction method (e.g. nSFOAE, Walsh et al., 2015) to preserve changes that could occur linearly. In addition, our low-level (55 dB SPL) spectrally notched noise reduced the possibility of inadvertent activation of the MEMR, and our low probe level of 35 dB SPL ensured that the probe itself did not elicit measurable MOCR activity.

Most previous studies using OAEs to probe cochlear function under auditory attention have presented OAE eliciting stimuli as task-irrelevant stimuli (Avan and Bonfils, 1992; de Boer and Thornton, 2007; Ferber-Viart et al., 1995; Maison et al., 2001; Walsh et al., 2014a, 2014b, 2015). This could lead to instances where participants tried to ignore the OAE stimuli, since they could be distracting in the context of the actual task. The current study utilized SFOAE stimuli in one of the behavioral tasks to compare the cochlear function at the attended frequency to that when attention is directed to the task focused on a different frequency region. This was done to maximize the size of the attentional effect, but the results did not reveal a frequency-specific effect of attention. This finding is consistent with other studies that required selective attention to DPOAE primaries (e.g. Srinivasan et al., 2012, 2014; Wittekindt et al., 2014). These studies demonstrated cross-modal effects of attention, even though the frequency of DPOAE was different from the attentional target in the auditory task. Although there is anatomical evidence from animals (Brown, 1989; Liberman and Brown, 1986) and OAE evidence from humans (Lilaonitkul and Guinan, 2012; Maison et al., 1997) supporting tuning of

MOC effects, the current study was unable to demonstrate that shifts in attention produce frequency-specific effects. There was a trend for lower emission magnitudes in the AH condition relative to the AL condition in the original group (see the mean bars in panel A of Figure 5), but like the AL-V difference, this effect was not reproduced in our second group of participants.

The frequency specificity of MOC effects in previous reports were attributed to changes in the stimulus properties of the MOC elicitor (Lilaonitkul and Guinan, 2012; Maison et al., 1997) which were held constant in the current study. This suggests that the attentional modulation of MOC activity behaves differently than the acoustic activation of the MOC system, as attention was unable to produce an observable frequency-specific effect in this study.

Several previous reports have compared a state of attention, where participants were actively involved in some perceptual task, to a passive state of inattention (Avan and Bonfils, 1992; Puel et al., 1988; Walsh et al., 2014a, 2014b, 2015). These comparisons present two primary inherent difficulties. First, it is not possible to interpret the results found in these studies as strictly due to selective attention, as general arousal likely differs between an active (attention) and a passive (inattention) condition. Second, while task performance can provide some measure of how attentive participants are while completing the behavioral task, inattention does not have a corresponding behavioral metric. This means that participants could continue to attend to aspects of their environment (including OAE-evoking stimuli) during the passive task, and so any comparisons made to this condition would suffer due to the variability in attentional state across participants.

Although the methodological changes made in the current study eliminated a few potential sources of confounding factors due to the experimental design, the new design was not able to overcome the significant inter-subject variability in selective attention effects on OAE magnitude. Furthermore, this experimental design did not allow us to elucidate what is different about the participants who showed significant differences in SFOAEs between V and AL conditions. Some limitations of our design are discussed in the following section.

2. Limitations of the current design

Our current design is somewhat limited in its ability to relate behavioral performance to changes in SFOAE due to our choice of response method. While responding at the end of a run should reduce motion-related artifacts during the recording, it also makes it impossible to do more than a very coarse comparison between OAE magnitudes and performance. A future improvement could be to incorporate responses to targets in real time.

A second limitation comes from the fact that we used short gated stimuli to measure SFOAEs. Unlike the continuous probe tones used in most previous studies (Backus and Guinan, 2006; Lilaonitkul and Guinan, 2012, 2009), short tones make it difficult to observe slow changes due to factors such as drift in the position of the probe assembly in the ear canal. We looked for drift across conditions by comparing both baseline SFOAE measurements and noise floor across the three attention conditions and found no significant differences in either measure. We also looked for trends in stimulus pressure in the ear canal across each of the runs in each participant to see if there were

any systematic changes. While there was some variability across the runs, we noted no systematic change in stimulus pressure in the ear canal that could produce the observed selective attention effects. This is also supported by a lack of order as a significant factor influencing emission magnitudes.

The use of a low-frequency tone for eliciting SFOAE may have limited our ability to observe MOC effects by limiting the range of cochlear gain (and potential reduction) applied to the stimulus. In animal studies, innervation of the OHCs by MOC efferents has been shown to be greater at mid-frequencies than low or high frequencies (Brown, 2011; Liberman and Brown, 1986). Physiological measurements in animals (for review, see Recio-Spinoso and Oghalai, 2017; Robles and Ruggero, 2001) as well as indirect measurements of cochlear gain in humans also suggest that there is less gain at lower frequencies (Gorga et al., 2007; Plack et al., 2008). Future work utilizing the methodological improvements noted in the previous section could be further improved by measuring at less apical frequencies that may have a larger dynamic range for measuring MOC effects.

3. The importance of adequate sampling

As concerns over the replicability of psychological science are currently a topic of active discussion (e.g., Open Science Collaboration, 2015) a key contribution of the current work is to highlight the need for higher power in experimental designs seeking to use SFOAEs to investigate MOC efferent effects.

Several of the more recent studies of selective attention on cochlear function (e.g. Srinivasan et al., 2012, 2014; Walsh et al., 2015) used relatively small (n < 15) samples

of participants, which limited their ability to detect significant differences in emission

magnitudes that were due to attention. The current study used a simple power analysis to

estimate the required number of participants to observe an attention effect based on pilot

data, but a sample of this size (n = 15) failed to secure a reproducible finding. The

surprising difference in the outcome of the attention condition study between two

participant samples, despite the same design and stimuli, suggests that the reported

attentional effects on OAEs in humans may have been due to spurious effects in a

specific sample. It is also possible that the differences between samples simply reflect

large inter-subject variability in the magnitude of attentional effects, even in a group of

young normally hearing listeners. Individual variability in OAE magnitudes have been

explored by studies looking at biological factors like race and sex, as well as

psychophysical performance (McFadden et al., 2012, 2018a, 2018b), but future work is

necessary to explore the individual variability revealed in our study.

D. Conclusions

This study provides some evidence that cochlear responses to sound may change

under conditions of selective attention. Specifically, visual attention (relative to auditory

attention) is accompanied by a modest reduction (~ 2.5 dB) of SFOAE magnitude,

suggesting a reduction in cochlear gain of the unattended auditory stimulus. Although

statistically significant for the combined pool of participants, the effect of attention

condition was driven by a large significant effect in the first group of participants tested.

None of the considered differences between the two participant groups provides a

compelling explanation for the different outcomes. Recent studies of the attentional

effects utilize sample sizes that are likely too small to reliably detect attention effects, which may explain the lack of consistent results seen in the current literature. It is possible that the efferent system is capable of modulating cochlear responses by attention, but the OAE-based measures are inadequate to detect the attentional effects because their inherent variability swamps any systematic differences related to attention. It is also possible that even with the anatomical efferent structures in place, attentional control does not extend all the way to the cochlea and instead it arises at central sites in the auditory pathways. This possibility has support from a study showing no evidence of attentional modulation in auditory-brainstem steady-state responses (Varghese et al., 2015a).

Chapter 3: Examining the effects of perceptual load on cochlear function during selective attention.

I. Introduction

In chapter 2, we designed a study to incorporate advancements in measuring medial olivocochlear reflex (MOCR) effects with otoacoustic emissions (OAEs) and improvements to attentional manipulations to provide more compelling evidence of attentional modulation of cochlear sound processing (Beim et al., 2018). In that study participants attended to different aspects of an audiovisual stimulus to detect target stimuli in sequences of tones and visually presented letters. The auditory stimuli were simultaneously used to measure stimulus frequency OAEs (SFOAEs) and behavioral performance. While our previous study focused primarily on designing stimuli for the optimal measurement of OAEs and MOCR effects it is of equal importance to consider the literature surrounding selective attention to the task design that allows for optimal chances of detecting an attentional modulation of MOCR activity. However, even in such studies, differences in factors such as perceptual or cognitive load have rarely been considered. These factors may be important in determining the relationship between cochlear effects and selective attention, particularly with respect to whether selection occurs at an early or late stage of processing.

Early debates on the nature of selective attention yielded evidence for both early selection (i.e., selectively filtering out unattended stimuli at an early stage of perceptual processing) and late selection (i.e., filtering only after considerable processing of the unattended stimuli had occurred). Late attentional selection is supported by studies showing that unattended stimuli can influence behavioral responses in attentional tasks (e.g. Miller, 1987; Stroop, 1935), suggesting that unattended stimuli are processed, and

that attentional filtering occurs shortly before the behavioral response. In contrast, other

studies have provided evidence in favor of early selection theories of selective attention

by showing that certain aspects of unattended stimuli seem to be consciously inaccessible

(Broadbent, 1958; Treisman, 1969). The perceptual load theory of selective attention

(Lavie, 1995, 2005; Lavie and Tsal, 1994) proposes that the early and late selection

theories can be reconciled by considering the perceptual or cognitive load involved in the

specific tasks. According to the perceptual load theory, stimuli are processed obligatorily

if enough attentional resources exist, resulting in late selection. However, when

attentional resources are limited by a high perceptual load, selection operates earlier so

that more resources can be devoted to attending the target stimuli. Evidence of the effects

of perceptual load on selective attention can be found in many studies in vision. For

instance, increasing the perceptual load of a linguistic judgement task eliminated

signatures of motion processing in visual cortex produced by the motion of dots in the

participants peripheral vision compared to a condition where the linguistic judgment task

had a low perceptual load (Rees et al., 1997). The attenuation of neural representations of

distractor stimuli under high perceptual load has also been demonstrated in subcortical

structures (e.g. Lateral Geniculate Nucleus; O'Connor et al., 2002) demonstrating the

possibility that this attentional modulation is not limited to cortical processing. In the

auditory system, some form of early selection could potentially occur via MOCR

modulation of cochlear gain, which would propagate forward through the auditory

pathways, resulting in large changes in the cortical representations of unattended stimuli

(e.g. Mesgarani and Chang, 2012).

It is critical to note that perceptual load is not increased by reducing stimulus fidelity or by performing near-threshold stimulus judgements, but rather by increasing the processing demands (Lavie et al., 2004). In the case of Rees et al. (1997) the high-load task involved determining whether visually presented words were bi-syllabic, which would require phonemic understanding of the presented word (an increase in the linguistic processing load). The low-load task was to identify whether words were presented in uppercase letters, which does not even require that participants attempt to read or recognize the word. Increasing the number of stimuli presented to a participant (e.g. increasing the number of possible targets and/or distractors in a set of stimulus) is another method by which perceptual load can be increased (Lavie, 1995, 2005). As it relates to our hypothesis, a high perceptual load could lead to earlier attentional filtering and may produce a more robust medial olivocochlear (MOC) efferent effect than a low perceptual load.

Most previous studies of the effects of selective attention on MOCR activity have utilized tasks that involve a low perceptual load. Examples of these low load tasks include detecting a change in the intensity of a single repeated tone pip (Giard et al., 1994), detecting a change in visual grating orientation (Srinivasan et al., 2012), or detecting a target "Q" among a series of "O" stimuli in a serially presented stream of letters (Lukas, 1980). These tasks involve attending to a single stimulus without the presence of distractors in order to detect a change in a single feature. Recent studies by Walsh et al. (2014a, 2014b, 2015) involved listening to or reading a series of digits, which imposes more perceptual load than simply detecting a deviant response. This task also imposed a memory load, as subjects had to memorize the series of digits and

compare them to a probe. In contrast to a high perceptual load, a high working memory

load has been shown to increase the influence of distractor stimuli (Lavie et al., 2004)

suggesting that high memory load may interfere with early selection. Another recent

study (Beim et al., 2018) also involved a memory load in the form of counting target

stimuli, and the visual 1-back task was a memory-based task, but little in the way of a

perceptual load. In addition, the task involved gaps of 2 seconds between targets, which

could have led to momentary lapses in attention that would not have been detected but

may have contributed to variability in SFOAE measures both within and across listeners.

Here we present two experiments designed to improve the attentional demands of

our tasks by ensuring a high perceptual load in both the auditory and visual attention

tasks, and ensuring sustained attention is necessary for optimal performance. Experiment

1 focuses on improving on the methodology used in Chapter 2 in two key ways: 1)

increasing the perceptual load of the visual task, 2) collecting behavioral responses to

target stimuli immediately after detection rather than the end of a trial. Experiment 2

further improves the experimental design by utilizing continuous auditory stimuli to

better insure sustained attention.

Both experiments also examined whether effects of selective attention depend on

the presence of the acoustic MOCR elicitor. Previous studies of selective attention effects

on cochlear function, including our the experiment reported in Chapter 2, have utilized

noise to elicit the MOCR to examine whether or not the effect of the MOCR is modulated

by attention (e.g. de Boer and Thornton, 2007; Harkrider and Bowers, 2009; Walsh et al.,

2014b, 2014a, 2015). When investigating auditory attention, Walsh et al. (2014a) claimed

that larger attention effects were seen than in previous studies, due to the presence of

their MOCR elicitor noise. It is also possible that the corticofugal projections to the medial olivary complex could allow attention to activate the MOC efferent in the absence of an acoustic elicitor. Although some studies have examined attention effects on cochlear processing without the use of an acoustic MOCR elicitor, (Avan and Bonfils, 1992; Giard et al., 1994; Puel et al., 1988), the findings as they relate to attention are not conclusive. Avan and Bonfils (1992) found no effect of an attentional task on OAEs, while Giard et al. (1994) saw larger transient-evoked OAEs when an auditory stimulus was attended. No study has directly compared the effects of attention measured with and without an acoustic MOCR elicitor in the same participants under otherwise identical conditions.

II. Experiment 1

A. Rationale

The auditory stimuli in Beim et al. (2018) and consisted of sequences of tones where listeners had to detect a conjunction of two features, the tone frequency and amplitude modulation (AM), in order to complete a task. In that study, there could be both on- and off-frequency distractor stimuli, as on-frequency tones without modulation were still possessed a salient onset cue that listeners would need to ignore. According to literature on perceptual load (for review see Lavie, 2005; Lavie and Tsal, 1994) both the conjunction of the features and the similarity of distractor to target stimuli would increase the perceptual load. In contrast the visual task used in Beim et al. (2018) presented single letters in isolation and relied on the participants memory of the previous stimulus to completed. We improve the perceptual load of the visual task by using a multiple object

tracking (MOT; Pylyshyn and Storm, 1988) paradigm where participants monitor multiple targets in the presence of distractors that possess identical features.

The auditory stimuli from Beim et al. (2018) were combined with the visual MOT stimuli and task to ensure that perceptual load remains high across all attention conditions. The auditory behavioral tasks used in Experiment 1 involved detecting AM tone pips at an attended frequency, while ignoring tone pips and AM occurring at a distractor frequency. A second key improvement in Experiment 1 over Beim et al. (2018) was the collection of behavioral responses immediately after detecting targets rather than at the end of a run. This change decreases the dependence of our behavioral performance metric on memory and allows for a better examination of the relationship between the performance on the behavioral tasks and the SFOAE magnitude measured during the same runs.

B. Method

1. Overview

Participants were presented with a series of auditory and visual stimuli that remained present across three attentional conditions. The auditory stimuli were comprised of short low-and high-frequency tones. The tones were presented both with and without a broadband noise with two spectral notches designed to elicit the MOCR. Visual stimuli consisted of moving red squares presented on a black background. At the beginning of a run, participants were instructed to attend to one aspect of the combined stimulus: the low frequency tones (hereafter: AL), the high frequency tones (AH) or the

visual stimuli (V) and perform a behavioral task as described in the following sections.

During each of the six conditions, (3 attention conditions x 2 MOCR elicitor conditions)

SFOAEs were measured using the low frequency auditory stimuli to investigate task

dependent changes in SFOAE magnitude. Emission magnitudes were compared across

conditions both within- and across individuals to look for effects of selective attention

and the MOCR elicitor.

2. Participants

A total of 24 NH listeners (6 male) participated in Experiment 1. Nine listeners

were excluded from the final sample due to inability to pass SFOAE screening

procedures. The final sample of 15 participants contained 3 males and had a mean age of

25.4 years (range 18-60). All participants in both Experiment 1 and Experiment 2 had

normal hearing (> 20 dB HL at octave frequencies between 0.25- and 8-kHz), provided

written informed consent before participating in the study, and were monetarily

compensated for their participation.

3. Auditory Stimuli

Auditory stimuli were used both to evoke SFOAEs and to perform the auditory

behavioral tasks. Each experimental trial began with a baseline measurement of SFOAE

used to extract SFOAE from the task stimuli. The stimuli used for this baseline

measurement were identical to those used in Chapter 2. A 2-s gap followed the offset of

the baseline SFOAE measurement. Task relevant stimuli began after the gap with the

presentation of a 1-s cue tone at either a low frequency determined by the SFOAE

screening procedure or 4 kHz to indicate which frequency participants should attend to complete the auditory task. When present, the MOCR elicitor began 500 ms after the onset of the cue tone. The elicitor was a 1.7-s Gaussian noise with two 1-octave notches centered at 0.75- and 4-kHz. The elicitor was presented at an overall root -mean-square (RMS) level of 55 dB SPL. Following the offset of the cue tone, a total of four 300-ms tone pips (two at each frequency) were presented at 35 dB SPL. The tone pips were gated with 5-ms ramps to reduce spectral splatter. The two tones at each frequency had randomized inter-stimulus intervals (ISIs). These random ISIs were used to make the timing of the stimuli less predictable to ensure that listeners would need to vigilantly attend the sequence of tones to detect the target stimuli. The ISIs were constrained to have a minimum duration of 100 ms and to ensure that the offset of the final tone in the sequence occurred no later than the end of the MOCR elicitor. The sequence of the cue tone, tone pips, and elicitor noise together formed a block (shown in Fig 3.1) and repeated four times per run with a 2-s gap between blocks to allow for a decay of MOCR effects produced by the elicitor (when present). Up to three of the eight tone pips at each frequency in a single run were amplitude modulated at a rate of 10 Hz with 100% modulation depth. The random selection of which tones were modulated was done independently for each tone frequency and was constrained so that only one tone from each pair could be modulated, ensuring the other tone would be a valid SFOAE probe. During the auditory tasks, participants were required to respond within one second via button press to the modulated tones at the frequency indicated by the cue tone while ignoring the modulated tones at the opposite frequency. A schematic representation of the auditory stimuli with the MOCR elicitor present is shown in Fig. 3.1.

