DISENTANGLING THE INFLUENCE OF ATTENTION IN THE AUDITORY EFFERENT SYSTEM DURING SPEECH PROCESSING

Heivet Hernandez Perez
Bachelor in Biology
Master in Physiology

A thesis submitted in fulfilment of the requirements for the degree of:

DOCTOR OF PHILOSOPHY

Audiology Section, Department of Linguistics
Faculty of Human Sciences
Macquarie University
Sydney, Australia

Submitted: 2nd of March 2018

Abstract

The physiological mechanisms allowing humans to selectively attend to a single conversation in acoustically adverse situations, such as overlapping conversations or background noise, are poorly understood. In particular, the extent to which goal-directed, top-down processes of auditory attention can modulate the inner ear activity via the auditory efferent system remains unclear. This thesis investigates the relationship between degraded speech and the auditory efferent control of the cochlea. Young, normal-hearing, participants were assessed in a series of three experiments where speech intelligibility was manipulated during Active and Passive listening to: 1) noise vocoded speech; 2) speech in babble noise and 3) speech in speech-shaped noise. A lexical decision task was used in the “Active” listening condition where subjects were instructed to press a button each time they heard a non-word. In the “Passive” listening condition they were instructed to ignore all auditory stimuli and watch a movie. Click-evoked OAEs (CEOAEs) were obtained from the ear contralateral to the speech stimuli, allowing the measurement of cochlear-gain changes. A 64-channel EEG was synchronized with the CEOAE recording system, enabling the simultaneous measurement of cortical speech-onset event-related potentials (ERPs), click-evoked auditory brainstem responses (ABRs) and behavioural responses. Behavioural results showed that accuracy declined as the speech signals were degraded, while ERPs components were enhanced during the Active condition compared to the Passive condition. A decrease in cochlear gain (reduction in CEOAE amplitudes) with increasing task difficulty was observed for noise vocoded speech, but not for speech in babble or speech-shaped noise. Brainstem components showed decreased activity linked to CEOAE suppression. These findings contribute to an integrative view of auditory attention as an adaptive mechanism that recruits cochlear gain control via the auditory efferent system in a manner dependent upon the auditory scene encountered.

Statement from Author

I state that this work has been submitted exclusively to Macquarie University in Sydney, Australia, for the consideration of a PhD degree.

Ethical review, guidance and approval have been obtained from Macquarie University Ethics Review Committee (Human Research). No. 5201500235 (see Appendix section).

I certify that I developed the original idea in collaboration with my three supervisors (Prof Catherine McMahon, Dr Jessica Monaghan and Prof Sumitrajit Dhar) and Co-authors (Prof David Poeppel and A/Prof Sriram Boothlingam). I have taken the leadership to conduct all parts of this research work, including writing the content of this thesis. My supervisors and collaborators have assisted in improving the research protocol, analyses, and interpretation of the data, as well as the quality of the written work. I conducted the data collection, with support from my supervisor Dr Jessica Monaghan. Statistical support has been obtained from Dr Jessica Monaghan and Dr Peter Humburg.

Heivet Hernandez Perez

Table of Contents

Acknowledgements

"It always seems impossible until it's done" Nelson Mandela

This is exactly how my PhD adventure felt but it was possible with the inspiration, support and encouragement of many amazing people.

I would like to thank my dear parents who showed what hard work and discipline are by example. To my brother and uncle who supported me in this PhD journey which brought me to Australia (whether or not they agreed). To Professor Timothy Devoogd, who since we first met in Cuba in an IBRO school, encouraged and supported me to pursue PhD studies overseas…which sounded like a dream at the time. I would also like to thank my dear friends and former boss/supervisor back in Cuba, MD. PhD Alejandro Torres Fortuny and co-worker and mentor Eduardo Eimil for teaching me so much about hearing and electrophysiology.

I'm deeply grateful to my supervisors, Prof Catherine McMahon, Prof Sumitrajit Dhar and Dr. Jessica Monaghan, without whose expertise, experience, knowledge and patience this project would not have been possible. Catherine, thanks so much for fighting so fiercely for my transfer from University of Queensland to Macquarie University, here is where I was meant to be. You inspire dedication for your work and your family and I admire you for that. To Professor Sumitrajit Dhar, thanks for accepting me as your PhD student from afar and for making me feel at home every time I visited Northwestern University. Your family (Julia, Sachi and Suvali) is beautiful in every way. Without your unconditional patience and availability for skype meetings and remotely fixing my mistakes (taking into consideration the time differences between Chicago and Sydney), the OAE recording setup would not exist. To my dear friend and supervisor, Dr. Jessica Monaghan, you have been a constant and unconditional support, even when Oscar was only days old, you were there for me. I

feel tremendously honoured to be your 1st PhD student. You are a brilliant scientist and I only hope to be able to work with you for many more years.

To my closest collaborators, Prof David Poeppel and A/Prof Sriram Boothalingam, I'm sincerely thankful for the time you have invested in my PhD project. David, you are one of the reasons why I ended up at Macquarie University, thanks so much for backing me up. Sriram, thanks so much for always being so positive and willing to share your knowledge.

I would like to express my appreciation to Professor David Ryugo's laboratory (Mike, Cath, Kirupa, Giedre, Kiera, Femi and Tan). I have learnt so much about anatomy and physiology of the auditory system from you. Thank you for making me part of the gang, you are great fun to be around!

To the great people at Macquarie University and the Australian Hearing Hub. Isabelle Boisvert, thanks so much for receiving me in your home when I first moved to Sydney. You have been incredible supportive and helpful in assisting me in supplies purchases, sharing stimuli database and just being there for me. Ronny, if I know something about fieldtrip after this PhD is all because of you, thanks so much for always being so helpful and kind. I want to say thank you to the McAlpine group. David, thanks so much for your incredible support. Jaime and Nick, you are amazing researchers and co-workers, thanks for always help me out since you first arrived! Maca and Lindsey you are great lab-mates, thanks for always cheering me up! To our amazing Audiology and Hearing Research Group, thanks so much for such great scientific discussions. Carmen Kung, thanks so much for sharing your knowledge in Presentation programming. Elaine Schmidt and Chi Lo thanks for your linguistic assistance during the speech tokens selection. Bram, Fabrice, Peter De Lissa and Joaquin thanks so much for helping with the initial Neuroscan setup! Greg Stuart,

nothing could have ever been connected if it were not for you… “literally”. Thanks so much for always being so helpful and always available for my weird/urgent requests of equipment, calibration and cables. John Newall and Phillip Nakad thanks so much for assisting me with the Titan! Louise Dodd, Collette Ryan, Rick Filmer, I really appreciate your help with purchases and dealing with equipment/room issues on this journey. Finally, to all my participants, thanks so much for showing up and being so patient and understanding!

To all my PhD brothers and sisters: Chi, Kelly, Mariana, Chris, Fadwa, Amanda, Adam, Rakshita, Shivali, David Meng, Chavi, Remi, Rebecca: thanks for always be willing to have a chat and put the EEG cap on to listen to all my weird experimental sounds.

I would also like to say thank you to friends, whom, in a way or another have encouraged and support me during this adventure: my dear professors Olga and Martha, my French-mum Monique, my dear friend in NYC Eileen, Prof Elyse Sussman, Joann, Aert, Ara, Lachy, Su, Laura, Yaise and Yeni.

Last but not least, to my dear husband Jason, THANK YOU! You have been incredibly patient, caring, helpful and supportive all along this PhD journey!

List of Tables

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Appendix

List of Abbreviations

ABRs: Auditory Brainstem Responses
AC: Auditory Cortex
ACh: Acetylcholine
AN: Auditory Nerve
ARC: Australian Research Council
ASSRs: Auditory Steady State Responses
AVCN: Antero-Ventral Cochlear Nucleus
BBN: Broadband Noise
BM: Basilar Membrane
BN: Babble Noise
BN10: Speech in Babble Noise +10 dB SNR
BN5: Speech in Babble Noise +5 dB SNR
BN7: Speech in Babble Noise +7 dB SNR
Ca++: Calcium ions
CAP: Compound Action Potentials
CEOAEs: Click-Evoked Otoacoustic Emissions
CF: Characteristic Frequency
CM: Cochlear Microphonics
CN: Cochlear Nucleus
CNC: Consonant-Nucleus-Consonant
CNS: Central Nervous System
COCB: Crossed OlivoCochlear Bundle
DCN: Dorsal Cochlear Nucleus
DPOAE: Distortion Products Otoacoustic Emissions
EEG: Electroencephalography
ERPs: Event Related Potentials
FA: False Alarm rate
GABA: Gamma-AminoButyric Acid
HR: Hit Rate
IC: Inferior Colliculus
IHC: Inner Hair Cells
K+: Potassium ions
LOC: Lateral Olivary Complex
MEMR: Middle Ear Muscle Reflex

MGB: Medial Geniculate Body of the thalamus
MOC: Medial OlivoCochlear
OAEs: OtoAcoustic Emissions
OHC: Outer Hair Cells
PVCN: Postero-Ventral Nucleus
RL: Reticular Lamina
RMANOVA: Repeated Measured Analysis of Variance
RMS: Root Mean Square
RT FA: Reaction Times to False Alarm responses
RT Hit: Reaction Times to Hit responses
SE: Standard Error
SFOAE: Stimulus Frequency Otoacoustic Emissions
SNR: Signal to Noise Ratio
SOC: Superior Olivary Complex
SSN: Speech-Shaped Noise
SSN3: Speech in Speech-Shaped Noise +3 dB SNR
SSN5: Speech in Speech-Shaped Noise +5 dB SNR
SSN8: Speech in Speech-Shaped Noise +8 dB SNR
TM: Tectorial Membrane
TOAE: Transient OtoAcoustic Emissions
TTLs: Transistor–Transistor Logic signals
UOCB: Uncrossed OlivoCochlear Bundle
VNTB: Ventral Nucleus of the Trapezoid Body
Voc12: 12 channels noise Vocoded speech
Voc16: 16 channels noise Vocoded speech
Voc8: 8 channels noise Vocoded speech

Chapter 1 Introduction

1.1 Preamble

The “cocktail party effect”, first coined by Cherry in 1953, describes one’s ability to selectively focus attention on a single conversation amid multiple simultaneous and overlapping conversations. However, to date, the physiological mechanisms which contribute to this ability remain unclear. Research has revealed that top-down attention is likely to operate via a neural gain control mechanism, whereby signals of interest are perceptually enhanced while the background scene is suppressed in order to maintain goal-directed behaviour (Ahveninen et al., 2006; Johnson & Zatorre 2006; Knudsen 2007). Some studies in human subjects have concluded that attention affects the initial stages of cortical processing (Karns & Knight, 2009; Saupe et al., 2009), but not subcortical stages (Picton et al., 1971; Connolly et al., 1989). However, this concept has been challenged since 1946 when Rasmussen first demonstrated the bidirectional nature of the auditory pathway, describing the existence of descending neuronal projections from the brainstem to the cochlea: an auditory efferent system. Moreover, in 1956 Hernandez-Peon described how sound-evoked neural responses in the cochlear nucleus (CN) (i.e., the 1st relay station in the auditory pathway) were reduced when cats were presented with olfactory or visual stimuli. More recently, several studies have demonstrated ways in which the auditory periphery is modulated by attention (Meric & Collet, 1994; Giard et al., 2000; De Boer & Thornton, 2007).

Multiple steps have been highlighted as key to the perception and comprehension of speech: these include the efficient transmission of sounds through the external and middle ears; their subsequent spectral analysis and transduction into electrical signals in the cochlea as well as the integration of the extracted acoustic information along the auditory pathway till reaching the auditory cortex (AC) (Plomp 2001; Davis & Johnsrude 2007). However, for this to happen, several active mechanisms need to occur during the

sound transduction in the cochlea. Active and highly-localized increases in cochlear gain result in high-frequency selectivity and specificity which, as a by-product, generate low-level acoustic emissions from the cochleae (known as otoacoustic emissions: OAEs) (Kemp, 1978). Auditory information then ascends to higher auditory centres and, ultimately, forms complex descending loops between auditory structures. These descending pathways can modulate the cochlear gain (and, hence, the magnitude of OAEs), and may help modulate perception and cognition to facilitate speech comprehension in complex listening environments. Within this thesis, we seek to understand how attention to spoken words can modulate the auditory efferent system at both the level of the cochlea and auditory brainstem.