Auditory stimuli were digitally generated using custom MATLAB software (The

Mathworks, Natick, MA), converted to analog signals by a Lynx Two-B sound card

(Lynx Studio Technology, Costa Mesa, CA), and presented to listeners via transducers in

the ER10x (Etymotic Research, Elk Grove, IL) acoustic probe assembly. Recordings

were made using the microphones in the ER10x probe assembly, digitized by the Lynx

Two-B and stored on a computer hard drive for analysis. Simultaneous audio playback

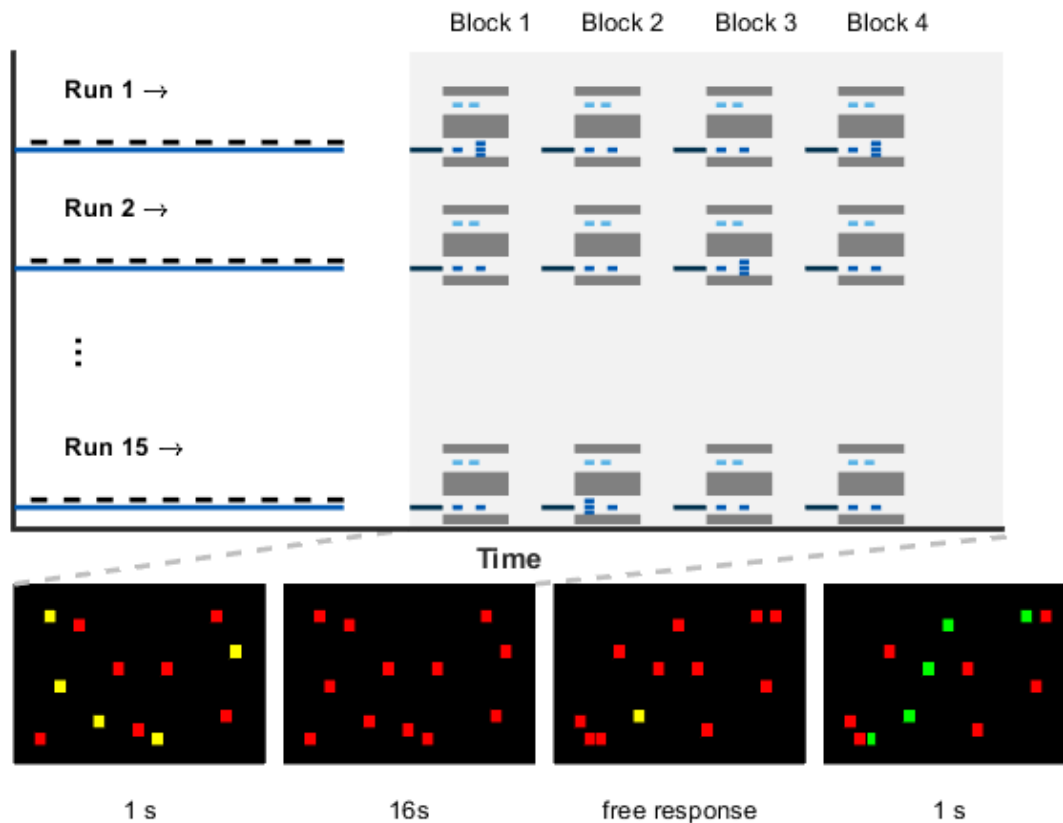and recording was controlled using the Psychophysics toolbox (Brainard, 1997). The

Figure 3.1

Schematic representation of the stimuli used in experiment 1. The top portion of the figure shows time frequency representations of the auditory stimuli. The SFOAE probe is shown in dark blue. Suppressor presentations shown in black are in alternating polarity. MOC elicitor noise is illustrated by the shaded gray regions. AM segments are illustrated by their flanking frequency components for both the low-frequency tone (shown in dark blue) and the high-frequency tone (shown in light blue). The inset at the bottom shows the timing for the visual stimuli.

sound delivery system was calibrated with a Type 4153 artificial ear and 2 cc coupler

(Bruel & Kjaer NA, Duluth, GA) to verify output sound pressure in an average ear canal.

4. Visual Stimuli

The visual stimuli began concurrently with the sequences of short tones used for

the auditory behavioral task. We used a modified version of a classical set of multiple

object tracking (MOT; Pylyshyn and Storm, 1988) stimuli that was adapted from

(Makovski and Jiang, 2009). At the onset of visual stimulus presentation, 12 squares were

presented at random locations on a computer monitor. Each square subtended

approximately 0.28° of visual angle. Squares were presented in color on a black

background that subtended approximately 15°x18°. A subset of 5 squares were initially

presented in yellow, while the remaining 7 were presented in red. Yellow squares were

used to cue participants to attend to them throughout their motion. Yellow squares turned

red and all squares began movement on random trajectories 2 seconds after appearing on

the screen. Motion trajectories were calculated to avoid collision or overlap between

nearby squares. After 15 seconds, all squares ceased movement simultaneously with the

offset of the auditory stimuli. When motion stopped, a free response interval began, and a

mouse cursor appeared so that participants could select the squares that were initially

yellow at the beginning of the trial. After the participant had selected a total of 5 squares,

the cursor was removed from the screen and feedback was provided by highlighting any

correct responses in green. Any missed targets were not shown to participants. The free

response interval and the performance feedback were only present when participants

completed the visual task. During the auditory tasks visual stimulus presentation ended 1

s after motion ceased. The visual stimuli and their trajectories were generated prior to

each trial and presented at a rate of 30 frames per second on a computer monitor. Sample

frames depicting the stimuli during various stages of the MOT task are shown at the

bottom of Fig 3.1

5. Procedure

All procedures were conducted in a double-walled sound attenuating chamber

(Industrial Acoustics Company, Bronx, NY) with participants seated in a semi-reclined

position in a chair to minimize the presence of motion related artifacts in the SFOAE

recordings. Participants kept a keyboard on their lap to make responses during the

auditory attention conditions and used a computer mouse to respond to the visual

attention conditions. Before beginning the experimental tasks, participants first

completed screening procedures to ensure that spontaneous otoacoustic emissions were at

least 100 Hz away from the experimental frequency and that SFOAEs could be measured

with high signal to noise ratios. The screening procedures are described in full detail in

Chapter 2. Participants then completed 15 training runs for each of the attention

conditions to familiarize themselves with the task. Training stimuli did not contain AM at

the distractor frequency for each of the auditory attention conditions. The MOCR elicitor

noise was also not used during training. Target detection rates of at least 75% were

required using the training stimuli in order for participants to move on to the full

experiment. In a second session after the training, participants completed the 6

experimental conditions (3 attentional conditions *x* 2 MOC-reflex elicitor conditions).

Each participant completed 15 runs of each condition in a counterbalanced order in order

to minimize order effects.

6. SFOAE Analysis

The same method for extracting SFOAE from the recorded waveform as in Beim

et al. (2018) to analyze the data from this study. The baseline measurement of SFOAE

was used to estimate SFOAE evoked by the task stimuli. Vector subtraction to extract

emission from the task stimuli used 200-ms windows centered on the duration of the

300ms tone pips. Tone pips containing AM were excluded from the analysis such that an

unmodulated tone pip from each of the four blocks present in a trial was combined in the

averaging procedure. Tone pips from all runs that passed artifact rejection procedures

identical to those in Chapter 2 were averaged across blocks to produce a single average

waveform from which the SFOAE residual was extracted.

7. *Bootstrapping procedure*

A bootstrapping procedure was used to quantify any significant changes in

SFOAE that make occur across attention conditions within an individual listener.

Artifact-free pairs of stimuli were resampled with replacement from the available pool of

artifact free recordings such that the new sample of recorded data contained the same

number of recordings as the original artifact-screened data. The SFOAE analysis

described above was carried out on the resampled data to extract SFOAE magnitudes for

each attention condition separately. The resampling procedure was repeated 10,000 times

per experimental condition to yield distributions of SFOAE magnitudes for each individual. Pairwise comparisons were made using the bootstrap procedure, and differences between conditions were considered significant when the mean SFOAE magnitude estimated from one distribution fell outside the 95% confidence interval of the second distribution.

C. Results

1. Behavioral Results

Hit and false alarm rates were used to calculate participants' sensitivity, d', to the presence of AM on the attended carriers (Green and Swets, 1966). We did not estimate d' for the visual task because participants were required to make 5 responses corresponding to 5 targets, meaning that misses and false alarms would not be independent of one another. Instead performance on the visual task was scored as the percentage of correctly identified targets.

Performance on the auditory tasks in the conditions without the MOC elicitor was well above chance with mean d' scores of 2.31 and 2.51 for the AL and AH conditions respectively. When the MOCR elicitor noise was present performance for the AL condition dropped to 1.90. Mean performance in the AH did not decline (mean: 2.68). A repeated measures ANOVA confirmed an interaction between the effect of the MOC elicitor and attention conditions [$F(1,14) = .5.27$, $p = .038$]. There was no significant main effect of attention condition [$F(1,14) = 3.3$, $p = .091$] and no significant effect of the MOC elicitor [$F(1,14) = .284$, $p = .603$] on behavioral performance.

Visual task performance was 72% and 70% correct in the presence and absence of the MOC elicitor respectively and there was no significant effect of the MOC elicitor on visual task performance as expected [$t(14) = 1.08$, $p = .301$].

2. SFOAE Results

The bootstrap procedure was used to evaluate the significance of shifts in SFOAE magnitude within listeners. One goal of this analysis was to find any significant patterns in SFOAE magnitude across attention conditions that may not be obviously reflected in the data pooled across listeners. The bootstrap analysis revealed significant shifts in emission magnitudes across attention conditions in multiple listeners, but these effects
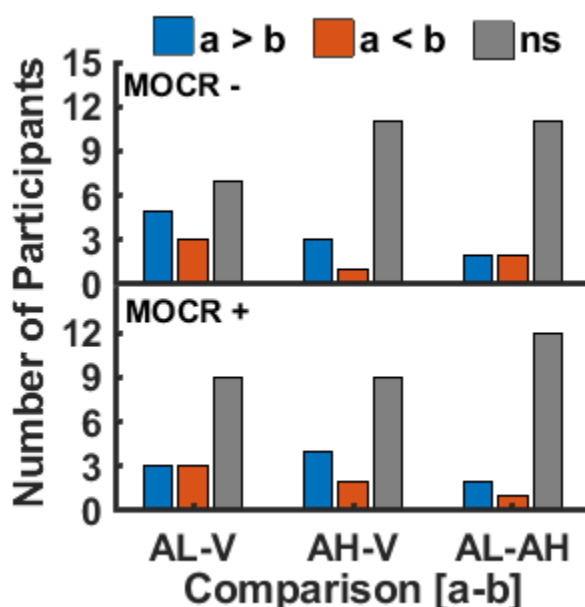
Figure 3.2

Summary data from the bootstrap analysis. Each set of three bars shows number of participants exhibiting each possible directional relationship between SFOAE magnitudes for the 3 attentional condition comparisons of interest. Grey bars show proportion of cases for each paired comparison where there was no significant difference. Blue bars show proportion of cases where the first condition produced larger SFOAE magnitudes than the second (e.g. in the first set of bars showing the AL-V comparison, the blue bar indicates the proportion of cases where SFOAEs were significantly larger in the AL condition). Orange bars show a significant effect in the opposite direction of the blue bars. The top panel shows the results obtained with the MOCR elicitor absent (MOCR-), while the bottom panel shows results for the MOCR elicitor present (MOCR+).

were not consistent across listeners in either of the MOC-reflex elicitor conditions. Figure

3.2 is a summary of the bootstrap results, illustrating the number of significant shifts in

SFOAE magnitudes for each pairwise attention condition comparison.

The mean SFOAE magnitudes for each participant across all attention and elicitor

conditions are shown in Fig 3.3. We used a 2x3 repeated measures ANOVA with factors

of MOC elicitor and attention condition to look for changes in SFOAE magnitude across

participants. The ANOVA revealed a significant main effect of the MOCR elicitor

$[F(1,14) = 14.1, p = .002]$ with lower emission magnitudes observed when the MOCR

elicitor was present. There was no significant main effect of attention $[F(2,28) = 1.54, p =$

.232] and no interaction between the MOC elicitor and attention condition $[F(1.21,16.95)$
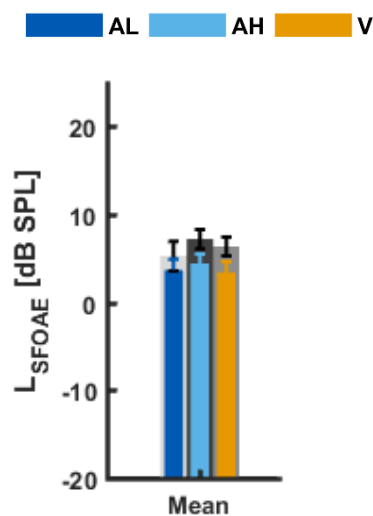
$= .435, p = .556]$.

Figure 3.3

Mean magnitudes across participants (n=15) of SFOAE extracted from the behavioral task stimuli. Colored bars indicate different attention conditions. Grey bars behind each colored bar indicate SFOAE magnitudes when the MOC elicitor was not present for the same attentional condition. Error bars show standard error of the mean.

3. Relationship between behavioral and SFOAE data

A correlational analysis was undertaken to look for evidence of a relationship between performance in the behavioral tasks and SFOAE magnitudes evoked by the low-frequency tone during the three behavioral tasks. Scatter plots of the measures of behavioral performance plotted against the corresponding SFOAE magnitudes are shown in Fig 3.4. There were significant correlations between behavioral performance and emission magnitude in the AL condition when the MOC elicitor was not present [$r(13) = .65$, $p = .008$], but no other significant correlations were observed in the remaining conditions [$r(13) < .45$, $p > .094$]. Across all attention conditions there was a trend towards increasing SFOAE magnitude accompanying better performance.
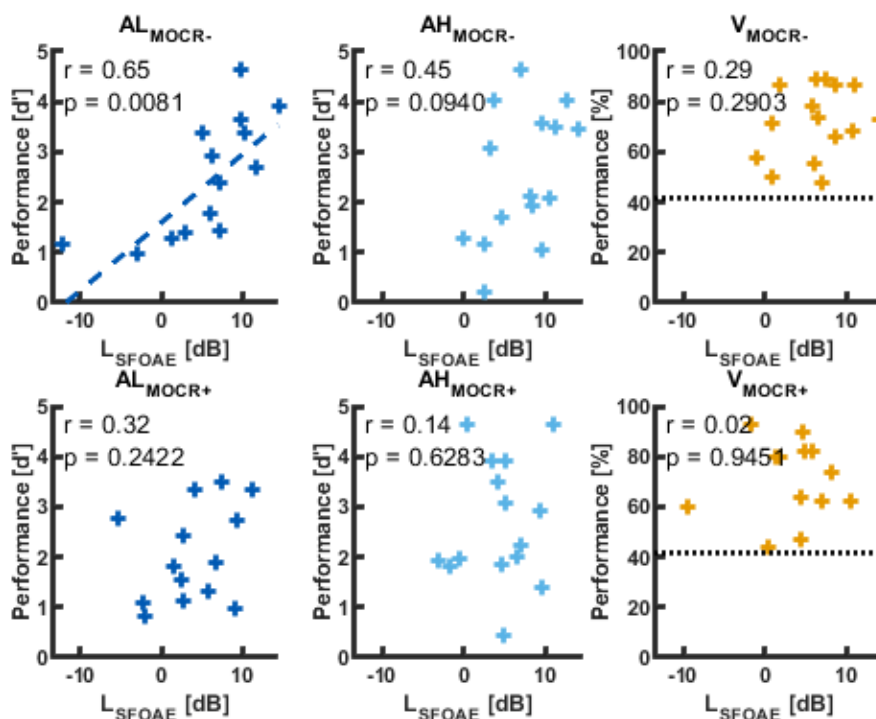

Figure 3.4

Results of correlational analyses between task performance and SFOAE magnitude. Each row of panels presents data for each attention condition. The MOCR elicitor is absent for the data presented in the top panels, and present for the data shown in the bottom panels. Black dotted lines in the panels for the V conditions correspond to chance performance on the visual task.

D. Discussion

  Despite increasing the perceptual load of the auditory stimuli by using a sequence of tone pips, we failed to observe any consistent effect of selective attention on cochlear function as assessed by SFOAEs. The data from Experiment 1 is quite similar to the replication sample from Beim et al. (2018). Both the results from that study and the current experiments show some evidence of significant shifts in SFOAE magnitude across attention conditions, but no consistent pattern in those shifts. The correlational analysis yielded a single significant relationship between SFOAE magnitude and behavioral performance in the AL condition when the MOCR elicitor was absent. The fact that the relationship was only present for one of three attentional conditions suggests that the relationship is likely spurious, as significant relationships would be expected across all three attention conditions if the MOCR effects were modulated under selective attention.

  The lack of evidence demonstrating an effect of attention in Experiment 1 suggests that either our manipulations to the auditory stimuli had no consistent effect on the magnitude of SFOAEs or that the SFOAE measure itself is not sensitive to the manipulations. One factor that may decrease the sensitivity of this measure is variability caused by participants learning the approximate timing of the behavioral stimuli. Due to the cyclic presentation of the cue tone and the probes that follow it, participants may begin to anticipate the downtime between blocks of stimuli and allow lapses in their attention. These attentional lapses could produce unwanted variability in the data making it more difficult to detect a true effect of selective attention if it exists. Experiment 2

addresses this limitation by modifying the experimental stimuli and behavioral tasks to better require sustained attention.

III. Experiment 2

A. Rationale

Our first experiment focused on using tasks with high perceptual load and measuring cochlear function in a high-frequency region, where cochlear gain should be high, to maximize chances of observing an attentional effect. The target tones were relatively brief and were separated by random silent intervals. It is possible that attention lapsed somewhat during these intervals, only to be reengaged by the tone onset. In this experiment, the sequences of tone pips were replaced by long tones that could contain short bursts of amplitude modulation (AM) throughout their duration, eliminating the silent gaps where attention was not required. The continuous auditory stimuli also better matched the visual MOT task used in Experiment 1 since both tasks required continuous sustained attention.

B. Method

1. Overview

Similar to Experiment 1, participants were presented with a series of auditory and visual stimuli that remained identical across three attentional conditions. Cochlear function during each of the conditions was assessed by SFOAEs evoked by tones at a fixed frequency. The task stimuli consisted of continuous low- and high-frequency tones

as well as a set of visual MOT stimuli. During a run, the low- and high-frequency tones were amplitude modulated for a short duration at random intervals. Participants completed the same attentional conditions (AL, AH, and V) as in Experiment 1. When attending to the tones, participants were required to respond with a keypress as soon as they detected an AM segment on the tone they were attending. The participants completed the three selective attention conditions with and without a notched-noise MOCR elicitor to examine whether any attention effects are influenced by acoustic activation of the MOCR. The order in which participants completed the attentional conditions was counterbalanced to control for order effects. As in Experiment 1, the effects of attention and the MOCR elicitor were examined by comparing the levels of the SFOAE evoked by the same stimulus across the conditions.