1.2 Background

Descending fibres from the AC and inferior colliculus (IC) target two types of auditory brainstem neurons in and around the superior olivary complex (SOC): lateral olivocochlear (LOC) and medial olivocochlear (MOC) neurons, which innervate the outer (OHCs) and inner (IHCs) hair cells respectively (Rasmussen 1946, Warr et al. 1979, Suga & Ma 2003). OHCs have a unique electromotile property, which amplifies the vibration of the basilar membrane (BM) and underpins the cochlear gain (Ashmore et al., 2010). Previous animal studies have shown that electrical stimulation of neurons in the AC (Xiao & Suga, 2002; Perrot et al., 2006; Dragicevic et al., 2015) and IC (Mulders & Robertson, 2000a; Popelar et al., 2002; Ota et al., 2004; Zhang & Dolan 2006) can induce changes in the cochlear gain. Such changes can be quantified by measuring a by-product of OHC electromotility, the low-level sounds emitted from the ear canal known as OAEs (Kemp 1978). It is well accepted that when transient sounds (such as brief tones or speech tokens) are presented in noise, auditory efferent activity (via MOC neurons) suppresses OAE amplitudes, reflecting a reduction in the cochlear gain. This has been hypothesised to be a mechanism that acts to reduce the cochlear gain to

noise, while enhancing the perception of the transients (Garinis et al., 2011; De Boer et al., 2012). This may be further modulated by auditory attention, providing a cochlear gain control mechanism which could, in part, explain how auditory attention can enable humans to selectively listen to one conversation within a “cocktail party” of sounds.

Multiple studies have explored how attention modulates MOC inhibitory effects in the cochleae of normal hearing subjects: however, the results have proven inconsistent: while some studies have shown a reduction in the MOC inhibitory drive to the OHCs (Ferber-Mart et al., 1995; Harkrider & Bowers, 2009), others have demonstrated increases (Maison et al., 2001; De Boer & Thornton, 2007; Garinis et al., 2011; Wittekindt et al., 2014). Despite these results appearing contradictory, they would suggest that cortical or higher-order processing involved in the focus of auditory attention such as the temporal and frontal areas (Alho et al., 2006; Alho & Vorobyev 2007), parietal (Wu et al., 2007) and cingulate cortical areas (Bidet-Caulet & Bertrand, 2005) may help modulate the MOC inhibitory drive. It is well accepted that auditory attention is task-specific and physiological mechanisms involved in auditory attention operate above the auditory thalamus, mainly in the AC (Karns & Knight, 2009; Saupe et al., 2009); however, it remains unclear whether such attentional mechanisms influence the auditory efferent system. Within this chapter, we present a broad overview of cochlear function, describing the mechanisms that underpin OAE generation and their modulation via the MOC. We also demonstrate how OAEs can be used to objectively measure the magnitude of the MOC effects on the cochlear gain. Finally, we discuss the evidence for the anatomical pathways which could support a model of attentional modulation of the auditory pathway, as well as the current research which shows how different attentional paradigms, across different sensory modalities (auditory and visual mainly) have influenced auditory function. Herein we pursue the goals of disentangling the effects of attending to speech tasks on the auditory efferent system.

1.3 Cochlear function

The inner ear is a highly sophisticated structure serving both hearing and balance. The cochlear hair cells are located inside a bony structure, filled with fluid and partitioned by membranes, called the cochlea. The three distinguishable compartments are the scala vestibuli, scala media and scala tympani (Figure 1.1). The scala media is filled with endolymph, an extracellular fluid that is chemically distinct from perilymph (present in the other two scalae) yet is similar in ionic composition to intracellular fluids (Waltner & Raymond, 1950). The scala media is a tubular structure that contains the organ of Corti, which sits on the BM. Both the scalae vestibuli and tympani have direct communication with the middle ear cavity through the oval and round window respectively, see Figure 1.1.

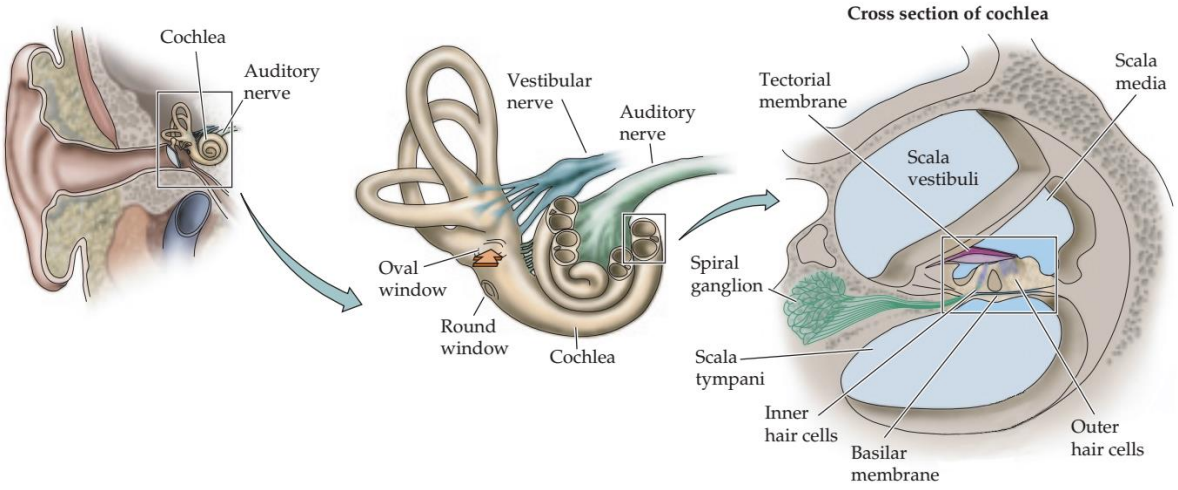

Figure 1.1. Schematic representation of the cross section of the cochlea. Auditory and vestibular nerves are represented as well as the oval and round windows that connect the inner and middle ears. The three partitions are shown: Scala vestibuli, Scala tympani and the Scala media. Scala media is full of endolymph while scalae vestibuli and tympani contain the perilymph. The rectangle shows the organ of Corti. (Edward, 2006), used with permission.