2. Participants

A total of 20 normal-hearing (NH) participants were recruited for this experiment. Participants needed to pass both an SFOAE screening procedure and training on the behavioral task to be included in the experiment. Five listeners (1 male, 4 female) were excluded from the sample because they did not have SFOAEs with a sufficiently high SNR around the probe frequency used to measure the effect of selective attention. The remaining 15 participants (11 female, 4 male), aged 18-33 (mean: 23.5), were included in the final dataset. None of the listeners that participated in Experiment 1 took part in Experiment 2.

3. Stimuli and procedure

The visual stimuli were identical to those used in Experiment 1. The bottom panels of Fig. 3.5 show the timing of the visual stimulus presentation as it relates to the auditory stimuli. Changes were made to auditory stimuli involved using continuous tones rather than tone pips. The baseline SFOAE was measured using the suppression paradigm described in Chapter 2. The emission was measured only in the 750-Hz region and only the SFOAE at this frequency was used to monitor changes to cochlear responses across the three attentional tasks.

The task-relevant stimuli began after a 2-s silent gap that followed the offset of the probe tone. After the silent gap, either a low (approximately .75k Hz) or high (4 kHz) tone with a duration of 17 s was presented. The first 1 s of this tone served as the cue for participants to selectively attend to the low or high frequency (AL or AH condition, respectively). After 1 s, a second 16-s tone at the other frequency began so that both tones ended simultaneously. Each tone was presented at 40 dB SPL and could contain multiple segments that were amplitude modulated at full (100%) depth at a rate of 10 Hz for a duration of 300 ms. The timing of the AM segments was randomized for both the high and low frequency tones with the mean time of 1.7 s between AM segments at a given frequency. The only constraint on the timing of the modulations was the inclusion of a protected window that was used for extracting SFOAEs. Each tone could have up to 5 AM segments within a run. Participants were required to attend to one frequency and respond within 1 s via button press to each modulated segment in the attended tone while ignoring the modulations on the other tone.
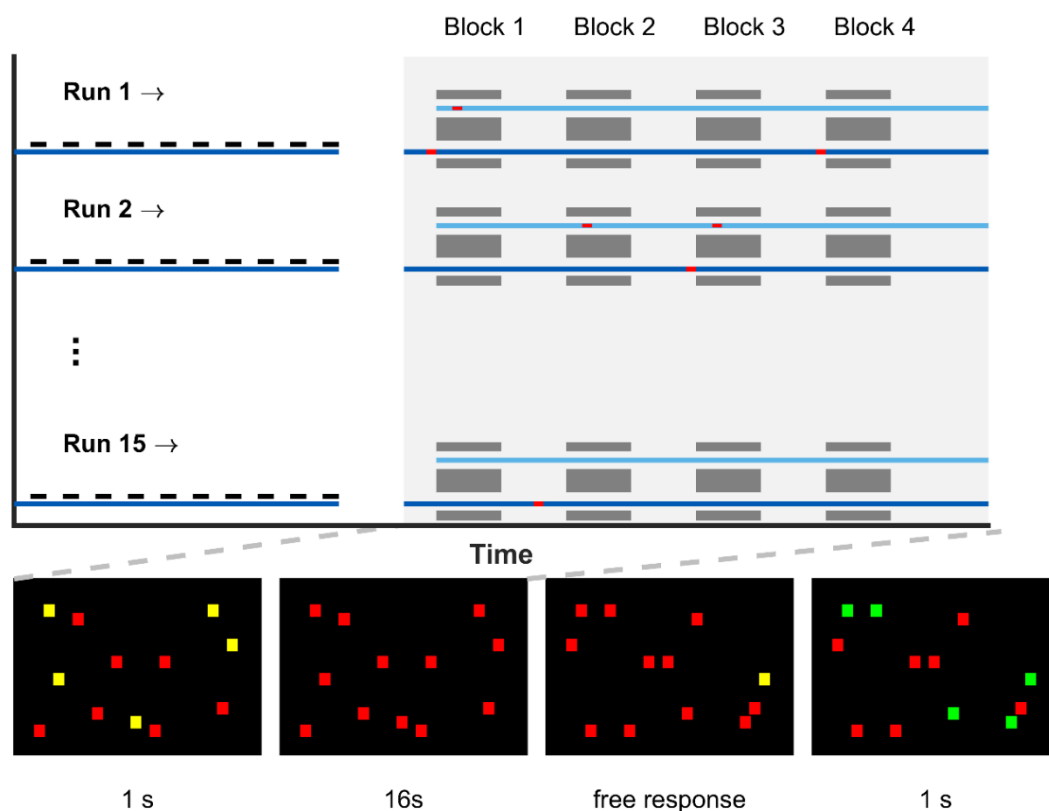
Block 1   Block 2   Block 3   Block 4

Run 1 →

Run 2 →

Run 15 →

Time

1 s   16s   free response   1 s

Figure 3.5

Schematic representation of the stimuli used in experiment 2. The top portion of the figure shows time frequency representations of the auditory stimuli. The SFOAE probe is shown in dark blue. Suppressor presentations shown in black are in alternating polarity. MOC elicitor noise is illustrated by the shaded gray regions. AM segments are illustrated short red segments during both the low and high frequency tones. The inset at the bottom shows the timing for the visual stimuli.

Attention effects were measured using the stimuli described above presented alone in one condition and with an MOCR elicitor in another condition. When present, the elicitor began 500 ms after the onset of the first tone. The elicitor was a 2-s noise that spanned a range from .1- to 10 kHz in frequency with two 1-octave notches centered on the frequencies of the tones. The noise was gated on and off every 2 seconds and was presented at an overall level of 60 dB SPL. A schematic representation of the auditory stimuli is shown in the top half of Fig. 3.5.

Prior to the experiment, participants completed training as described in Experiment 1. During the experiment, a total of six conditions were tested (3 attention conditions x 2 MOCR elicitor conditions). The experimental runs were grouped by condition and participants completed 15 runs for each condition. The order in which conditions were completed was counterbalanced across participants to attempt to control for any potential order effects. Participants completed all six attention conditions within a single experimental session, during which the probe assembly was not removed from their ears. The methods for the generation and recording of the stimuli and the equipment were identical to those in Experiment 1.

4. SFOAE Analysis

The emission extraction technique was the same as in Experiment 1. Adjustments to averaging windows were made to accommodate the continuous auditory stimuli. The 16 s of the low-frequency tone presented during the behavioral task portion of the experiment was broken into 4 segments per run lasting 4 s each. The stimuli for both the baseline SFOAE measurement and the task stimuli were averaged across the number of accepted runs. The same heterodyne procedure outlined in Experiment 1 was used to estimate the ear-canal sound pressure during the segments with the suppressor tone present. Next, the same mean of this sound pressure was subtracted from the 4-s segments corresponding to the behavioral task stimuli, yielding the SFOAE residual magnitude and phase. For each participant the mean emission magnitude was calculated from a 300-ms window centered temporally in the averaged 4-s task stimulus waveform.

This timing was chosen because AM segments were distributed such that they did not occur during this portion of the waveform across 4-s blocks and across runs. Example of the extracted SFOAE magnitude and phase from a single participant are shown in Fig. 3.6. The same bootstrapping procedure outlined in sec. II.B.7 was used to evaluate whether or not significant shifts in SFOAE magnitude occurred within individuals.
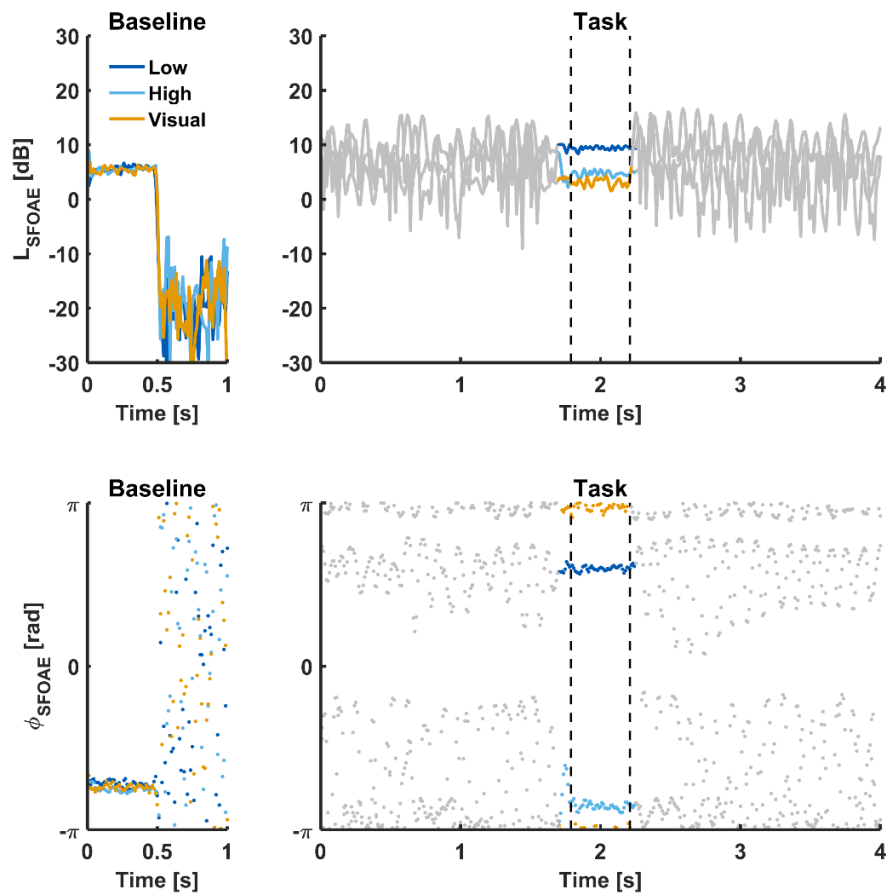

Figure 3.6

Example OAE data extracted from a sample participant in experiment 2. Left panels show magnitude and phase during the baseline measurement of SFOAE. In the right panels, extracted SFOAEs are shaded grey during the regions in which AM can occur and are shaded in color during the analysis window. Fluctuations in level due to averaging segments of the tone containing AM can be clearly seen in the grey regions. Different colored traces correspond to differing attention conditions. Note that in the left panels SFOAE magnitudes for each condition overlap but diverge when this participant performs the task (as shown in the right panels).

C. Results

1.   Behavioral Results

We quantified the behavioral performance in Experiment 2 using the same

methodology as Experiment 1. Across listeners auditory task performance ranged from d'

values of 0.96 to 4.65 (representing perfect performance after correction) with mean

values of 2.78 and 2.88 for the AL and AH conditions, respectively. In conditions with

the MOCR elicitor present, performance in the auditory tasks was worse overall. Mean d'

values were 1.59 and 2.05 for the AL and AH conditions, respectively. A repeated-

measures ANOVA with factors corresponding to the presence or absence of the MOCR

elicitor and attended frequency (low or high) revealed no significant effect of the

attended frequency [$F(1,14) = 1.14$, $p = .304$]. There was a significant main effect of the

MOCR elicitor on behavioral performance [$F(1,14) = 34.5$, $p < .001$] confirming worse

performance when the MOCR elicitor was present. There was no significant interaction

between the elicitor and attention conditions. [$F(1,14) = 1.4$, $p = .256$].

Visual task performance was analyzed separately from the auditory task

performance and ranged from 65-96% correct (from 3.25 to 4.8 targets out of 5 correctly

identified) with a mean of 80% correct and was nearly identical between the elicitor

present and elicitor absent conditions.

2.   SFOAE Results

First, an individual analysis was undertaken using the bootstrap technique, as

described in Experiment 1. Figure 3.7 shows a summary of the bootstrap analysis. Most

participants did not show significant shifts in OAE magnitude across attention conditions.

Those who showed significant effects did not exhibit effects in a consistent direction.

This was true for both MOCR elicitor present (MOCR+) and absent (MOCR-)
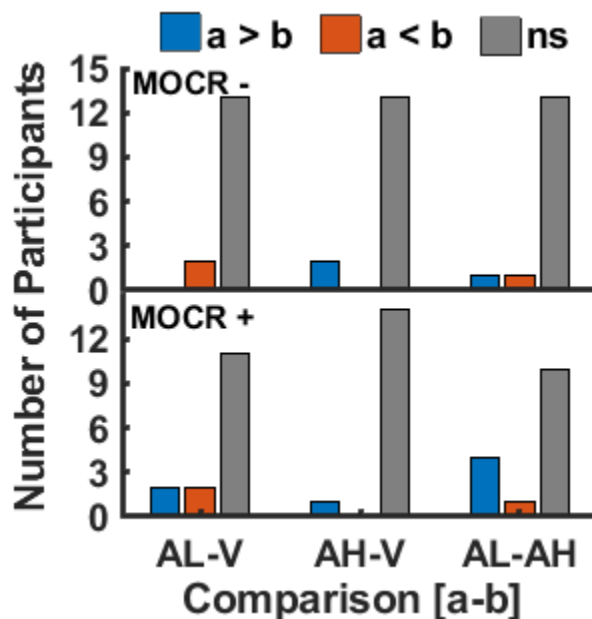
conditions.

Figure 3.7

Bootstrapping summary data for experiment 2. Bars as described in Fig. 3.2. Top panel shows results with the MOCR elicitor absent (MOCR-), while bottom panel shows the results with the MOCR elicitor present (MOCR+).

Next, we examined the effect of the MOCR elicitor and attention conditions

across listeners. Figure 3.8 shows levels of SFOAE extracted from the low-frequency

tone for each attentional condition. Greyscale bars behind each colored bar show the level

of SFOAEs when the elicitor noise was absent. We conducted a repeated-measures

ANOVA with factors MOCR elicitor and attention condition. The ANOVA revealed a

significant effect of MOCR elicitor [$F(1,14) = 35.1$, $p < .001$] with lower SFOAE

magnitudes when the elicitor was present. However, there was no main effect of attention

condition on SFOAE magnitudes [$F(2,28) = 2.17$, p = .133], and no interaction [$F(2,28) =$ 1.40, p = .263], suggesting that acoustically eliciting the MOCR does not change the nature of observed attentional effects. There were no significant effects of, or interactions between, attention and MOCR elicitor presence on the baseline SFOAE and the noise floor recorded during each measurement [F < 1.39, p > .258].
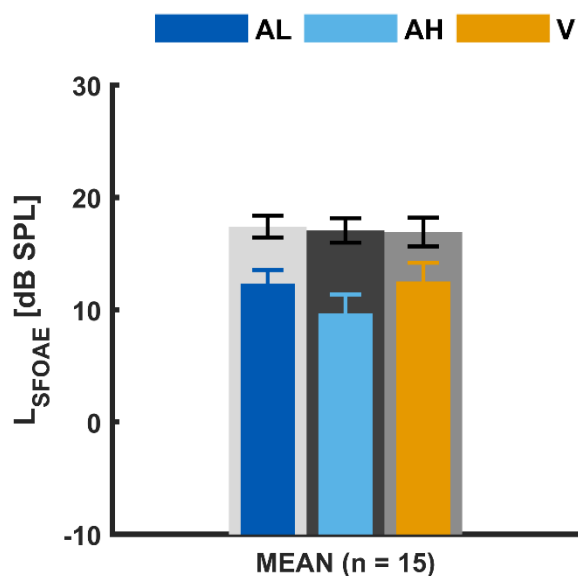

Figure 3.8

Mean magnitudes across participants of SFOAE extracted from the behavioral task stimuli. Colored bars indicate different attention conditions. Grey bars behind each colored bar indicate SFOAE magnitudes when the MOC elicitor was not present for the same attentional condition. Error bars denote standard error of the mean.

3. *Relationship between behavioral and SFOAE data*

Results of the correlational analysis are shown in Fig 3.9. There were no significant correlations between performance in the auditory conditions and the level of respective SFOAEs extracted from those trials [$-.263 < r(13) < .422$, p > .08]. This was true for both MOCR elicitor conditions. There was also no significant correlation

between performance on the visual task and SFOAE magnitude [r(13) < .032, p =.453],

although there was a negative trend that neared significance when the MOCR elicitor was

present [r(13) = -.453 p = .06].
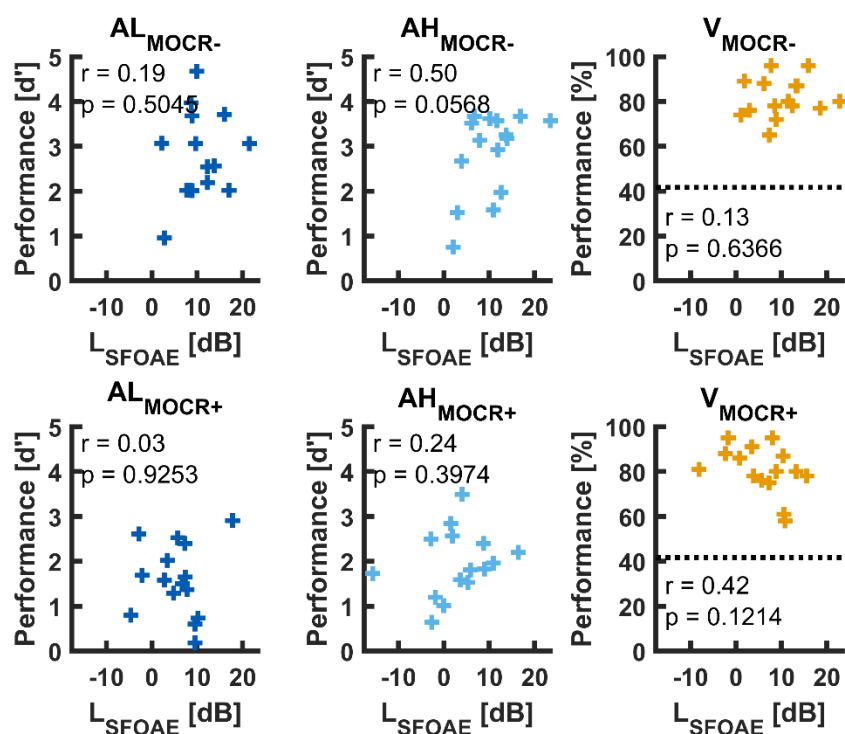
Figure 3.9

Correlational analysis between behavioral performance in experiment 2 and SFOAE magnitudes extracted from the task stimuli. Each panel shows data from a different experimental condition.

D. Discussion

In this experiment, we wanted to test whether the presence of an MOCR elicitor

noise influenced the effect of selective attention on cochlear responses, and whether

sustained attention (via a continuous probe tone) might better reveal attentional

modulation of cochlear function. The MOCR elicitor produced the expected effect of

reducing SFOAE magnitude when it was present, but we found no evidence of attentional

modulation in either the individual or the group data. The trend in the correlational analysis towards lower SFOAE magnitudes with increasing performance on the visual task might be consistent with the same type of reduction in SFOAE observed in Beim et al. (2018), but it was not statistically significant, even before any correction for multiple comparisons, and the reduction in SFOAE magnitude during the V condition observed in Beim et al. (2018) was not visible in the bootstrap analysis or the average data across attentional conditions. Taken as a whole, the data from this experiment suggest that MOCR effects were detectable in our participants while they were performing a perceptual task, but that the MOCR effects do not appear to differ in any systematic way across attentional conditions and were not systematically related to performance on the tasks used in the study. This outcome suggests that the use of continuous stimuli did not improve our ability to detect attentional changes, as the overall pattern of results mirrors that of Experiment 1.