The organ of Corti contains IHCs and OHCs as well as the supporting cells that surround the receptors. Hair cells are modified epithelial cells with small stereocilia that protrude from their apical sections. Despite both types of hair cells contributing to the transduction of the sound in the inner ear, they differ in their number, morphology and location as well as their afferent and efferent innervations (Bredberg et al., 1972). The stimulus for generating a receptor potential in the hair cells is the displacement of stereocilia; however, several cascading events precede this. That is, incident sound waves evoke movement of the BM, which behaves as a spectrum analyser due to its physical properties changing along its length (width and stiffness) (Voldrich, 1978; Kössl & Russell 1995, Emadi et al., 2004).

The combined up-and-down motion of the BM and tectorial membrane (TM) causes a sinusoidal voltage modulation in both OHCs and IHCs in the organ of corti as their stereocilia move back and forth (Patuzzi, 1996). The stereocilia movement in the OHCs is due to the direct connection of the stereocilia to the TM whereas the flow of the endolymph induces movement of the IHC stereocilia (shear between the TM and the reticular lamina (RL) (a thin and stiff lamina that runs parallel to the TM and underneath the OHC stereocilia) (Chadwick et al., 1996; Nowotny & Gummer, 2006, 2011). Deflections of stereocilia in the direction of the largest cilia open non-selective mechanically gated ion channels that increases the depolarizing influxes of calcium ions (Ca++) and potassium ions (K+) into the cell (Roberts & Howard 1988). Contrarily, displacement of the organ of Corti in the opposite direction (away from the tectorial membrane) leads to hyperpolarization.

Depolarization of IHCs leads to the basolateral release of vesicular glutamate into the synaptic cleft and onto primary afferent dendrites (Chan & Hudspeth, 2005). These are the bipolar, first-order neurons of the afferent auditory pathway and this excitatory neurotransmitter increases their firing rates. Approximately 35,000 spiral ganglion axons

constitute the auditory nerve (AN) in the young human auditory system (Otte et al., 1978). Of these, 95% are Type I that synapse with individual IHCs, while the remaining 5% (Type II) are branched and synapse with multiple OHCs (Liberman et al., 1990). While it is clear that the Type I afferent neurons transmit acoustic information to higher auditory centres, the role of the Type II afferent neurons is less clear (Brown, 1994; Robertson et al., 1999; Weisz et al., 2009).

However, cochlear vibration is not only a passive mechanical resonance as it also involves active processes. OHCs are capable of generating movements that continuously deliver energy back into the organ of Corti and alter the mechanical properties of the cochlea (Robles & Ruggero, 2001; Ashmore et al., 2010). Because the cochlea is a bony structure, increases in the cochlear partition pressure need to be released through the inner ear's connection with the middle ear i.e. the oval and round windows. In particular, the movement of the oval window evokes a retrograde displacement of the ossicles that eventually reaches the eardrum. These pressure perturbations in the eardrum can be recorded in the external ear canal as low intensity sounds which are the OAEs (Kemp, 1978).

1.4 Generation of otoacoustic emissions

OAEs were hypothesized by Thomas Gold as a physical-acoustic phenomenon in 1948 when he proposed the existence of a mechanism for cochlear gain amplification (termed the "cochlear amplifier") (Gold & Pumphrey 1948; Gold, 1948). However, it was not until 1978 that David Kemp demonstrated the physiologically active cochlear mechanism leading to the amplification of the sound. In this study, Kemp described how a healthy cochlea responds to an acoustical stimulus with the retrograde sound transmission that can be measured with a sensitive microphone in the external ear canal.

The active electrical generation of contractions by the OHCs' somata and sterocilia is frequency-specific and produces the OAEs (Brownell et al., 1985; Ashmore et al., 2010). Here individual OHCs can act as non-linear oscillators in a limited dynamic range where the contribution to overall basilar membrane vibration becomes relatively less important at higher intensities (Patuzzi et al., 1986). Recently, it has been suggested that the OHC amplification causes more movements in the RL than in the BM, particularly at frequencies below the characteristic frequency (CF: frequency at which the BM respond to the smallest sound intensity) (Guinan 2017).

Two different mechanisms in the cochlea help generate evoked OAEs: linear reflection and nonlinear distortion (Shera & Guinan 1999). Evoked OAEs require sound stimulation for their generation and can be sub-classified according to the type of sound used to evoke them. When a pair of tones Is delivered to the cochlea, they evoke distortion products OAEs (DPOAEs) which are backward-travelling waves that arise from nonlinear distortion sources (Kemp & Brown 1983; Shera, 2004). On the other hand, transient pips, noise-bursts, clicks or chirps can evoke transient OAEs (TOAEs) while a single pure tone presented with a second suppressor tone induce stimulus-frequency OAEs (SFOAEs). These all arise from linear reflections of the forward-travelling wave most likely due to anatomical variations in IHC and OHC number and geometry (Engström et al., 1966; Wright, 1984) as well as variations in the density or structure of the prestin (the OHC motor proteins) (Zheng et al., 2000; Shera, 2004).

1.5 Medial Olivocochlear Reflex

The MOC reflex is comprised of afferent and efferent pathways that, when activated, can result in the modulation of OAE magnitudes. MOC neurons are central to this auditory brainstem reflex, their neural output ultimately affecting those changes to OAE magnitudes by manipulation of OHC function. In this section, the anatomy and physiology of the MOC reflex is described in depth, providing a solid foundation for

understanding how attentional control may influence MOC nuclei and therefore affect OAE magnitudes.

1.5.1 Medial Olivocochlear Reflex – anatomy

The first stage in the MOC reflex involves the AN conveying sound-evoked electrical activity from the periphery to the auditory brainstem. While it has been recently suggested that Type II AN-fibres may help convey this auditory signal (Froud et al., 2015), Type I AN-fibres are likely the main afferent input to the reflex pathway (Maison et al., 2016).

After the AN, the auditory signal is relayed via an intermediate nucleus, the CN, to the MOC neurons. Located bilaterally at the pontomedullary junctions (Figure 1.2), the CN consists of three sub-nuclei in both humans and non-primates: the dorsal cochlear nucleus (DCN), the antero-ventral cochlear nucleus (AVCN) and the postero-ventral nucleus (PVCN) (Ryugo & Parks 2003). Although afferent projections from the AVCN (Ye et al., 2000) and PVCN (Thompson & Thompson, 1991; Smith et al., 1993; Doucet & Ryugo, 2003) have been identified as direct inputs to MOC neurons, only lesions of the caudal PVCN have been shown to effectively disrupt the MOC reflex (De Venecia et al., 2005) and this is therefore likely the location of the MOC reflex's interneurons. Indeed, some axonal projections of stellate/multipolar cells have been shown to exit the PVCN (Doucet & Ryugo, 2003), cross the midline and provide excitatory input to contralateral and ipsilateral MOC neurons (Smith et al., 1993; Darrow et al., 2012).