The lack of an effect suggests that the attentional modulation of cortical activity observed in many studies of attention arises at higher levels of processing in the auditory pathways. However, it is also possible that the SFOAE-based measure of cochlear function contains too much variability to reliably detect effects of selective attention.

IV. General Discussion

In this chapter we conducted two experiments designed to follow up on the study by Beim et al. (2018), in which attentional effects on cochlear processing were investigated using SFOAEs. In Experiment 1, we incorporated perceptual load theory into our experimental design, with the hypothesis that a task involving a high perceptual

load would increase early attentional effects. We also examined whether MOCR effects of attention are different at higher frequencies where larger cochlear gain may improve the dynamic range of MOCR effects. Our results with a low-frequency probe tone, finding no effect of attention, replicated the findings of the replication sample in Beim et al. (2018).

Our second experiment sought to answer the questions of whether more sustained attention or the absence of the acoustic MOCR elicitor would produce observable attentional effects on SFOAEs. We observed no differences in the pattern of results between the continuous stimuli used in Experiment 2 and the discrete stimuli used in Experiment 1.

In both experiments, we saw significant shifts in SFOAE magnitude between various attention conditions in a minority of participants. The direction of these differences was roughly evenly split in each comparison, making it difficult to ascribe any functional importance to them. In the vast majority of participants there were no significant shifts in SFOAE magnitude between attention conditions. The acoustic MOCR elicitor did produce the expected decrease in SFOAE levels when present, but attentional effects were not observed either with or without the elicitor. This suggests that there may not be a role for MOC efferents in selective attention. The lack of low-level attentional effects is also supported by work showing no evidence of attentional modulation of electrophysiological signals at the level of the brainstem (Varghese et al., 2015b). Electrophysiological evidence on the attentional modulation of peripheral auditory processing is also inconclusive, as an animal model demonstrated a reduction in the compound action potentials measured while animals attended to visual stimuli,

relative to auditory stimuli (Delano et al., 2007). These differences could reflect species differences in attentional effects, but may also simply arise due to methodological differences between studies. It would be important to attempt to replicate the results reported by Delano et al. (2007), given the large inter-subject variability observed in humans under similar conditions (Beim et al., 2018).

These findings contradict other recent studies reporting an effect of selective attention on cochlear responses (Maison et al., 2001; Srinivasan et al., 2012, 2014, Walsh et al., 2014a, 2014b, 2015; Wittekindt et al., 2014), but agree with other otoacoustic emission studies that did not find such effects (Avan and Bonfils, 1992; Michie et al., 1996). Part of the lack of agreement amongst the existing literature is due to different stimulus and task design choices.

The findings of this study contradict other recent studies reporting an effect of selective attention on cochlear responses (Maison et al., 2001; Srinivasan et al., 2012, 2014, Walsh et al., 2014a, 2014b, 2015; Wittekindt et al., 2014), but agree with other otoacoustic emission studies that did not find such effects (Avan and Bonfils, 1992; Michie et al., 1996). Part of the lack of agreement amongst the existing literature may be due to different stimulus and task design choices.

The results of the current study are in apparent disagreement with recent work by Walsh et al. (2014a, 2014b, 2015), but there are important differences between our study and this previous work that could explain the difference. Perhaps most critically, the Walsh et al. reports compared conditions where participants are not engaged in any task with conditions where participants selectively attended to task-related stimuli. The differences in OAE magnitudes could therefore be due to differences in general arousal

caused by engagement in the task or lack thereof. In contrast, our current study compared only conditions in which participants were engaged in a task, which would not be expected to produce strong differences in arousal.

Studies using distortion product otoacoustic emission (DPOAE) measures of cochlear function have reported effects that differ in direction. For instance, Srinivasan et al. (2012, 2014) reported decreases in DPOAE with auditory attention, whereas Wittekindt et al. (2014) saw decreases in DPOAE magnitude with visual attention. Researchers have often explained this discrepancy in direction of effects seen in DPOAEs as resulting from the MOCR activation producing both increases and decreases in DPOAE magnitude in both humans and animals (e.g. Maison and Liberman, 2000; Müller et al., 2005). An alternative interpretation of the disagreement in data from Srinivasan et al. (2012, 2014) and Wittekindt et al. (2014) is that significant shifts of DPOAE magnitude in different directions are due to noise in the measurement of OAEs in general. The current study examined a large sample of listeners (45 participants, 90 measurements of SFOAE during behavioral tasks for each attention condition) and found no statistically significant shifts in SFOAE magnitude under different conditions of selective attention, but nevertheless revealed large differences between participants. Given this large variability and small (or non-existent) effect size, it is possible that small sample sizes will produce spurious false positive results.

There are many possible causes of noise in SFOAE measurements. The current study tried to control for several sources of variability in measurement by using only high SNR measurements of SFOAE, optimizing stimulus parameters individually for each subject, and by calibrating stimuli using forward pressure to ensure that the level of

sound reaching the cochlea is as invariant as possible across listeners. We also used tasks with high perceptual load as this has been shown in the attention literature to produce evidence of attentional selection earlier in cortical processing streams (Schwartz et al., 2005) and at subcortical sites (O'Connor et al., 2002). Despite these measures, no evidence for effects of selective attention at the level of the cochlea was found here.

We analyzed performance data in conjunction with SFOAE magnitudes as we expected to see evidence of selective attention in the relationship between these variables. For example, if the data from the initial sample of participants in Beim et al. (2018) reflects a true effect, whereby greater cochlear gain reduction is applied while attending a visual stimulus, we might expect to see a negative relationship between task performance and SFOAE magnitude (i.e. best performance accompanied by smallest SFOAE magnitudes) for the visual condition, but not necessarily for either of the auditory conditions. The correlational analyses from experiment 2 show a trend that would support this idea, but the correlation was not significant, and was not consistent with the analyses of data from Experiment 1 which showed no trend.

One potential reason for a failure to observe an attentional effect could be our choice of a low SFOAE probe frequency around 750 Hz. Animal studies have shown that substantially less cochlear amplifier gain at low than at higher frequencies (Recio-Spinoso and Oghalai, 2017). It is possible that a larger gain is needed to observe robust gain changes due to selective attention. In addition, density of MOC innervation of OHCs in animals has been shown to be greater at mid rather than low frequencies (Brown, 2011; Liberman and Brown, 1986). Using a higher frequency SFOAE probe to assess cochlear

function in a higher frequency may result in a larger dynamic range for MOC effects due to the increased density of MOC innervation and a greater cochlear gain.

In conclusion, across two experiments designed to produce optimal chances for observing an effect of selective attention on cochlear function we did not observe any robust changes in SFOAEs that could provide such evidence. The results, combined with other recent studies highlighting the difficulty in interpreting otoacoustic measures (e.g. Francis et al., 2018; Guinan, 2014), suggest that otoacoustic-emission-based measures of MOC activity may not be an ideally sensitive tool to measure peripheral effects of selective attention in humans or that attentional control does not extend all the way to the cochlea.

Chapter 4: Stimulus frequency effects on otoacoustic emission measures of cochlear function during selective attention

I. Introduction

In chapters 2 and 3 experimental methodology for assessing cochlear responses to sound during selective attention tasks was refined in several ways across different experiments. We used state-of-the-art techniques to ensure robust measures of otoacoustic emission that would not be contaminated by confounding factors such as the middle ear muscle reflex. Attention was manipulated carefully in a design that avoided comparing attention to inattention and manipulated the focus of attention across conditions while leaving stimuli constant. Behavioral tasks were designed to produce high perceptual load and ensure that sustained attention was required to perform optimally. Despite these advances, none of the experimental results discussed in chapters 2 and 3 provided compelling evidence that selective attention modulates in a systematic way. This evidence suggests that the medial olivocochlear (MOC) efferent system may not play a role in auditory processing during selective attention. It is also possible that the choice to measure cochlear function with low-frequency SFOAEs has limited the ability to detect MOC reflex (MOCR) effects because the dynamic range of such effects may be smaller at low frequencies like those employed in the experiments in chapters 2 and 3. If the dynamic range of MOCR effects is smaller, significant changes may be difficult to detect due to floor or ceiling effects in MOCR activation. The potentially limited dynamic range for MOCR effects stems from two key factors: less cochlear gain at low frequencies compared to mid frequencies, and potentially less MOC innervation at low frequencies compared to mid frequencies.

Cochlear gain has been demonstrated to smaller at low frequencies than mid to high frequencies both physiologically in animals (Recio-Spinoso and Oghalai, 2017; Robles and Ruggero, 2001) as well as indirectly in humans (Gorga et al., 2007; Plack et al., 2008). Higher cochlear gain would provide a greater range for MOC based gain reduction, potentially increasing the dynamic range of any attentional modulation of the MOCR.

While no anatomical data exists from humans, several animal studies have demonstrated that MOC innervation of outer hair cells (OHCs) is most dense in the medial region of the cochlea and that density decreases both basally and apically to this location (e.g. Guinan et al., 1984; Maison et al., 2003). Other studies have shown that measure of functional MOC inhibition in these animals are also largest for auditory nerve fibers with mid to high characteristic frequencies as opposed to low frequencies (e.g. mice; Vetter et al., 1999).

In contrast to this evidence, one study found that acoustically elicited MOCR effects in humans were smaller at 4 kHz than at 0.5- and 1- kHz (Lilaonitkul and Guinan, 2012). Even though the acoustically elicited MOCR effects were smaller at 4 kHz in Lilaonitkul and Guinan, (2012), Walsh et al. (2015) reported effects of selective attention on OAEs using a non-linear SFOAE evoked by 4 kHz probe stimuli. The inconsistency of these findings warrants further exploration. The current experiment measured SFOAEs with both high and low frequency evoking tones. The low frequency data was collected to ensure that a difference between samples of participants between chapters 2 and 3 and the current experiment could not account for the differences if an effect of attention was observed in the higher frequency SFOAE data.

Here we present an experiment designed to test whether attentional effects manifest differently in separate cochlear regions where gain is expected to differ possibly leading to larger attentional effects at higher frequencies.

II. Method

A. Rationale

Like the experiments in chapters 2 and 3, the primary purpose of this experiment was to determine whether manipulations of attention result in changes to cochlear gain. The experimental design was similar to that used in the first experiment in chapter 3, but was refined in three key ways: 1) measuring cochlear responses in both low- and high-frequency regions, 2) using state of the art calibration techniques to ensure sound levels presented to the cochlea are as consistent as possible across conditions and participants 3) testing a larger sample of participants in order to be more certain about the experimental evidence we obtained. In this experiment, the continuous stimulus design used in chapter 3 experiment 2 was abandoned since that experiment revealed no observable difference in our SFOAE measures of cochlear function during identical attention conditions. The discrete stimuli may also produce higher perceptual load than the continuous stimuli because there are more highly-salient tone onsets that may serve as perceptual distractors in the behavioral task. In addition to implementing changes to the test stimuli, background noise in the ear contralateral to stimulus presentation was monitored during the behavioral task and SFOAE recordings, as some previous investigations of selective attention on cochlear function have noted differences in physiological noise depending on attention (Walsh et al., 2014a, 2014b). The involvement of efferent effects in changes to

noise level was recently disputed by Francis et al. (2018), who argued that the changes in noise are not likely to be due to MOCR activity but instead reflected the tendency of participants to reduce movements while performing a behavioral task. By comparing observed changes in SFOAEs to any observed changes in the noise in the unstimulated ear, it should be possible to determine the degree to which changes in SFOAE data might be related to noise generated from participant motion.

B. Overview

Participants were presented with a series of auditory and visual stimuli that remained present across three attentional conditions. The auditory stimuli comprised short low- and high-frequency tones embedded in a noise with two spectral notches. Noise was used to elicit the MOCR because attention effects have been observed more consistently in conjunction with acoustic efferent activation (Froehlich et al., 1990, 1993; Maison et al., 2001; Michie et al., 1996; Veuillet et al., 1991; Walsh et al., 2014b, 2014a, 2015) than without an acoustic elicitor (Picton et al., 1971; Picton and Hillyard, 1974). The visual stimuli comprised 12 red squares in a black background on a computer screen. At the beginning of each run, participants were instructed to attend to one aspect of the combined stimulus, low tones (hereafter term the AL condition), high tones (AH condition), or visual stimuli (V condition), and to perform a behavioral task, as described in the following sections. During each of the three attention conditions, SFOAEs were measured at one frequency to investigate task-dependent changes in SFOAE magnitude. Baseline measurements of the SFOAE were made at the beginning of each trial to ensure that a robust SFOAE with a signal-to-noise ratio (SNR) ≥ 20 dB (Goodman, 2013) was evoked by the tone used to monitor attention effects and that no systematic changes in the

baseline SFOAE occurred across attention conditions. The three attention conditions were each tested once to monitor the SFOAE at the lower frequency and were tested again to monitor the SFOAE at higher frequency. The effects of attention were examined both within and across listeners by comparing the levels of the SFOAE evoked by the same stimuli across the different attention conditions. In this way it was possible to assess the effects of attention, as well as the effects of probe frequency, with the working hypothesis that smaller SFOAE will be observed for attention directed away from the probe (V and auditory for the non-probe frequency) and larger SFOAE for attention directed to the SFOAE evoking probe frequency.

C. Participants

A total of 39 NH participants aged (19-61) were recruited to participate this experiment. Participants were required to pass both an SFOAE screening procedure and training on the behavioral task to be included in the experiment. After screening, a total of 30 participants (7 male, 23 female, aged 19-35, mean age: 23.1) remained in the participant sample. All the participants who failed screening were rejected due to an inability to measure robust SFOAEs in either the low-frequency (.75 kHz) region (3 participants) or high-frequency (4 kHz) region (6 participants). All participants in this study had normal hearing, defined by air conduction thresholds < 20 dB hearing level (HL) at octave frequencies between 0.25- and 8 kHz and measured using a calibrated Madsen Conera audiometer (GN Otometrics, Schaumberg, IL). All participants provided written informed consent before participating in the study and were either monetarily

compensated for their time or received course credit. All experimental procedures were approved by the Institutional Review Board at the University of Minnesota.

D. Calibration

The SFOAE-evoking stimuli used in both the screening procedure and the experimental tasks were presented using 40 dB forward pressure level (FPL), individually calibrated in each participant's ear. The suppressor stimulus used in the baseline measurement of SFOAE was presented at 60 dB FPL. Forward pressure calibration was used to moderate effects of standing waves on in-the-ear calibration. Standing waves can cause large variations in the sound level reaching the cochlea, depending on the insertion depth of the probe and the size of individual participants ear canals, especially at the higher nominal frequency of 4 kHz (Neely and Gorga, 1998; Scheperle et al., 2008, 2011; Siegel, 1994). In order to decompose the total recorded pressure from the microphone into forward and reverse pressure, we first estimated the Thévenin-equivalent source characteristics using a calibration tool specifically designed for use with the ER10X system ("ER10X Stepper", available from: http://audres.org/cel/thev/). Once source characteristics were estimated, in-ear measurements were made to calculate forward pressure using version 3.32 of the EMAV software package (Neely and Liu, 1994; available from: http://audres.org/rc/emav). In-ear calibration was performed at the beginning of every session, as well as any time participants removed the acoustic probe assembly from their ear.

E. Auditory Stimuli

The auditory stimuli were used both for the behavioral auditory task and for evoking the SFOAE used to assess the cochlear responses during all three selective-attention conditions (AL, AH, and V). Auditory stimuli were grouped into runs consisting of a baseline SFOAE measurement and a series of stimuli used for the behavioral task. The baseline SFOAE measurement was obtained using a suppression method (Shera and Guinan, 1999). A 10-s probe tone was presented at 40 dB FPL with a suppressor tone 50 Hz above the probe tone. The suppressor tone was gated on and off every 500 ms, ending at the same time as the probe tone. The frequencies of the probe and suppressor tones were near either .75 kHz or 4 kHz, depending on which SFOAE frequency was used to monitor selective-attention effects.

The task-relevant stimuli began after a 2-s silent gap that followed the offset of the probe tone used in the baseline measurement. Task-relevant stimuli consisted of a block of repeating sounds. Each block began with a 1-s cue tone presented at 40 dB FPL at the frequency (nominally 750 or 4000 Hz) that listeners were instructed to attend. A notched-noise MOCR elicitor began 500 ms after the onset of the cue tone and was presented for 1.7 seconds with a root mean square (RMS) level of 60 dB sound pressure level (SPL). This noise spanned from 0.1 to 10 kHz and included two 1-octave notches centered at 0.75- and 4 kHz. Following the offset of the cue tone, a total of four 300-ms tone pips (two at each frequency) were presented at 40 dB FPL. The tones were gated with 5-ms raised-cosine ramps to avoid spectral splatter. The two tones presented at each frequency had random inter-stimulus intervals (ISIs). These random ISIs were used to make the timing of the stimuli less predictable so that listeners would need to attend

vigilantly to the sequence of tones to detect the target stimuli. The ISIs were constrained

to have a minimum duration of 100 ms and the offset of the final tone in the sequence

occurred no later than the end of the noise. Each block ended with a 2-s silent interval to

allow for the decay of MOCR effects produced by the noise elicitor. A total of four

blocks were presented after each baseline measurement, resulting in a total of eight tone

pips at each frequency. Up to three of these eight tone pips at each frequency were

amplitude modulated at 10 Hz with 100% modulation depth. The selection of which tones

were amplitude modulated was constrained so that only one of the two tones at each

frequency could be modulated within a block. This constraint was imposed so that each

block contained an unmodulated tone pip which was used to measure the SFOAE at the

test frequency. A schematic representation of the auditory stimuli is shown Fig. 4.1.

All auditory stimuli were generated digitally with a sampling frequency of 48 kHz

with 24-bit resolution using custom MATLAB software (The Mathworks, Natick, MA),

converted to analog signals using a RME Fireface UC sound card (Haimhausen,

Germany), and delivered to participants' left ears via transducers in the ER10X acoustic

probe system (Etymotic Research, Elk Grove, IL). Playback and recording control were

performed using the Psychophysics toolbox (Brainard, 1997) in MATLAB. Recordings

were made using the microphones of the ER10X system, which provided 20 dB of gain to

the signal before it was digitized by the RME Fireface UC sound card and saved on the
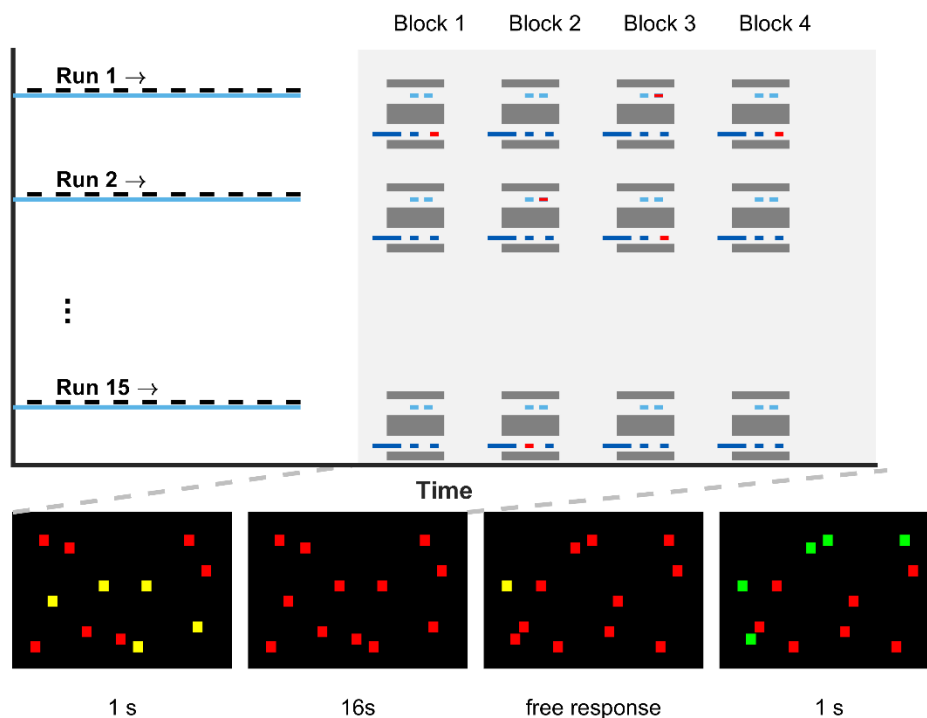
computer hard drive for offline analysis.

Figure 4.1

Schematic representation both auditory and visual stimuli. The top portion of the figure shows time frequency representations of the auditory stimuli. The SFOAE probe is shown in dark blue. Suppressor presentations shown in black are in alternating polarity. MOC elicitor noise is illustrated by the shaded gray regions. Tones with AM are illustrated in red. The inset at the bottom shows the timing for the visual stimuli.

F. Visual Stimuli

The visual stimuli were identical to those described in Chapter 3. The visual

stimuli began concurrently with the sequences of short tones used for the auditory

behavioral task. Sample frames depicting the stimuli during various stages of the MOT

task are shown at the bottom of Fig 4.1.

G. Procedure

All the procedures were conducted in a double-walled sound attenuating chamber (Industrial Acoustic Company, Bronx, NY) with the participants seated in a semi-reclined position in a chair with a head rest to minimize motion-related artifacts. Participants kept a computer keyboard on their lap to make responses during the auditory attention trials. A mouse was used to record responses to the visual stimuli after the end of stimulus presentation. Participants were instructed to remain as still and relaxed as possible during each run without closing their eyes.

Before testing began, participants completed a two-part screening procedure to ensure that: 1) any spontaneous otoacoustic emissions (SOAEs) were at least 100 Hz away from the experimental frequency range, and 2) high SNR measurements of SFOAE were obtainable within the frequency range of the experimental stimuli. These screening procedures are described in detail in Beim et al. (2018). The frequency of the low-frequency tone used for the SFOAE probe was selected as the frequency between 675-825 Hz that produced the largest SFOAE with an SNR greater than 25 dB. The frequency of the high-frequency tone was selected from the range of 3.6-4 kHz using the same criterion. Participants passed this screening procedure if at least one valid probe frequency was found in both the low- and high-frequency ranges.

After passing screening, participants completed 15 training runs of the experiment for each of the three attentional conditions. Training runs differed from experimental runs in two key ways: 1) AM was never imposed on the distractor frequency, and 2) the MOCR noise elicitor was not presented. Participants had to respond to the presence of AM at the target frequency within 1 s from the offset of AM for their response to be valid

and needed to achieve at least 75% correct detection of the AM targets across the 15 runs

for each condition to pass the training. The screening and training procedures were

typically completed over the course of one 2-h session.

In a second session, after completing training, participants completed the

experimental runs under a total of six conditions (3 attention conditions x 2 SFOAE

probe frequency conditions). Participants were instructed to attend the low-frequency

tone (AL), the high-frequency tone (AH) or the visual stimuli (V) and to perform the

relevant task for each stimulus (detecting AM for the auditory conditions and performing

MOT for the visual stimuli). The experimental runs were grouped by condition and

participants completed 15 runs for each condition. The order in which conditions were

completed was counterbalanced across participants to attempt to control for any order

effects or longer-term build-up effects. Participants completed all six attention conditions

within a single experimental session, during which the probe assembly was not removed

from their ears.

While participants were completing the experimental runs, a second ER10X probe

was used to record background noise in the contralateral ear. The contralateral recordings

were made during half of the total runs (either during the low- or high-frequency SFOAE

measurements) for each participant, selected randomly.

H. SFOAE Analysis

Before extracting emissions, recordings were manually scanned for noise and

movement-related artifacts. Recorded runs that contained visible artifacts were excluded
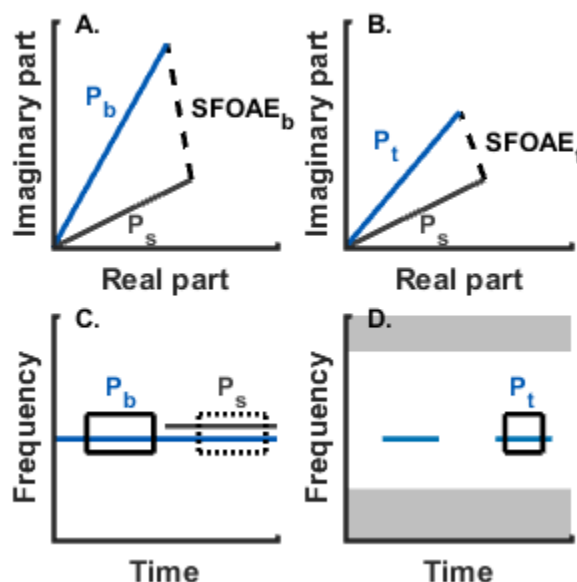
Figure 4.2

Phasor diagrams and time-frequency schematic illustrations depicting the emission extraction procedure from the averaged temporal waveform of recorded audio. A). A phasor diagram depicting the extraction of baseline SFOAE as the resultant of vector subtraction of baseline pressure Pb and suppressed baseline Ps. B). Phasor diagram depicting the extraction of SFOAE from the task relevant stimuli as the vector difference between pressure during the behavioral task Pt and Ps. Panels C and D are time-frequency schematics of the stimuli used to extract SFOAEs depicted in panels A and B, respectively. Shaded grey regions represent the notched noise elicitor. Dashed boxes highlight analyses windows where averaging occurs to perform the vector subtraction.

from analysis. Additional recordings were made if the magnitude of SFOAE in the

baseline measurement was less than 25 dB above the noise floor. On average 12

recordings were left after exclusion, yielding 120 probe-suppressor pairs for the baseline

OAE measurement. The baseline measurement was split into 1-s segments that contained

500 ms of the SFOAE probe alone and 500 ms of the probe plus the suppressor. The tone

pips presented during the behavioral task that did not contain amplitude modulation were

pooled across blocks and runs for averaging. The stimuli for both the baseline SFOAE

measurement and the task stimuli were averaged across the accepted runs. A heterodyne

procedure was then used to extract the sound pressure at the stimulus frequency (Guinan

et al., 2003). To estimate SFOAE magnitude the baseline measurement was extracted

using the suppression technique (Shera and Guinan, 1999). The suppression technique

uses the vector difference of the ear-canal sound pressure recorded with SFOAE probe

tone alone and the sound pressure recorded with the SFOAE probe tone presented with

the suppressor to estimate the SFOAE residual. A phasor diagram depicting this vector

subtraction is shown in Fig 4.2. The mean complex-valued sound pressure taken from a

300-ms window temporally centered on the duration of the suppressor tone was

subtracted from the heterodyned waveform. This yields the magnitude and phase of the
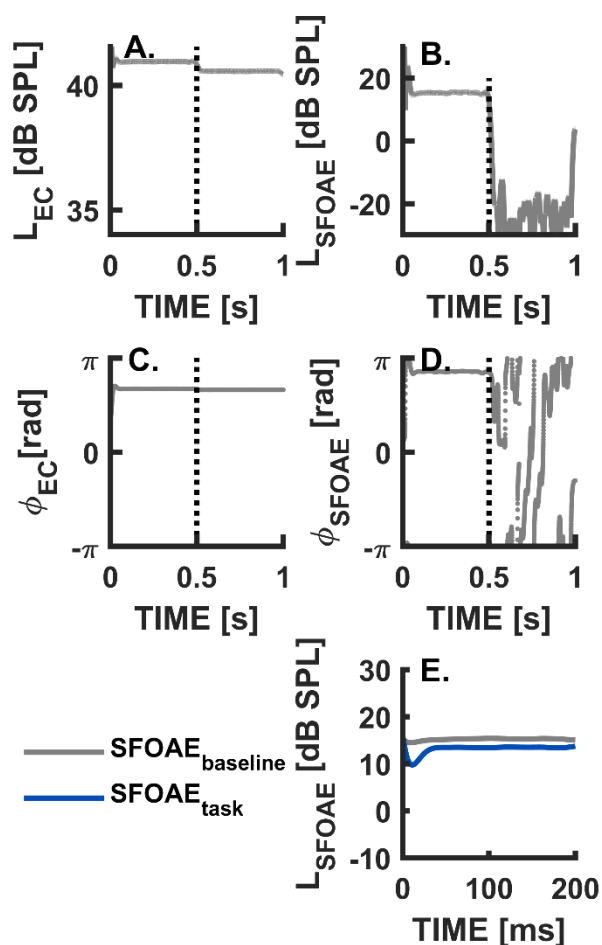

Figure 4.3

A representative result of the SFOAE extraction procedure with a 4 kHz probe tone. Panels A-D show the result of the extraction procedure from data recorded during the baseline measurement. Stimulus pressure recorded in the ear canal is decomposed into magnitude and phase in the left panels, while the extracted SFOAE is shown on the right. Panel E compares the emission magnitude evoked by the behavioral task stimulus (in blue) with the baseline SFOAE magnitude shown in grey (baseline replotted from panel B).

SFOAE residual during the first 500 ms of the waveform and the noise floor during the

final 500 ms, as shown in the panels A-D of Fig 4.3. Next, the same mean sound pressure

was subtracted from the segments corresponding to the behavioral task stimuli

(unmodulated tones), yielding the SFOAE residual magnitude and phase. For each

participant the mean emission magnitude was calculated from a 200-ms window centered

temporally in the task tone pip waveform. The SFOAE evoked during this window is

shown in panel E of Fig. 4.3 alongside the baseline SFOAE measurement. Note that the

difference between the baseline (grey) and task (blue) emissions reflects both the

contribution of our MOC elicitor as well as any effect of selective attention.

I. Bootstrap analysis procedure

To determine whether differences in SFOAE level across attentional conditions

were significant within an individual, a bootstrapping procedure was used to estimate

variability in the extracted SFOAE magnitudes. Artifact-free pairs of recorded segments

for both the baseline measurement and behavioral task were resampled randomly with

replacement, such that the total number of resampled segments was the same as the

original number of segments for that participant. The SFOAE extraction procedure

described previously was then repeated on each resampled set of recorded data and the

mean emission magnitudes were saved. This resampling procedure was repeated 10,000

times for each participant to construct estimated distributions of emission magnitude

across the attentional conditions. Differences between distributions were considered

significant when the mean magnitude of one distribution fell outside the 95% confidence interval of the second distribution.

III. Results

A. Behavioral Results

Hit and false alarm rates were used to calculate participants' sensitivity, d', to the presence of AM on the attended carriers (Green and Swets, 1966). We did not estimate d' for the visual task because participants were required to make 5 responses corresponding to 5 targets, meaning that misses and false alarms would not be independent of one another. Instead performance on the visual task was scored as percentage of correctly identified targets.

A paired samples t-test was used to compare performance between the two auditory attention conditions. There was a significant main effect of attended frequency [t(59) = 3.39, p = .001] indicating that performance in the AH condition (mean: 2.23) was better than the AL condition (mean: 1.50). Performance in the visual condition was similar across the two sets of recorded SFOAE frequencies with participants correctly identifying 73% of targets on average.
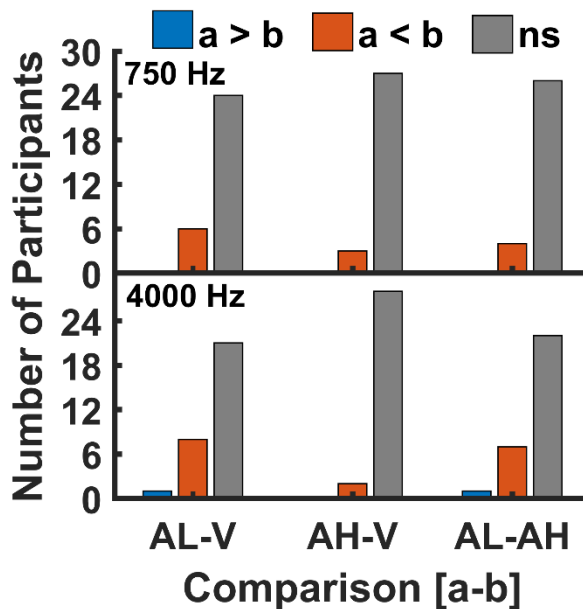
B. SFOAE Results

Figure 4.4

Each set of three bars shows the number of participants exhibiting each possible directional relationship between SFOAE magnitudes for the 3 attentional condition comparisons of interest. Grey bars show proportion of cases for each paired comparison where there was no significant difference. Blue bars show number of participants where the first condition produced larger SFOAE magnitudes than the second (e.g. in the first set of bars showing the AL-V comparison, the blue bar indicates the number of participants where SFOAEs were significantly larger in the AL condition). Orange bars show a significant effect in the opposite direction of the blue bars. The top panel shows the results using the low-frequency SFOAE probe, while the bottom panel shows results for the high-frequency probe tone.

The bootstrap analysis revealed that most participants failed to exhibit significant differences in SFOAE magnitude across attentional conditions at either probe frequency. The number of participants who did not exhibit significant shifts in emission magnitude between the three paired attention comparisons is shown by grey bars in Fig. 4.4. Some participants exhibited significant shifts in SFOAE magnitude across attentional conditions, but the direction of the shifts was not consistent across these individuals (blue bars represent cases in which the minuend was significantly larger than subtrahend in comparisons shown by labels on x axis and orange bars represent the opposite effect).

These within-subjects analyses do not support the hypothesis that a higher SFOAE

frequency might better reveal effects of selective attention because the pattern of results

and the total number of subjects exhibiting significant effects is similar across the two

probe frequencies. The bootstrap analysis also failed to provide evidence of any

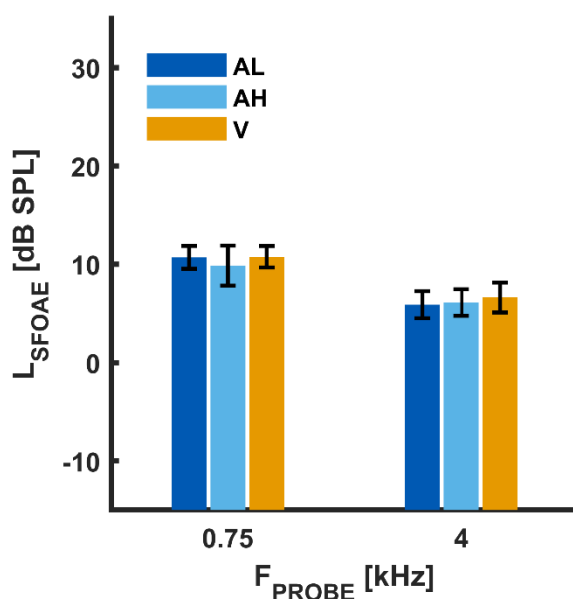systematic shift in SFOAE levels based on attention condition.

Figure 4.5

Mean data (n = 30) for SFOAE magnitudes extracted from the task stimuli. Sets of three bars are grouped by probe frequency. SFOAEs evoked during the auditory attention conditions are shown in shades of blue, while SFOAEs from the visual attention condition are shown in yellow. Error bars denote standard error of the mean.

Next, we examined attentional effects on SFOAE at a group level. Mean values of

SFOAE magnitude during each attention condition are plotted separately by SFOAE

frequency in Fig 4.5. A repeated-measures ANOVA with factors of probe frequency and

attention condition was conducted on the mean SFOAE magnitude values. The ANOVA

revealed a significant main effect of probe frequency on SFOAE magnitudes, indicating

that SFOAE magnitudes were greater at lower frequencies [$F(1,29) = 8.78$, $p = .006$], as

expected. However, there was no significant effect of attention [F(2,58) = .361, p = 699] and no significant interaction between the factors [F(1.35,39.27) = .18, p = .835].

C. Relationship between behavioral and SFOAE data
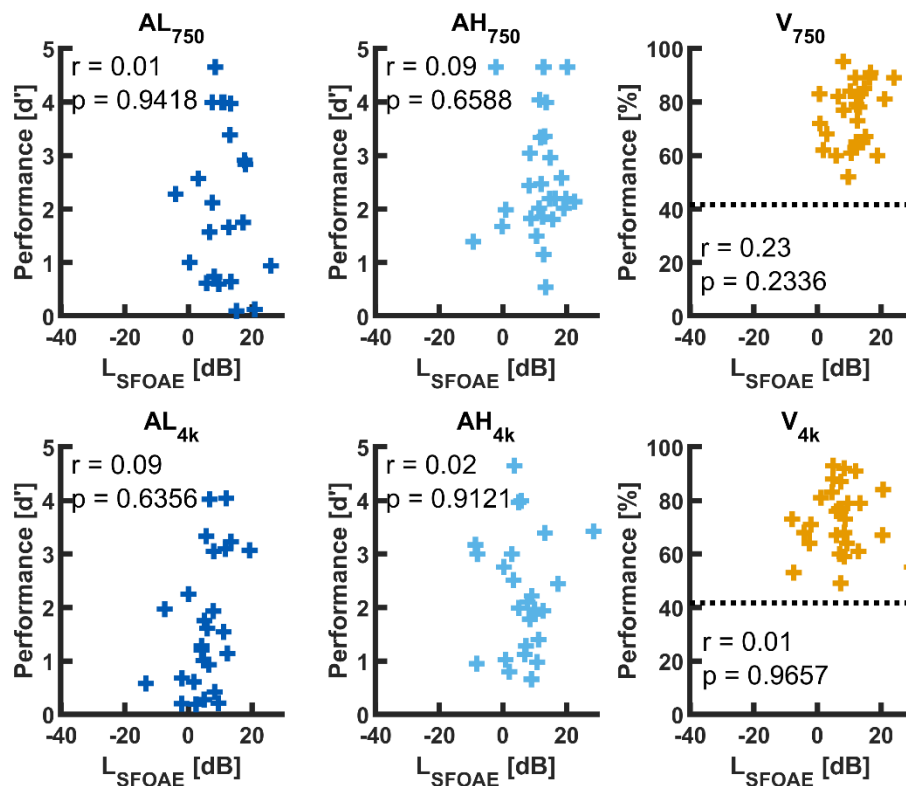
Figure 4.6

Results of correlational analyses between task performance and SFOAE magnitude. Trend lines show linear fits to the data. Each attention condition is plotted on separate axes. Top panels show magnitudes of SFOAE evoked by the 750 Hz tone while bottom panels show the 4 kHz results.