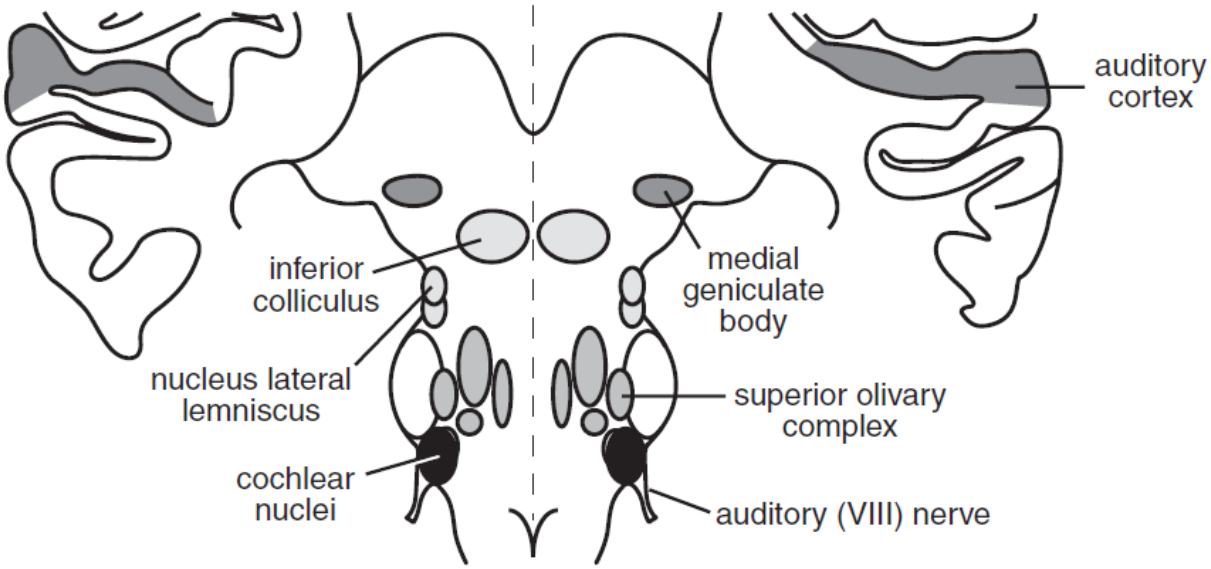

Figure 1.2. Schematic representation of the major relay structures in the auditory system. The first relay structure is the cochlear nucleus, followed by the superior olivary complex, nucleus of the lateral lemniscus, the inferior colliculus, medial geniculate body in the thalamus and finally the primary AC in the temporal lobe. Note that parallel pathways in the brainstem are responsible for the communication among the relay structures (not shown). (Brugge & Howard 2002), used with permission.

The final stage of the reflex is the activation of the MOC neurons in or around the ventral nuclei of the trapezoid body (VNTB) whose axons then project bilaterally back and innervate OHCs in the cochleae (Rasmussen, 1946; Warr & Guinan, 1979; Brown & Levine, 2008). The majority of MOC axons travel dorsally to the fourth ventricle where they cross the midline and join the olivocochlear bundle innervating the contralateral cochlea; a smaller number of MOC fibres do not decussate and instead innervate the ipsilateral cochlea (Warr et al., 1986; Azeredo et al., 1999; Maison et al., 2003), see Figure 1.3. Whether an MOC neuron's axon is crossed or uncrossed in the olivocochlear bundle likely determines its involvement in either the ipsilateral (blue arrows/neurons in Figure 1.3) or contralateral (red arrows/neurons in Figure 1.3) MOC reflex pathways.

While both reflex pathways incorporate the same initial stages where an incoming sound drives contralateral MOC neurons via decussating PVCN fibres (red/blue arrows from CN in Figure 1.3), they diverge in how the MOC axons then project towards the target cochlea. In the ipsilateral reflex, the MOC efferent projections (thin blue lines emanating from blue neurons, Figure 1.3) cross back via the crossed olivocochlear bundle (COCB/thick gold line, Figure 1.3) to innervate the cochlea originally stimulated by sound. In the contralateral reflex, the MOC efferent projections (thin red lines emanating from red neurons, Figure 1.3) do not cross back to the sound-stimulated cochlea but instead join the uncrossed olivocochlear cochlear bundle (UOCB/thick gold line, Figure 1.3), conveying the efferent signal to the unstimulated cochlea (De Venecia et al., 2005; Guinan, 2006).

In humans, the exact ratio of crossed Vs. uncrossed MOC axons is unknown (Guinan 2017), however it has been estimated that the average olivocochlear bundle of humans contains around 1,400 efferent fibres (Arnesen, 1984). 30% of these fibres are the thick, myelinated MOC fibres targeting OHCs (Guinan et al., 1983; Arnesen, 1984; Warr, 1992) while the remaining 70% are the thin, unmyelinated LOC fibres that contact IHCs and their Type I AN-fibres (Arnesen, 1984; Brown, 1987; Warr et al., 1997).

On entering the cochlea, individual MOC fibres can contact a number of OHCs, spanning a cochlear distance of up to an octave (Liberman & Brown, 1986). This innervation of OHCs is not random however, as an MOC fibre stimulated by a tone of specific frequency will project back to cochlear regions of matching tonotopy (Robertson & Gummer 1985; Liberman & Brown 1986; Brown 1989). Therefore, single MOC neurons likely affect sound processing in the cochlea in a frequency-specific manner. However they do so by affecting a frequency-band rather than a single frequency (Brown & Vetter 2009). This frequency-specificity appears to differ between ipsilateral and contralateral MOC reflex pathways: contralateral-reflex MOC axons innervating a larger,

more apical region of the cochlea compared to ipsilateral-reflex MOC axons (Brown, 2014).

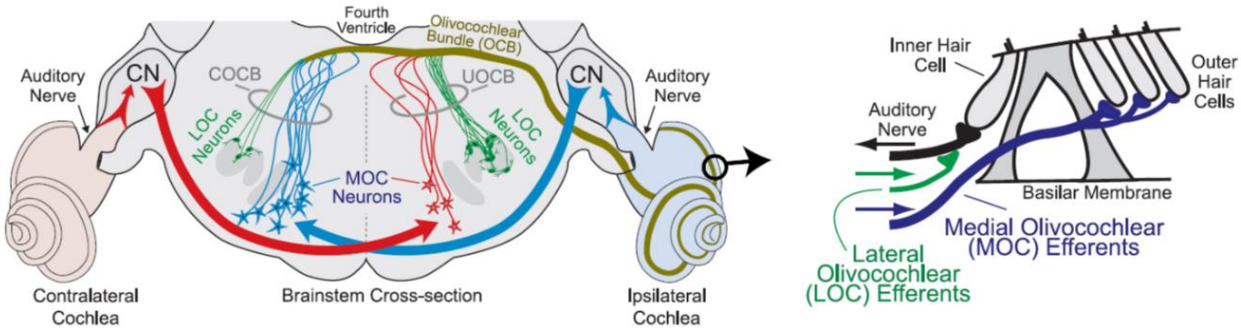

Figure 1.3. Schematic representation of the olivocochlear reflex in cats. Afferent innervation (represented in red and blue ascending arrows) from PVCN in the CN cross the midline to establish synapses with MOC neurons. MOC and LOC neurons in the crossed (COCB) and uncrossed (UOCB) olivocochlear bundle (in gold color) are also shown. On the right is represented the LOC and MOC innervation in the inner ear. (Guinan, 2006), used with permission.

1.5.2 Medial Olivocochlear Reflex – physiology

Electrical stimulation of MOC fibres at the floor of the fourth ventricle has been shown to reduce sound-evoked motion in the cochlea (Cooper & Guinan 2006) and decrease auditory nerve responses (Guinan & Gifford, 1988: Elgueda et al., 2011). These effects are indicative of reduced cochlear gain when MOC efferents are active; a phenomenon mediated by the release of the neurotransmitter, acetylcholine (ACh), from MOC terminals at their synapses with OHCs. ACh binds to postsynaptic α9 cholinergic receptors on OHCs (Robertson 1984; Liberman & Brown 1986; Wersinger & Fuchs, 2011), producing a large influx of Ca^{++} (Elgoyhen et al. 1994, 2001) that in turn triggers Ca^{++}- dependent K^{+} currents in the basolateral regions of OHCs (Housley & Ashmore

1991; 1996). The resulting net hyperpolarization of OHCs, together with the accompanying increase in the basolateral conductances, help to reduce the gain of the cochlear amplifier and diminish the endocochlear potential (Gifford & Guinan 1987). As a result, more energy is required to evoke BM vibrations as well as IHC activity and later AN-responses (Murugasu & Russell 1996; Cooper & Guinan 2003).

Two types of MOC efferent effect on inner ear function have been described: 1) a fast effect, with a duration of (~100 ms) and 2) a slow effect with a duration of (~10 sec) (Sridhar et al., 1995; Cooper & Guinan 2003). These are known as the classical MOC effects, both turning down the gain of the cochlear amplifier. MOC efferent activation also shows non-classical effects that cannot be explained either by simple vibrations of the organ of Corti nor MOC effects on the cochlear amplifier and whose function is not well understood (Gifford & Guinan 1983; Liberman & Klang 1984; Guinan et al. 2005). The classical MOC effects can be measured in the BM, RL, IHC, AN activity, cochlear microphonics (CM) (a measure of the gross OHC receptor potential) (Elgueda et al., 2011) and OAE production (Guinan 2011). MOC fibres responses are sound-level dependent and are able to respond to sounds with a latency of 5-10ms (Robertson & Gummer 1985; Liberman & Brown 1986; Brown et al., 2003).

When a transient sound, such as a tone-burst, is presented simultaneously with a continuous background noise, the AN-responses are shifted towards higher sound intensity levels (i.e., AN-responses to the tone burst are masked), meaning that more energy is required for the AN to respond to transient sound (Winslow & Sachs 1987), see Figure 1.4c. However, if MOC fibres are electrically stimulated while these two stimuli are being presented, MOC activity reduces cochlear amplification as well as AN-responses to the noise while restoring the dynamic range to the transient sound, see Figure 1.4d. This effect is known as the MOC antimasking effect and allows IHC and AN-responses to still be sensitive to small increases in the transient sound level

(Wiederhold, 1970; Winslow & Sachs 1987; Guinan & Gifford 1988; Kawase et al., 1993).