To relate SFOAE data to behavioral performance we correlated individual performance on each attention condition with the SFOAE magnitude evoked while participants were attending the target stimulus. If MOCR activity is linked to behavioral performance on these tasks, then we might expect SFOAEs to vary with individual listener performance. Measures of behavioral performance plotted against SFOAE

magnitudes are shown in Fig. 4.6, for each attention condition. As seen in the figure there is no systematic relationship between performance measures and the corresponding SFOAE measures at either stimulus frequency. Correlational analysis revealed no significant correlations between SFOAE magnitudes and behavioral task performance for the three attentional conditions and both probe frequencies [.01 < r(29) < .23, p > .234].

D. Noise measurements

Analysis of noise recorded in ear contralateral to stimuli presentation was conducted to examine whether changes in SFOAE level were related to evidence of participant-generated noise that was not detected as an artifact but might have contaminated the recordings. To analyze the noise, recorded samples from the unstimulated ear were obtained from the same time intervals as the SFOAE measurements, so that the noise levels could be temporally linked to their corresponding SFOAE measurements in the stimulated ear. A high-pass filter with a 250 Hz cutoff was used to eliminate low-frequency noise that was present within the sound attenuating chamber. The noise level was calculated from the RMS amplitude of the filtered noise waveform within 200-ms windows at the same times used to analyze the SFOAE stimuli. Mean values for the noise levels in the ear canal of participants measured during the behavioral task are plotted in Fig. 4.7. A repeated-measures ANOVA with within-subjects factors of task (whether the noise sample was from the baseline measurement or during the behavioral task) and attention condition was conducted on the noise values. The ANOVA revealed no significant effect of task [F(1,28) = 3.95, p = .057] indicating

that noise levels did not change between the baseline measurement of SFOAEs and the behavioral task portion of each trial. The ANOVA also revealed no significant effect of attention [$F(2,56) = .775$, $p = .466$]. There were no significant interactions between any of the factors [$F(2,56) = .776$, $p = .465$].

We also conducted a correlational analysis between noise recorded in the ears of participants and the SFOAEs extracted from the same time segments. The analysis revealed no significant correlation between noise levels in the unstimulated ear and SFOAE magnitudes in any attention condition [$-.19 < r(14) < .25$, $p > .419$] or pooled across attention conditions [$.065 < r(44) < .095$, $p > .565$].
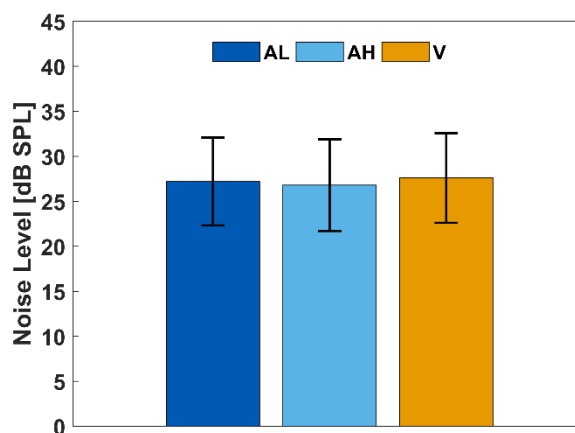
Figure 4.7

Mean contralateral-ear noise magnitudes across participants for each of the experimental conditions. Magnitudes are estimated from noise sampled during the same time periods as the task SFOAE. Colored bars denote the noise level measured in the ear canal for each attention condition. Error bars represent standard deviations.

IV. Discussion

This experiment was designed to improve the ability to observe attentional modulation of MOC efferent effects, relative to earlier studies in two important ways: 1) it used a high-frequency probe tone, which should be processed in a cochlear region with

higher gain and a possibly higher density of MOC efferent fiber innervation, making it more likely to observed efferent-induced changes in gain; and 2) it monitored the level of noise in the unstimulated contralateral ear to provide evidence of whether or not any changes seen in SFOAEs might be related to changes in participant motion. Our individual analyses revealed that statistically significant shifts of SFOAE magnitude occur with roughly equal frequency when measuring SFOAEs using low- (approximately 750 Hz) or high- (approximately 4000 Hz) frequency probe tones. A lack of improvement using high-frequency SFOAE probe tones may not be entirely surprising, as there is some human evidence indicating the MOC efferent effects are largest at frequencies around 0.5 and 1 kHz, and smaller near 4 kHz (Lilaonitkul and Guinan, 2012). In both cases we found no evidence of a relationship between participants behavioral performance on the task and the magnitudes of SFOAEs evoked by the task stimuli. We also used performance data to analyze SFOAEs only from correct trials (not shown), but this did not reveal a different pattern of results than SFOAEs obtained from both correct and incorrect trials.

Like the results discussed in Chapter 3, the results of the current study disagree with recent work by Walsh et al. (2014a, 2014b, 2015). There are important differences between our study and this previous work that could explain the difference. As discussed in chapter 3 the Walsh et al. reports compare results between attention and inattention. This results in a comparison between conditions that vary in task demands, potentially producing differences in general arousal. In contrast, our current study compared only conditions in which participants were engaged in a task, which would not be expected to produce as strong differences in arousal as the conditions in the Walsh et al. studies.

Along the same lines, Francis et al. (2018) suggested that the differences in physiological noise in the ear canal observed by Walsh et al. (2014a, 2014b) may be due to differences in participant movement between conditions.

Our current study showed no differences between attention conditions in physiological noise, which could indicate that any subtle participant movements were similar across all three attention conditions, as might be expected given similar levels of arousal. We also saw no differences in physiological noise between our baseline measurement of SFOAE (i.e. when there was no task for our participants to complete) and our task measurements, a comparison similar to the paradigm of Francis et al. (2018). A key difference between studies that could explain this difference is that Francis et al. (2018) collected the inattention data before the behavioral task was introduced and explained to participants, while in the current study the participants were always aware of the upcoming behavioral task and were likely anticipating the onset of the task stimuli during the measurement of the pre-task baseline.

Our study is limited in its ability to assess individual differences that could be related to our task as the samples were largely homogeneous (i.e. the vast majority of participants were young, female, normally hearing university students). Future work designed to investigate individual differences may be able to shed light on the sources of variability observed in the current dataset.

In conclusion, this study provides additional evidence that attentional modulation of cochlear function is not reliably detectable using SFOAEs, even in cochlear regions with larger gain and higher potential dynamic range for MOC effects. Combined with the

data presented in chapters 2 and 3 theses data provide a strong challenge to the existing

literature that suggests the MOCR is involved in the selective attention.

Chapter 5: Summary and conclusions

I. Summary

The work reported in this thesis sought to answer a fundamental question about the role of the medial olivocochlear reflex (MOCR) in the cognitive process of selective attention. Understanding a functional role for the MOCR in selective attention is important because this reflex is part of the auditory system that is impaired in sensorineural hearing loss. Difficulty in selectively attending to a target speaker in noisy backgrounds is a chief complaint among hearing aid users; understanding a potential method by which the normal auditory system aids in this process could help to inform our design of hearing assistive devices and computer speech understanding technologies.

Many previous studies using various methodologies have attempted to provide evidence of whether changes in cochlear function are linked to changes in selective attention. One study using electrophysiological evidence in animals provided strong evidence for attentional modulation of cochlear function by reporting changes in the compound action potential of the auditory nerve. (Delano et al., 2007) Several studies in humans have used otoacoustic emissions (OAEs) to assess cochlear function in non-invasive way also provide some evidence that attentional manipulations appear to produce modulations in cochlear function (Maison et al., 2001; Srinivasan et al., 2012, 2014; Walsh et al., 2015). A difficulty in interpreting the results from these studies is that the size and direction of MOCR effects has not been consistent. For example, (Srinivasan et al., 2012, 2014)report smallest distortion product OAE (DPOAE) magnitudes under auditory attention, indicating larger MOCR effects, but other studies report smaller MOCR effects (as shown by larger OAE magnitudes) under auditory attention conditions

(de Boer and Thornton, 2007; Harkrider and Bowers, 2009). Part of the differences between studies can be explained by differing stimulus design and which type of OAEs were measured. While small effect sizes seen in OAE changes may underlie larger perceptually relevant changes in cochlear function (Puria et al., 1996), the inconsistent directionality of effects and the fact that a number of studies fail to demonstrate any attentional modulation of cochlear function (Avan and Bonfils, 1992; Michie et al., 1996) make it difficult to draw any definitive conclusions from the existing body of work on this topic.

The work conducted as part of this thesis was designed to address the shortcomings of the previous literature in several ways. The use of stimulus-frequency OAEs (SFOAEs) provides a better measure of cochlear function than those used in other studies: DPOAEs (distortion and reflection components can interact), click-evoked OAES (CEOAEs; click trains can be MOC elicitors on their own), and the non-linear SFOAE (any effects that change linearly will also cancel in analysis). Additionally, all of the experiments presented here avoid comparing attention to inattention, as factors like general arousal may differ between the task and no-task conditions used in previous studies (e.g. s. In Chapter 2 we addressed these short comings and in our first sample of participants we found large effects of attention in a majority of participants individually and significant effects at the group level. The effects seen in this first sample were larger than any of those previously reported in the literature. Because of the striking difference between these results and the existing findings, we repeated our experiment on a second sample of participants and were unable to replicate results. In addition, a major limitation

in our first study was the inability to directly relate performance data to MOCR effects

because behavioral responses.

Chapter 3 addressed our ability to relate behavioral performance on the attention

tasks to our estimates of change in cochlear function and addressed the question of

whether perceptual load and sustained attention are important factors when looking for

cochlear effects of attention. Perceptual load theory (Lavie, 1995, 2005) states that

attentional resources are consumed obligatorily until the demands of a task exceed the

attentional capacity of the system. When this occurs, selective attention is expected to

operate early, limiting the processing of unattended stimuli to a minimum. The MOCR is

uniquely suited to this type of modification, as it modulates the gain of the peripheral

auditory system, aiding in the attenuation of distracting information. By requiring

sustained attention, the second experiment in Chapter 3 minimized opportunity for lapses

in attention. We tested participants both with and without an MOCR elicitor noise to

address how acoustically eliciting the MOCR system influences any potential attentional

effect. The data obtained from the both experiments in Chapter 3 failed to provide

evidence for a role of the MOCR in selective attention. We observed the expected effect

of our MOCR elicitor producing a reduction in SFOAE magnitudes, but this effect was

the same across all attention conditions. Our high perceptual load visual task produced

similar results to the low perceptual load working memory task in Chapter 2. There were

no detectable differences in the patterns of data obtained using the continuous stimuli,

suggesting that sustained attention did not improve our ability to detect an effect. By

collecting behavioral responses immediately after target presentation, we able to more

accurately relate behavioral performance to our SFOAE measures, but still did not find

significant linear relationships between the individual performance and SFOAE

magnitude across any of the attention conditions.

Chapter 4 asked whether or not the dynamic range of cochlear gain is important in

measuring MOCR effects under selective attention. By using a higher-frequency SFOAE

probe stimulus we measured cochlear function in a region where cochlear gain is greater,

improving our ability to detect MOCR effects by decreasing the probability that floor or

ceiling effects of MOCR activation would limit the ability to detect a relationship across

attentional conditions. Additionally, this study measured physiological noise in the ear

canal during the task performance. Attentional effects have been reported in

physiological noise measurements in unstimulated ears while participants performed

behavioral tasks. The changes in noise could reflect an MOCR effect, or they could be

related to participant motion, as suggested by Francis et al. (2018). Our data showed no

evidence for attentional effects when measuring cochlear function using low- nor high-

frequency stimuli. We also found no difference in noise measurements across attention

conditions. The lack of a noise effect suggests that participant motion was similar

between attention conditions, unlike the pattern of data observed in Francis et al. (2018).

In Francis et al. (2018) the authors saw a difference in noise recorded in the opposite ear

depending on when participants were passively sitting or engaged in an experimental

task. We compared noise levels during the baseline measurement of SFOAEs (before

behavioral task) and found that in addition to no significant effect of attentional

condition, noise levels also did not significantly differ between the baseline (passive) and

attentional task (active) conditions. Overall, therefore, using larger samples of subjects

than have been tested in most previous studies, and using more rigorous behavioral

techniques, we were not able to find any significant effects of selective attention on cochlear function.

II. Meta-analysis and Conclusions

Taken together, we have 120 sets of SFOAE measurements from 90 young normally hearing listeners using a (nominally) 750-Hz probe tone. In our initial subset of 15 participants we saw a significant effect of attention condition on SFOAE magnitudes, with visual attention (attention directed away from our SFOAE stimuli) producing the smallest SFOAE magnitudes. This pattern of data was not repeated across four further experiments. In each experiment there were participants for whom our bootstrapping procedure revealed significant shifts in SFOAE magnitude across attention conditions. In the first sample, a majority of participants exhibited significant shifts in SFOAE
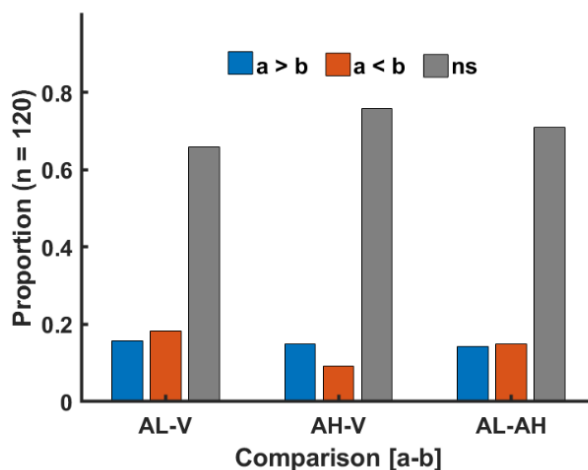
Figure 5.1

Grand summary of the bootstrap SFOAE magnitudes across all experimental conditions utilizing a (nominally) 750 Hz probe tone. Each set of bars shows the proportion of the sample that demonstrated each possible directional relationship between the attentional comparisons of interest. Grey bars show the proportion of cases for which there was no significant difference in SFOAE magnitudes between each condition. Blue and red bars indicate the proportion significant directional relationships in which either the minuend or the subtrahend respectively produced larger SFOAE magnitudes.

magnitude between the attend-low (AL) and visual (V) tasks. This was not true for the

subsequent samples, that only contained small numbers of participants with significant

shifts in SFOAE magnitude. A summary of the bootstrapping procedure results for all

participants tested with a 750 Hz tone is shown in Fig. 5.1.

As shown in the figure, there is no strong directionality of the significant effects

reported across the entire sample of participants. This makes it difficult to attribute the

shifts to any systematic change in MOCR activity that might be related to shifts in

attention.

In all experiments we tried to relate performance on the behavioral tasks to

SFOAE data to provide stronger evidence of attentional modulation than magnitude

differences across conditions. The experiments presented in chapter 2 were somewhat

limited in their ability to relate performance to SFOAEs because responses were made

after the presentation of several targets and relied on participants' memories. The

subsequent experiments in chapters 3 and 4 were conducted with a response window

immediately following presentation of the stimuli, but also failed to show any meaningful

relationship between performance on the behavioral tasks and SFOAE data. Another way

to demonstrate a relationship between MOCR activity and behavioral performance would

be to examine SFOAE magnitudes from trials with correct and incorrect responses

separately. This analysis was conducted using the data obtained from the experiment in

Ch. 4, and no significant difference was found in SFOAE magnitude between the correct

and incorrect trials.

The lack of a consistent directional relationship between SFOAE magnitudes

across attentional conditions, coupled with the lack of relationship between task

performance and SFOAE magnitudes, suggests that the MOCR does not operate as part of the system that produces large changes in the cortical representation of attended vs. unattended stimuli. This finding contradicts the results of many previous studies using similar methodologies (e.g. Maison et al., 2001; Srinivasan et al., 2012, 2014, Walsh et al., 2014b, 2014a, 2015).

In several of these cases (Maison et al., 2001; Walsh et al., 2014b, 2015) the authors compared attention to inattention to demonstrate their effects. While behavioral performance can be a good metric of attention to the task stimuli, there is no corresponding way to measure inattention. According to the perceptual load theory (Lavie, 1995; Lavie et al., 2004; Lavie and Tsal, 1994), attentional resources are obligatorily allocated to perception resulting in the influence of distractor stimuli (e.g. Eriksen and Eriksen, 1974; Stroop, 1935). This means that the OAE stimuli used in the studies by Walsh et al. (2014a, 2014b, 2015) and Maison et al. (2001) are likely still attended, and the actual difference between conditions in the case of Maison et al. (2001) is whether a task is performed or not.

In Walsh et al. (2014a, 2014b, 2015) the participants are still engaged in a task during the inattention condition so differences between the inattention and attention conditions could stem from differences between the tasks such as general arousal, task difficulty, or listening effort. The results from these studies and the differences between the experiments presented in this thesis suggest that the differences in measurements between attention and inattention conditions might simply reflect variability in the OAE measures themselves, or that MOCR activity can be modulated by the non-attentive differences between tasks (i.e. general arousal, difficulty).

The differences in attention effect seen between our first sample of participants and the subsequent experiments suggests that sufficiently large samples are required to ensure accurate measures of effect. The work presented in this thesis is the largest sample of data collected with the same experimental paradigm. The fact that the vast majority of this data (75 listeners measured at 750 Hz, 30 listeners at 4000 Hz) show no evidence of attentional modulation of SFOAE provides a strong motivation to interpret the conclusions of the existing body of work with skepticism.

In conclusion the experiments presented here challenge the conclusions of many recent studies suggesting a role of the MOCR in selective attention. A large sample of listeners completed tasks with rigorous manipulations of attention while cochlear function was estimated using state-of-the-art OAE measurement techniques. The results showed that both within and across listeners there were no systematic changes in SFOAE magnitude across conditions of auditory nor visual attention. The findings suggest that MOCR activity does not change depending on the target of attention, but do not rule out the possibility that differences in task demands or general arousal may produce changes in cochlear function. The results also highlight significant variability within individuals across attention condition that remains unexplained. Future work can more closely examine the individual differences with a large and diverse group of participants to ascertain whether they reveal any important factors influence MOCR activity.