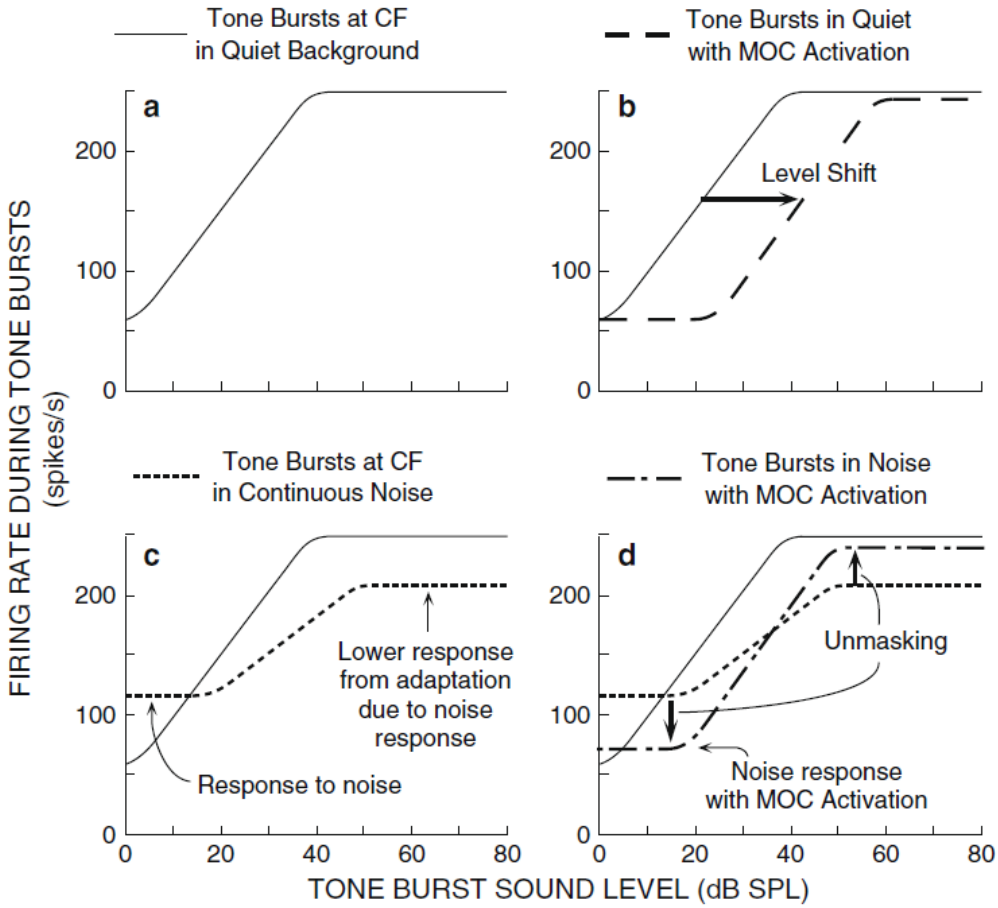

Figure 1.4. Antimasking effects of MOC in the AN. a. The firing rate of AN fibres when tone bursts are delivered at CF in quiet. b. The shift in firing rate after the electrical activation of MOC while tone bursts stimuli are still presented in quiet. c. The firing rate patterns when tone bursts are presented in noise. d. The restored firing rate of the AN fibres after the electrical activation of MOC fibres when tone bursts are presented in noise, (Guinan 1996), used with permission.

1.6 Using OAEs to measure MOC effects on the cochlear gain

OAEs can be used as an objective measure of MOC effects on the cochlear amplifier. The MOC reflex can be experimentally activated using broadband noise (BBN) presented simultaneously with an OAE-evoking acoustic stimulus. This technique

constitutes a unique non-invasive window of opportunity to explore how the MOC reflex modulates the cochlear amplifier in animal models and humans. To measure its effect on OAE recordings, it is necessary to obtain OAE amplitudes with and without MOC reflex activation (i.e., with and without noise). The difference between these two conditions is considered as a decrease in the cochlear amplification due to MOC efferent effects. The effects of the MOC reflex can be observed by presenting an ipsilateral, contralateral or bilateral noise simultaneously with the OAE elicitor (Collet et al., 1990; Perrot & Collet 2014). Collet et al. (1990) found that when a contralateral white noise was presented in increasing intensity steps, it increased MOC efferent inhibition through the contralateral MOC reflex. As a result, the gain of the cochlear amplifier, measured as decreases in TOAE amplitude in the ear contralateral to the noise, was reduced.

Most studies that explored MOC effects in TOAEs have found an average decrease in amplitude across subjects of 1-3 dB either using bilateral, ipsilateral or contralateral stimulation (Perrot & Collet, 2014). Although in humans, ipsilateral and contralateral MOC reflexes have quite similar effects, most studies have used contralateral sound to elicit MOC activity because it avoids the possible cancelations of the elicitor waveform when an ipsilateral acoustic stimulus is presented. The largest MOC effects are found at 1–2 kHz under contralateral stimulation (Liberman, 1988, Collet et al., 1990, Hood et al. 1996; Goodman et al. 2013). Such studies show that MOC efferent effects produce a significant averaged change in OAE magnitude across all subjects, but these effects within individuals are not as easy to observe due to the high variability of OAE measurements (Kemp, 2002, Mishra & Ludman 2013). However, Goodman et al. (2013) have shown consistent MOC effects of 1-3 dB in TOAEs only when the OAE magnitude is 9-22 dB higher that the noise floor (i.e. OAE signal to noise ratios (SNRs) of 9-22 dB) at an individual level. Nevertheless, according to Guinan

(2017), lower SNRs have successfully been used in group studies to explore MOC effects on OAE magnitudes.

When high intensity sounds are delivered to an ear they evoke contractions of both the stapedius and the tensor tympani muscles causing the stiffening of the ossicular chain and an increase in the impedance of middle ear sound transmission (middle ear muscle reflex: MEMR) (Murata et al.,1986; Liberman & Guinan,1998). Accordingly, retrograde middle ear transmission of OAE magnitude can be reduced, not as a result of MOC activation, but of MEMR or middle ear contractions (Lee et al., 2006). It has been shown that even sounds 10-15 dB below the MEMR threshold can cause contractions of the middle ear muscles (Feeney et al., 2003), therefore confounding the effects of MOC on OAEs magnitude.

In general, no matter which type of OAE is explored, it is possible to observe a reduction in OAE amplitude when the MOC reflex is activated (Guinan, 2006) that can last for several minutes. In particular, Moulin and Carrier (1998) showed DPOAE magnitude suppression for more than 20 min. TEOAEs and SFOAEs originate mainly from linear reflection mechanisms; therefore, it is possible to observe a pattern of OAEs magnitude reduction in the majority of these cases when the MOC reflex is active (Brown & Beveridge, 1997; Shera & Guinan, 1999). DPOAEs, on the other hand, seem more complex since the two tones can interact between themselves and decrease or increase OAEs amplitude when the MOC reflex is active (Siegel & Kim 1982; Wagner et al., 2007). However, click-evoked OAEs (CEOAEs) are the simplest TOAEs to use when measuring MOC effects mainly because the click stimulus can be easily separated in the time domain from the OAE signal (Guinan 2017).

1.7 Higher processing centres targeting MOC

For attention to modulate the auditory pathway, associations between higher order cortical centres and lower level auditory areas must exist. The auditory system is

formed by multiple loops and chains of ascending and descending information established between auditory processing centres (Terreros & Delano 2015). Notice how the AC and IC target the SOC where MOC neurons are located in Figure 1.5. It is believed that these interconnections control auditory inputs and may improve the ability to extract relevant information in multi-source auditory scene situations.