III. Outlook

The data presented in this thesis provide evidence that MOCR activity is unlikely to be modulated by selective attending to one stimulus over another. This is true both

within the auditory modality and across the auditory and visual stimulus modalities. The lack of an observable effect of attention on cochlear responses was replicated in each of the experiments presented in this thesis. Many of the previous experiments examining effects of attention on cochlear responses used relatively small samples of participants (e.g. 8 listeners in each experiment in Walsh et al., 2015; 8 listeners in Srinivasan et al., 2012). These samples are more likely than larger samples to incorrectly estimate the effect of experimental manipulations. For example, the first set of data presented in chapter two contains nine participants that exhibited a significant shift in SFOAE magnitude across attention conditions. This was enough to produce a small but significant effect of attention in the entire sample of listeners in Ch. 2., but no significant effect when pooled with the additional data from chapters 3 and 4. The replicability of research is becoming an increasingly important discussion in the psychological sciences (Open Science Collaboration, 2015). A future study aimed at replicated the paradigms used by many of the recent studies on attention (e.g. Maison et al., 2001; Srinivasan et al., 2014; Walsh et al., 2015) could yield insights when considering the differences between the currently reported findings and the current work.

Another important question arising from the current set of experiments is whether or not individual differences can explain the variance in MOCR effects observed within individual listeners across our series of experiments. As discussed in each chapter, each experiment contained a small subset of participants for whom bootstrapping revealed significant shifts in SFOAE magnitude across attention conditions. While we did not find evidence of a link to these participants behavioral performance, there are many other auditory and cognitive factors such as hearing threshold, age, and measures of cognitive

function like intelligence or working memory function that could potentially relate to the differences observed in the SFOAE data. Even though the data presented here is a large sample, it comes from a largely homogeneous group of participants. These young, female, normally hearing listeners are not a diverse enough sample to accurately assess the potential individual differences underlying the data we have observed thus far. Future work could create a more diverse sample to determine whether any important individual differences might underlie the current pattern of results.

Lastly, one question that was not tested in the current set of experiments is whether attending to stimuli from one ear in a dichotic listening task might produce an MOCR effect. Although the MOCR is a bilateral system composed of ipsilateral and contralateral pathways (see Guinan, 2006) it is possible that directing attention to one ear may produce an attenuating effect in the opposite ear. While our current data suggest that this is unlikely to reveal a link between attention and MOCR function, testing this hypothesis will either provide stronger evidence to support the current findings or a reason to reconsider our current conclusion that there is no role for MOCR in selective attention.

References

Ahveninen, J., Jaaskelainen, I. P., Raij, T., Bonmassar, G., Devore, S., Hamalainen, M., Levanen, S., et al. (**2006**). "Task-modulated 'what' and 'where' pathways in human auditory cortex," Proc. Natl. Acad. Sci. U. S. A., **103**, 14608–14613. doi:10.1073/pnas.0510480103

Avan, P., and Bonfils, P. (**1992**). "Analysis of possible interactions of an attentional task with cochlear micromechanics," Hear. Res., **57**, 269–275. doi:10.1016/0378-5955(92)90156-H

Backus, B. C., and Guinan, J. J. (**2006**). "Time-course of the human medial olivocochlear reflex," J. Acoust. Soc. Am., **119**, 2889–2904. doi:10.1121/1.2169918

Baiduc, R. R., Lee, J., and Dhar, S. (**2014**). "Spontaneous otoacoustic emissions, threshold microstructure, and psychophysical tuning over a wide frequency range in humans," J. Acoust. Soc. Am., **135**, 300–14. doi:10.1121/1.4840775

Beim, J. A., Elliott, M., Oxenham, A. J., and Wojtczak, M. (**2015**). "Stimulus Frequency Otoacoustic Emissions Provide No Evidence for the Role of Efferents in the Enhancement Effect," J. Assoc. Res. Otolaryngol., **16**, 613–629. doi:10.1007/s10162-015-0534-8

Beim, J. A., Oxenham, A. J., and Wojtczak, M. (**2018**). "Examining replicability of an otoacoustic measure of cochlear function during selective attention," J. Acoust. Soc. Am., **144**, 2882–2895. doi:10.1121/1.5079311

Beim, J. A., and Wojtczak, M. (**2014**). "Effect of low-intensity highpass noise on SFOAE group delays for low frequency evoking tones," J. Acoust. Soc. Am., **137**, 2409. doi:10.1073/pnas.0703993104

de Boer, J., and Thornton, A. R. D. (**2007**). "Effect of subject task on contralateral suppression of click evoked otoacoustic emissions," Hear. Res., **233**, 117–123. doi:10.1016/j.heares.2007.08.002

Brainard, D. H. (**1997**). "The Psychophysics Toolbox," Spat. Vis., **10**, 433–436. doi:10.1163/156856897X00357

Bregman, A. S. (**1990**). *Auditory scene analysis: the perceptual organization of sound*, MIT Press, Cambridge, MA, MA.

Broadbent, D. E. (**1952**). "Failures of Attention in Selective Listening," J. Exp. Psychol., **44**, 428–433.

Broadbent, D. E. (**1958**). "Selective listening to speech," Percept. Commun., , doi: 10.1037/10037-002. doi:10.1037/10037-002

Brown, M. C. (**1989**). "Morphology and response properties of single olivocochlear fibers in the guinea pig," Hear. Res., **40**, 93–109. doi:10.1016/0378-5955(89)90103-2

Brown, M. C. (**2011**). "Anatomy of Olivocochlear Neurons," In D. K. Ryugo and R. R.

Fay (Eds.), Audit. Vestib. Efferents, Springer New York, New York, NY, pp. 17–37. doi:10.1007/978-1-4419-7070-1_2

Carlyon, R. P., Cusack, R., Foxton, J. M., and Robertson, I. H. (**2001**). "Effects of attention and unilateral neglect on auditory stream segregation," J. Exp. Psychol. Hum. Percept. Perform., **27**, 115–27. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11248927

Carlyon, R. P., Plack, C. J., Fantini, D. A., and Cusack, R. (**2003**). "Cross-Modal and Non-Sensory Influences on Auditory Streaming," Perception, **32**, 1393–1402. doi:10.1068/p5035

Charaziak, K. K., and Siegel, J. H. (**2015**). "Tuning of SFOAEs Evoked by Low-Frequency Tones Is Not Compatible with Localized Emission Generation," J. Assoc. Res. Otolaryngol., **16**, 317–329. doi:10.1007/s10162-015-0513-0

Cherry, E. C. (**1953**). "Some experiments on the recognition of speech, with one and with two ears," J. Acoust. Soc. Am., **25**, 975–979. doi:10.1121/1.1907229

Choi, I., Rajaram, S., Varghese, L. A., and Shinn-Cunningham, B. G. (**2013**). "Quantifying attentional modulation of auditory-evoked cortical responses from single-trial electroencephalography," Front. Hum. Neurosci., **7**, 115. doi:10.3389/fnhum.2013.00115

Delano, P. H., Elgueda, D., Hamame, C. M., and Robles, L. (**2007**). "Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas," J. Neurosci., **27**, 4146–4153. doi:10.1523/JNEUROSCI.3702-06.2007

Deutsch, J. A., and Deutsch, D. (**1963**). "Attention: Some theoretical considerations," Psychol. Rev., **70**, 51–60. doi:10.1037/h0042712

van Dijk, P., and Wit, H. P. (**1990**). "Synchronization of spontaneous otoacoustic emissions to a 2f1−f2 distortion product," J. Acoust. Soc. Am., **88**, 850–856. doi:10.1121/1.399734

Ding, N., and Simon, J. Z. (**2012**). "Neural coding of continuous speech in auditory cortex during monaural and dichotic listening," J. Neurophysiol., **107**, 78–89. doi:10.1152/jn.00297.2011

Dragicevic, C. D., Aedo, C., León, A., Bowen, M., Jara, N., Terreros, G., Robles, L., et al. (**2015**). "The Olivocochlear Reflex Strength and Cochlear Sensitivity are Independently Modulated by Auditory Cortex Microstimulation," J. Assoc. Res. Otolaryngol., **16**, 223–240. doi:10.1007/s10162-015-0509-9

Eriksen, B. A., and Eriksen, C. W. (**1974**). "Effects of noise letters upon the idetification of a target letter in a nonsearch task," Percept. Psychophys.,.

"Estimating the reproducibility of psychological science," (**2015**). Science (80-.)., **349**, aac4716-aac4716. doi:10.1126/science.aac4716

Ferber-Viart, C., Duclaux, R., Collet, L., and Guyonnard, F. (**1995**). "Influence of

auditory stimulation and visual attention on otoacoustic emissions," Physiol. Behav., **57**, 1075–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/7652027

Francis, N. A., Zhao, W., and Guinan Jr., J. J. (**2018**). "Auditory Attention Reduced Ear-Canal Noise in Humans by Reducing Subject Motion, Not by Medial Olivocochlear Efferent Inhibition: Implications for Measuring Otoacoustic Emissions During a Behavioral Task," Front. Syst. Neurosci., , doi: 10.3389/fnsys.2018.00042. doi:10.3389/fnsys.2018.00042

Fritz, J. B., Elhilali, M., David, S. V, and Shamma, S. A. (**2007**). "Auditory attention— focusing the searchlight on sound," Curr. Opin. Neurobiol., **17**, 437–455. doi:10.1016/j.conb.2007.07.011

Froehlich, P., Collet, L., Chanal, J.-M., and Morgon, A. (**1990**). "Variability of the influence of a visual task on the active micromechanical properties of the cochlea," Brain Res., **508**, 286–288. doi:10.1016/0006-8993(90)90408-4

Froehlich, P., Collet, L., and Morgon, A. (**1993**). "Transiently evoked otoacoustic emission amplitudes change with changes of directed attention," Physiol. Behav., **53**, 679–682. doi:10.1016/0031-9384(93)90173-D

Giard, M.-H., Collet, L., Bouchet, P., and Pernier, J. (**1994**). "Auditory selective attention in the human cochlea," Brain Res., **633**, 353–356. doi:10.1016/0006-8993(94)91561-X

Gifford, M. L., and Guinan, J. J. J. (**1987**). "Effects of electrical stimulation of medial olivocochlear neurons on ipsilateral and contralateral cochlear responses," Hear. Res., **29**, 179–94. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3624082

Gilbert, A. N., and Wysocki, C. J. (**1992**). "Hand preference and age in the United States," Neuropsychologia, **30**, 601–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1528408

Goodman, S. S., Mertes, I. B., Lewis, J. D., and Weissbeck, D. K. (**2013**). "Medial olivocochlear-induced transient-evoked otoacoustic emission amplitude shifts in individual subjects," J. Assoc. Res. Otolaryngol., **14**, 829–842. doi:10.1007/s10162-013-0409-9

Gorga, M. P., Neely, S. T., Dierking, D. M., Kopun, J., Jolkowski, K., Groenenboom, K., Tan, H., et al. (**2007**). "Low-frequency and high-frequency cochlear nonlinearity in humans," J. Acoust. Soc. Am., **122**, 1671–1680. doi:10.1121/1.2751265

Green, D., and Swets, J. A. (**1966**). *Signal detection theory and psychophysics*, Wiley, New York. Retrieved from https://www.worldcat.org/title/signal-detection-theory-and-psychophysics/oclc/890266

Guinan, J. J. (**2006**). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans," Ear Hear., **27**, 589–607. doi:10.1097/01.aud.0000240507.83072.e7

Guinan, J. J., Backus, B. C., Lilaonitkul, W., and Aharonson, V. (**2003**). "Medial

olivocochlear efferent reflex in humans: Otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs," JARO - J. Assoc. Res. Otolaryngol., **4**, 521–540. doi:10.1007/s10162-002-3037-3

Guinan, J. J. J. (**1996**). "Physiology of Olivocochlear Efferents," In P. Dallos, A. N. Popper, and R. R. Fay (Eds.), The Cochlea, Springer New York, p. 435 502.

Guinan, J. J. J. (**2010**). "Cochlear efferent innervation and function," Curr. Opin. Otolaryngol. Head Neck Surg., **18**, 447–453. doi:10.1097/MOO.0b013e32833e05d6

Guinan, J. J., Warr, W. B., and Norris, B. E. (**1984**). "Topographic organization of the olivocochlear projections from the lateral and medial zones of the superior olivary complex," J. Comp. Neurol., **226**, 21–7. doi:10.1002/cne.902260103

Guinan Jr., J. J. (**2014**). "Olivocochlear efferent function: issues regarding methods and the interpretation of results," Front. Syst. Neurosci., **8**, 1–5. doi:10.3389/fnsys.2014.00142

Gutschalk, A., Micheyl, C., and Oxenham, A. J. (**2008**). "Neural Correlates of Auditory Perceptual Awareness under Informational Masking," (T. D. Griffiths, Ed.) PLoS Biol., **6**, e138. doi:10.1371/journal.pbio.0060138

Hansen, R., Santurette, S., and Verhulst, S. (**2014**). "Effects of spontaneous otoacoustic emissions on pure-tone frequency difference limens," J. Acoust. Soc. Am., **136**, 3147–3158. doi:10.1121/1.4900597

Harkrider, A. W., and Bowers, C. D. (**2009**). "Evidence for a cortically mediated release from inhibition in the human cochlea," J. Am. Acad. Audiol., **20**, 208–215. doi:10.3766/jaaa.20.3.7

Hernández-Peón, R., Scherrer, H., and Jouvet, M. (**1956**). "Modification of Electric Activity in Cochlear Nucleus during 'Attention' in Unanesthetized Cats," Science (80-.)., **123**, 331–332. doi:10.1126/science.123.3191.331

James, W. (**1890**). *The principles of psychology*, Henry Holt and Company, New York, xviii,1393 pages.

Jennings, S. G., Heinz, M. G., and Strickland, E. A. (**2011**). "Evaluating Adaptation and Olivocochlear Efferent Feedback as Potential Explanations of Psychophysical Overshoot," J. Assoc. Res. Otolaryngol., **12**, 345–360. doi:10.1007/s10162-011-0256-5

Johnson, J. A., and Zatorre, R. J. (**2006**). "Neural substrates for dividing and focusing attention between simultaneous auditory and visual events," Neuroimage, **31**, 1673–81. doi:10.1016/j.neuroimage.2006.02.026

Kahneman, D., and Treisman, A. M. (**1984**). "Changing views of attention and automaticity," In R. Parasuraman and D. R. Davies (Eds.), Var. Atten., Academic Press, New York, pp. 29–61.

Kalluri, R., and Shera, C. A. (**2001**). "Distortion-product source unmixing: a test of the

two-mechanism model for DPOAE generation," J. Acoust. Soc. Am., **109**, 622–637. doi:10.1121/1.1334597

Kalluri, R., and Shera, C. a (**2007**). "Comparing stimulus-frequency otoacoustic emissions measured by compression, suppression, and spectral smoothing," J. Acoust. Soc. Am., **122**, 3562–75. doi:10.1121/1.2793604

Kawase, T., Delgutte, B., and Liberman, M. C. (**1993**). "Antimasking effects of the olivocochlear reflex II Enhancement of auditory-nerve response to masked tones," J. Neurophysiol., **70**, 2533–49. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8120597

Keefe, D. H. (**1998**). "Double-evoked otoacoustic emissions I Measurement theory and nonlinear coherence," J. Acoust. Soc. Am., **103**, 3489–3498. doi:10.1121/1.423057

Kemp, D. T. (**1978**). "Stimulated acoustic emissions from within the human auditory system," J. Acoust. Soc. Am., **64**, 1386–1391.

Kemp, D. T., and Chum, R. (**1980**). "Properties of the generator of stimulated acoustic emissions," Hear. Res., **2**, 213–232. doi:10.1016/0378-5955(80)90059-3

Khalfa, S., Veuillet, E., and Collet, L. (**1998**). "Influence of handedness on peripheral auditory asymmetry," Eur. J. Neurosci., **10**, 2731–2737. doi:10.1046/j.1460-9568.1998.00286.x

Knudsen, E. I. (**2007**). "Fundamental components of attention," Annu. Rev. Neurosci., **30**, 57–78. doi:10.1146/annurev.neuro.30.051606.094256

Kruschke, J. K. (**2013**). "Bayesian estimation supersedes the t test," J. Exp. Psychol. Gen., **142**, 573–603. doi:10.1037/a0029146

Lavie, N. (**1995**). "Perceptual load as a necessary condition for selective attention," J. Exp. Psychol. Hum. Percept. Perform., **21**, 451–468. doi:10.1037//0096-1523.21.3.451

Lavie, N. (**2005**). "Distracted and confused?: Selective attention under load," Trends Cogn. Sci., **9**, 75–82. doi:10.1016/j.tics.2004.12.004

Lavie, N., Hirst, A., de Fockert, J. W., and Viding, E. (**2004**). "Load theory of selective attention and cognitive control," J. Exp. Psychol. Gen., **133**, 339–354. doi:10.1037/0096-3445.133.3.339

Lavie, N., and Tsal, Y. (**1994**). "Perceptual load as a major determinant of the locus of selection in visual attention," Percept. Psychophys., **56**, 183–197. doi:10.3758/BF03213897

Liberman, M. C., and Brown, M. C. (**1986**). "Physiology and anatomy of single olivocochlear neurons in the cat," Hear. Res., **24**, 17–36. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/3759672

Liberman, M. C., Liberman, L. D., and Maison, S. F. (**2014**). "Efferent Feedback Slows

Cochlear Aging," J. Neurosci., **34**, 4599–4607. doi:10.1523/JNEUROSCI.4923-13.2014

Lilaonitkul, W., and Guinan, J. J. (**2012**). "Frequency tuning of medial-olivocochlear-efferent acoustic reflexes in humans as functions of probe frequency," J. Neurophysiol., **107**, 1598–1611. doi:10.1152/jn.00549.2011

Lilaonitkul, W., and Guinan, J. J. J. (**2009**). "Human medial olivocochlear reflex: effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths," JARO - J. Assoc. Res. Otolaryngol., **10**, 459–470. doi:10.1007/s10162-009-0163-1

Lukas, J. H. (**1980**). "Human Auditory Attention: The Olivocochlear Bundle May Function as a Peripheral Filter," Psychophysiology, **17**, 444–452. doi:10.1111/j.1469-8986.1980.tb00181.x

Maison, S. F., Adams, J. C., and Liberman, M. C. (**2003**). "Olivocochlear innervation in the mouse: immunocytochemical maps, crossed versus uncrossed contributions, and transmitter colocalization," J. Comp. Neurol., **455**, 406–16. doi:10.1002/cne.10490

Maison, S. F., and Liberman, M. C. (**2000**). "Predicting vulnerability to acoustic injury with a noninvasive assay of olivocochlear reflex strength," J. Neurosci., **20**, 4701–4707.

Maison, S. F., Usubuchi, H., and Liberman, M. C. (**2013**). "Efferent feedback minimizes cochlear neuropathy from moderate noise exposure," J. Neurosci., **33**, 5542–5552.

Maison, S., Micheyl, C., and Collet, L. (**1997**). "Medial olivocochlear efferent system in humans studied with amplitude-modulated tones," J. Neurophysiol., **77**, 1759–68. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9114234

Maison, S., Micheyl, C., and Collet, L. (**2001**). "Influence of focused auditory attention on cochlear activity in humans," Psychophysiology, **38**, 35–40. doi:10.1111/1469-8986.3810035

Makovski, T., and Jiang, Y. V (**2009**). "The role of visual working memory in attentive tracking of unique objects," J. Exp. Psychol. Hum. Percept. Perform., **35**, 1687–97. doi:10.1037/a0016453

McFadden, D., Pasanen, E. G., Leshikar, E. M., Hsieh, M. D., and Maloney, M. M. (**2012**). "Comparing behavioral and physiological measures of combination tones: sex and race differences," J. Acoust. Soc. Am., **132**, 968–83. doi:10.1121/1.4731224

McFadden, D., Pasanen, E. G., Maloney, M. M., Leshikar, E. M., and Pho, M. H. (**2018**). "Correlations between otoacoustic emissions and performance in common psychoacoustical tasks," J. Acoust. Soc. Am., **143**, 2355–2367. doi:10.1121/1.5030999

McFadden, D., Pasanen, E. G., Maloney, M. M., Leshikar, E. M., and Pho, M. H. (**2018**). "Differences in common psychoacoustical tasks by sex, menstrual cycle, and race," J. Acoust. Soc. Am., **143**, 2338–2354. doi:10.1121/1.5030998

Meric, C., and Collet, L. (**1992**). "Visual attention and evoked otoacoustic emissions: A slight but real effect," Int. J. Psychophysiol., **12**, 233–235. doi:10.1016/0167-8760(92)90061-F

Mesgarani, N., and Chang, E. F. (**2012**). "Selective cortical representation of attended speaker in multi-talker speech perception," Nature, **485**, 233–236. doi:10.1038/nature11020

Michie, P. T., LePage, E. L., Solowij, N., Haller, M., and Terry, L. (**1996**). "Evoked otoacoustic emissions and auditory selective attention," Hear. Res., **98**, 54–67. doi:10.1016/0378-5955(96)00059-7

Miller, J. (**1987**). "Priming is not necessary for selective-attention failures: Semantic effects of unattended, unprimed letters," Percept. Psychophys., **41**, 419–434.

Moleti, A., Al-Maamury, A. M., Bertaccini, D., Botti, T., and Sisto, R. (**2013**). "Generation place of the long- and short-latency components of transient-evoked otoacoustic emissions in a nonlinear cochlear model," J. Acoust. Soc. Am., **133**, 4098–108. doi:10.1121/1.4802940

Mukerji, S., Windsor, A. M., and Lee, D. J. (**2010**). "Auditory brainstem circuits that mediate the middle ear muscle reflex," Trends Amplif., **14**, 170–91. doi:10.1177/1084713810381771

Mulders, W. H., and Robertson, D. (**2000**). "Evidence for direct cortical innervation of medial olivocochlear neurones in rats," Hear. Res., **144**, 65–72. doi:10.1016/S0378-5955(00)00046-0

Müller, J., Janssen, T., Heppelmann, G., and Wagner, W. (**2005**). "Evidence for a bipolar change in distortion product otoacoustic emissions during contralateral acoustic stimulation in humans," J. Acoust. Soc. Am., **118**, 3747–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16419819

Neely, S. T., and Gorga, M. P. (**1998**). "Comparison between intensity and pressure as measures of sound level in the ear canal," J. Acoust. Soc. Am., **104**, 2925–34. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9821338

Neely, S. T., and Liu, Z. (**1994**). "EMAV: Otoacoustic emission average," Tech. Memo No. 17, Boys T. Natl. Res. Hosp. Omaha, NE.,.

Norman, D. A. (**1968**). "Toward a theory of memory and attention," Psychol. Rev., **75**, 522–536. doi:10.1037/h0026699

Norrix, L. W., and Glattke, T. J. (**1995**). "Distortion product otoacoustic emissions created through the interaction of a spontaneous otoacoustic emission and an externally generated tone," J. Acoust. Soc. Am., **97**, 3281.

Norton, S. J., Schmidt, A. R., and Stover, L. J. (**1990**). "Tinnitus and otoacoustic emissions: Is there a link?," Ear Hear., **11**, 159–166. doi:10.1097/00003446-199004000-00011

O'Connor, D. H., Fukui, M. M., Pinsk, M. A., and Kastner, S. (**2002**). "Attention modulates responses in the human lateral geniculate nucleus," Nat. Neurosci., **5**, 1203–1209. doi:10.1038/nn957

O'Sullivan, J. A., Power, A. J., Mesgarani, N., Rajaram, S., Foxe, J. J., Shinn-Cunningham, B. G., Slaney, M., et al. (**2015**). "Attentional selection in a cocktail party environment can be decoded from single-trial EEG," Cereb. Cortex, **25**, 1697–1706. doi:10.1093/cercor/bht355

Open Science Collaboration, O. S. (**2015**). "PSYCHOLOGY Estimating the reproducibility of psychological science," Science, **349**, aac4716. doi:10.1126/science.aac4716

Penner, M. J., Glotzbach, L., and Huang, T. (**1993**). "Spontaneous otoacoustic emissions: Measurement and data," Hear. Res., **68**, 229–237. doi:10.1016/0378-5955(93)90126-L

Penner, M. J., and Zhang, T. (**1997**). "Prevalence of spontaneous otoacoustic emissions in adults revisited," Hear. Res., **103**, 28–34. doi:10.1016/S0378-5955(96)00162-1

Picton, T. W., Hillyard, S. A., Galambos, R., and Schiff, M. (**1971**). "Human auditory attention: a central or peripheral process?," Science, **173**, 351–3. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/5559908

Picton, T. W., and Hillyard, S. a (**1974**). "Human auditory evoked potentials II: effects of attention," Electroencephalogr. Clin. Neurophysiol., **36**, 191–199. doi:10.1016/0013-4694(74)90156-4

Plack, C. J., Oxenham, A. J., Simonson, A. M., O'Hanlon, C. G., Drga, V., and Arifianto, D. (**2008**). "Estimates of compression at low and high frequencies using masking additivity in normal and impaired ears," J. Acoust. Soc. Am., **123**, 4321–30. doi:10.1121/1.2908297

Posner, M. I. (**1980**). "Orienting of attention," Q J Exp Psychol, **32**, 3–25. doi:10.1080/00335558008248231

Puel, J.-L., Bonfils, P., and Pujol, R. (**1988**). "Selective attention modifies the active micromechanical properties of the cochlea," Brain Res., **447**, 380–383. doi:10.1016/0006-8993(88)91144-4

Puria, S., Guinan, J. J. J., and Liberman, M. C. (**1996**). "Olivocochlear reflex assays: effects of contralateral sound on compound action potentials versus ear-canal distortion products," J. Acoust. Soc. Am., **99**, 500–507. doi:10.1121/1.414508

Pylyshyn, Z. W., and Storm, R. W. (**1988**). "Tracking multiple independent targets: evidence for a parallel tracking mechanism," Spat. Vis., **3**, 179–197. Retrieved from http://booksandjournals.brillonline.com/content/journals/10.1163/156856888x00122

Rajan, R. (**1988**). "Effect of electrical stimulation of the crossed olivocochlear bundle on temporary threshold shifts in auditory sensitivity I Dependence on electrical stimulation parameters," J. Neurophysiol., **60**, 549–68. Retrieved from

http://www.ncbi.nlm.nih.gov/pubmed/3171641

Rasmussen, G. L. (**1946**). "The olivary peduncle and other fiber projections of the superior olivary complex," J. Comp. Neurol., **84**, 141–219. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20982804

Rasmussen, G. L. (**1953**). "Further observations of the efferent cochlear bundle," J. Comp. Neurol., **99**, 61–74. doi:10.1002/cne.900990105

Recio-Spinoso, A., and Oghalai, J. S. (**2017**). "Mechanical tuning and amplification within the apex of the guinea pig cochlea," J. Physiol., **595**, 4549–4561. doi:10.1113/JP273881

Rees, G., Firth, C. D., and Lavie, N. (**1997**). "Modulating Irrelevant Motion Perception by Varying Attentional Load in an Unrelated Task," Science (80-.)., **278**, 1616–1619. doi:10.1126/science.278.5343.1616

Robinette, M. S., and Glattke, T. J. (**2007**). *Otoacoustic Emissions: Clinical Applications*, (B. Brandenburg, Ed.) Thieme Medical Publishers, New York, 3rd ed.

Robles, L., and Ruggero, M. A. (**2001**). "Mechanics of the mammalian cochlea," Physiol. Rev., **81**, 1305–52. doi:10.1152/physrev.2001.81.3.1305

Ruggero, M. A., Rich, N. C., and Freyman, R. (**1983**). "Spontaneous and impulsively evoked otoacoustic emissions: indicators of cochlear pathology?," Hear. Res., **10**, 283–300. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/6874602

Scharf, B., Magnan, J., and Chays, A. (**1997**). "On the role of the olivocochlear bundle in hearing: 16 case studies," Hear. Res., **103**, 101–122. doi:10.1016/S0378-5955(96)00168-2

Scharf, B., Magnan, J., Collet, L., Ulmer, E., and Chays, A. (**1994**). "On the role of the olivocochlear bundle in hearing: a case study," Hear. Res., **75**, 11–26. doi:10.1016/0378-5955(94)90051-5

Scheperle, R. A., Goodman, S. S., and Neely, S. T. (**2011**). "Further assessment of forward pressure level for *in situ* calibration," J. Acoust. Soc. Am., **130**, 3882–3892. doi:10.1121/1.3655878

Scheperle, R. A., Neely, S. T., Kopun, J. G., and Gorga, M. P. (**2008**). "Influence of in situ , sound-level calibration on distortion-product otoacoustic emission variability," J. Acoust. Soc. Am., **124**, 288–300. doi:10.1121/1.2931953

Schwartz, S., Vuilleumier, P., Hutton, C., Maravita, A., Dolan, R. J., and Driver, J. (**2005**). "Attentional Load and Sensory Competition in Human Vision: Modulation of fMRI Responses by Load at Fixation during Task-irrelevant Stimulation in the Peripheral Visual Field," Cereb. Cortex, **15**, 770–786. doi:10.1093/cercor/bhh178

Shahnaz, N. (**2008**). "Transient evoked otoacoustic emissions (TEOAEs) in Caucasian and Chinese young adults," Int. J. Audiol., **47**, 76–83. doi:10.1080/14992020701711029

Shamma, S. A., Elhilali, M., and Micheyl, C. (**2011**). "Temporal coherence and attention in auditory scene analysis," Trends Neurosci, **34**, 114–123. doi:10.1016/j.tins.2010.11.002

Shera, C. A. (**2003**). "Mammalian spontaneous otoacoustic emissions are amplitude-stabilized cochlear standing waves," J. Acoust. Soc. Am., **114**, 244–262. doi:10.1121/1.1575750

Shera, C. A., and Guinan, J. J. (**1999**). "Evoked otoacoustic emissions arise by two fundamentally different mechanisms: A taxonomy for mammalian OAEs," J. Acoust. Soc. Am., **105**, 782–798. doi:10.1121/1.426948

Shera, C. A., Tubis, A., and Talmadge, C. L. (**2008**). "Testing coherent reflection in chinchilla: Auditory-nerve responses predict stimulus-frequency emissions," J. Acoust. Soc. Am., **124**, 381–95. doi:10.1121/1.2917805

Shinn-Cunningham, B. G. (**2008**). "Object-based auditory and visual attention," Trends Cogn. Sci., **12**, 182–186. doi:10.1016/j.tics.2008.02.003.Object-based

Shinn-Cunningham, B. G., and Best, V. (**2008**). "Selective attention in normal and impaired hearing," Trends Amplif., **12**, 283–299. doi:10.1177/1084713808325306

Siegel, J. H. (**1994**). "Ear-canal standing waves and high-frequency sound calibration using otoacoustic emission probes," J. Acoust. Soc. Am., **95**, 2589–2597. doi:10.1121/1.409829

Siegel, J. H., Cerka, A. J., Recio-Spinoso, A., Temchin, A. N., Van Dijk, P., and Ruggero, M. A. (**2005**). "Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering," J. Acoust. Soc. Am., **118**, 2434–2443. doi:Doi 10.1121/1.2005867

Sisto, R., Moleti, A., and Shera, C. A. (**2015**). "On the spatial distribution of the reflection sources of different latency components of otoacoustic emissions," J. Acoust. Soc. Am., **137**, 768–776. doi:10.1121/1.4906583

Smith, D. W., Aouad, R. K., and Keil, A. (**2012**). "Cognitive Task Demands Modulate the Sensitivity of the Human Cochlea," Front. Psychol., **3**, 1–8. doi:10.3389/fpsyg.2012.00030

Sridhar, T. S., Liberman, M. C., Brown, M. C., and Sewell, W. F. (**1995**). "A novel cholinergic 'slow effect' of efferent stimulation on cochlear potentials in the guinea pig," Neuroscience, **15**, 3667–3678.

Srinivasan, S., Keil, A., Stratis, K., Osborne, A. F., Cerwonka, C., Wong, J., Rieger, B. L., et al. (**2014**). "Interaural attention modulates outer hair cell function," Eur. J. Neurosci., **40**, 3785–3792. doi:10.1111/ejn.12746

Srinivasan, S., Keil, A., Stratis, K., Woodruff Carr, K. L., and Smith, D. W. (**2012**). "Effects of cross-modal selective attention on the sensory periphery: Cochlear sensitivity is altered by selective attention," Neuroscience, **223**, 325–332. doi:10.1016/j.neuroscience.2012.07.062

Stroop, J. R. (**1935**). "Studies of interference in serial verbal reactions," J. Exp. Psychol., **28**, 643–662. doi:10.1037/h0054651

Talmadge, C. L., Tubis, A., Long, G. R., and Tong, C. (**2000**). "Modeling the combined effects of basilar membrane nonlinearity and roughness on stimulus frequency otoacoustic emission fine structure," J. Acoust. Soc. Am., **108**, 2911–32. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11144584

Thompson, G. C., and Thompson, A. M. (**1986**). "Olivocochlear neurons in the squirrel monkey brainstem," J. Comp. Neurol., **254**, 246–58. doi:10.1002/cne.902540208

Treisman, A. M. (**1969**). "Strategies and models of selective attention," Psychol. Rev., **76**, 282–299. doi:10.1037/h0027242

Varghese, L., Bharadwaj, H. M., and Shinn-Cunningham, B. G. (**2015**). "Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses," Brain Res., , doi: 10.1016/j.brainres.2015.06.038. doi:10.1016/j.brainres.2015.06.038

Varghese, L., Bharadwaj, H. M., and Shinn-Cunningham, B. G. (**2015**). "Evidence against attentional state modulating scalp-recorded auditory brainstem steady-state responses," Brain Res., **1626**, 146–164. doi:10.1016/j.brainres.2015.06.038

Vetter, D. E., Liberman, M. C., Mann, J., Barhanin, J., Boulter, J., Brown, M. C., Saffiote-Kolman, J., et al. (**1999**). "Role of alpha9 nicotinic ACh receptor subunits in the development and function of cochlear efferent innervation," Neuron, **23**, 93–103. doi:10.1016/S0896-6273(00)80756-4

Veuillet, E., Collet, L., and Duclaux, R. (**1991**). "Effect of contralateral acoustic stimulation on active cochlear micromechanical properties in human subjects: dependence on stimulus variables," J. Neurophysiol., **65**, 724–35. doi:10.1152/jn.1991.65.3.724

Viemeister, N. F. (**1980**). "Adaptation of masking," In G. van den Brink and F. A. Bilsen (Eds.), Physiol. Behav. Stud. Hear., Noordwikerhout, The Netherlands, 190–199.

Walsh, K. P., Pasanen, E. G., and McFadden, D. (**2014**). "Selective attention reduces physiological noise in the external ear canals of humans II: Visual attention," Hear. Res., **312**, 160–167. doi:10.1016/j.heares.2014.03.013

Walsh, K. P., Pasanen, E. G., and McFadden, D. (**2014**). "Selective attention reduces physiological noise in the external ear canals of humans I: Auditory attention," Hear. Res., **312**, 143–159. doi:10.1016/j.heares.2014.03.012

Walsh, K. P., Pasanen, E. G., and McFadden, D. (**2015**). "Changes in otoacoustic emissions during selective auditory and visual attention," J. Acoust. Soc. Am., **137**, 2737–2757. doi:10.1121/1.4919350

Walsh, K. P., and Wojtczak, M. (**2014**). "Time-cource of recovery from the effects of a notched-noise on the ear-canal pressure at different frequencies," J. Acoust. Soc. Am., **135**, 2386.

Warr, W. B. (**1992**). "Organization of Olivo cochlear Efferent Systems in Mammals," pp. 410–448.

Warr, W. B., and Guinan, J. J. J. (**1979**). "Efferent innervation of the organ of Corti: Two separate systems," Brain Res., **173**, 152–155. doi:10.1016/0006-8993(79)91104-1

Webster, D. B., Popper, A. N., and Fay, R. R. (**1992**). *The Mammalian auditory pathway : neuroanatomy*, (D. B. Webster and R. R. Fay, Eds.) Springer Handb. Audit. Res. v. 1, Springer Science & Business Media, 485 pages. doi:10.1007/978-1-4612-4416-5

Whitehead, M. L., Kamal, N., Lonsbury-Martin, B. L., and Martin, G. K. (**1993**). "Spontaneous otoacoustic emissions in different racial groups," Scand. Audiol., **22**, 3–10. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8465138

Wittekindt, A., Kaiser, J., and Abel, C. (**2014**). "Attentional modulation of the inner ear: A combined otoacoustic emission and EEG study," J. Neurosci., **34**, 9995–10002. doi:10.1523/JNEUROSCI.4861-13.2014

Wojtczak, M., Beim, J. A., and Oxenham, A. J. (**2015**). "Exploring the Role of Feedback-Based Auditory Reflexes in Forward Masking by Schroeder-Phase Complexes," J. Assoc. Res. Otolaryngol., **16**, 81–99. doi:10.1007/s10162-014-0495-3

Wojtczak, M., and Oxenham, A. J. (**2009**). "On- and Off-Frequency Forward Masking by Schroeder-Phase Complexes," J. Assoc. Res. Otolaryngol., **10**, 595–607. doi:10.1007/s10162-009-0180-0

Zweig, G., and Shera, C. A. (**1995**). "The origin of periodicity in the spectrum of evoked otoacoustic emissions," J. Acoust. Soc. Am., **98**, 2018–2047. doi:10.1121/1.413320