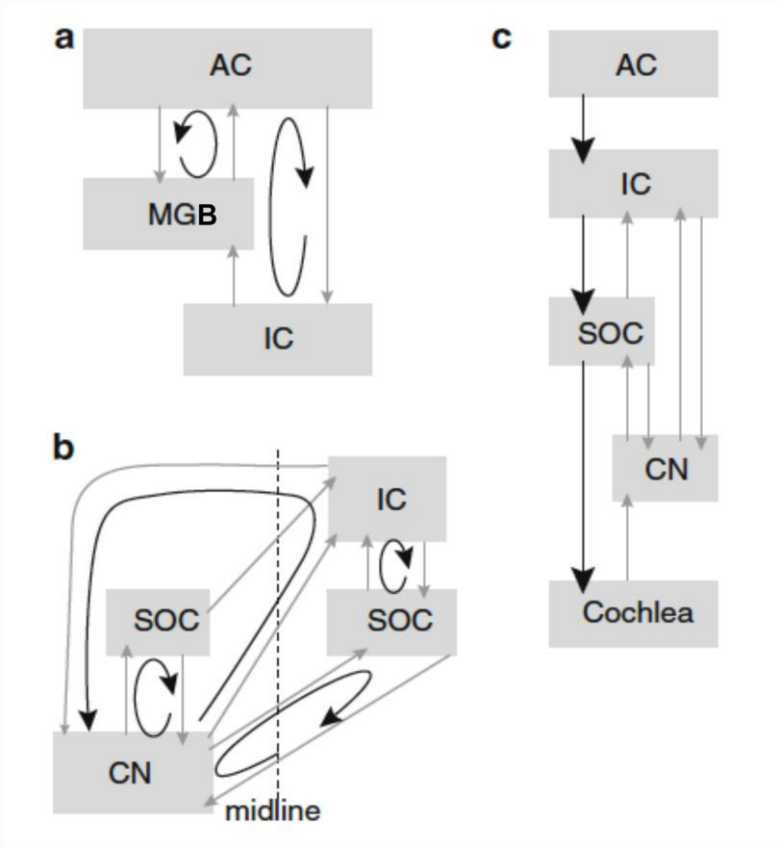

Figure 1.5. Schematic representation of afferent and efferent connections in the auditory system. Notice the chains and loops established between the AC and subcortical auditory areas including the cochlea. (Ryugo et al., 2010), used with permission.

It is worth mentioning that the CN, which constitutes the afferent input in the MOC reflex, also receives efferent innervations from AC, MGB, IC and SOC. Consequently, these efferent inputs can modulate CN function through a variety of neurotransmitters (glutamate, glycine, ACh, GABA) (Schofield 2011) as well as modulating the MOC reflex

at any time. Nevertheless, the main focus of this section will be on the efferent inputs targeting the MOC neurons from the IC and AC.

Anatomical efferent connections between MOC neurons and the IC (Faye-Lund 1986; Thompson & Thompson 1993; Vetter et al., 1993) as well as MOC neurons and the AC (Mulders & Robertson 2000b) have been well described. Electrophysiological studies where IC has been electrically stimulated have shown the excitatory nature of the IC innervation to MOC neurons (Popelar et al., 2002, Zhang & Dolan 2006). In these experiments, the ipsilateral MOC neurons have been recruited and as a by-product the cochlear gain is reduced in a frequency-specific manner

The primary AC cortex presents a microcolumnar organization which means that neurons located along a radial axis show similar spectral sensitivities (Oonishi & Katsuki, 1965; Merzenich et al., 1975; Jones, 2002). It consists of a prominent layer I; dense and well-developed layers II and III; a granular layer IV (receiving most of the thalamic inputs mentioned before) and layers V and VI (populated by pyramidal cells and small cell bodies respectively) (Winer, 1992, Linden & Schreiner 2003), see Figure 1.6

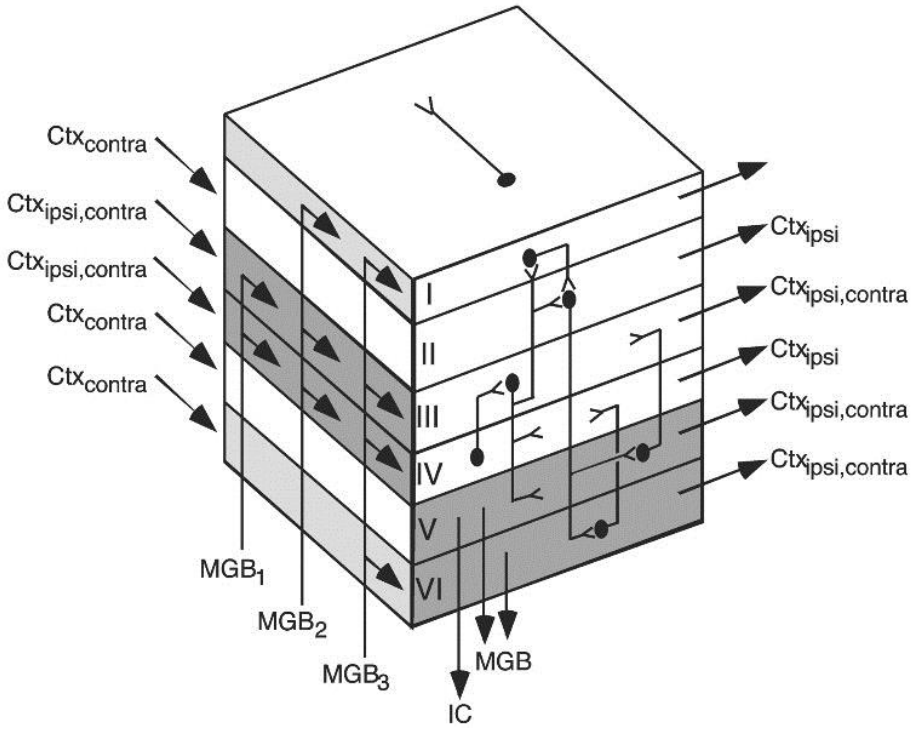

Figure 1.6. Columnar organization of the primary AC. The left cube face shows thalamic (MGB1, MGB2 and MGB3) and cortico-cortical inputs (Ctx_{ipsi} and Ctx_{contra}). The right cube face displays interlayer connections as well as outputs to thalamus (MGB), IC and others cortical structures. Adapted from Linden & Schenider, 2003, used with permission.

Cortical efferent projections from pyramidal neurons in layers V and VI (see Figure 1.6) target IC (Bajo & Moore 2005), SOC (Feliciano et al., 1993; Mulders & Robertson 2000b) and CN (Schofield & Coomes 2005) (see also Saldaña 2016 for a review). The AC efferent projections originating in layers V and VI, project ipsilaterally to the SOC where they contact MOC neurons that innervate the contralateral cochlea (Mulders & Robertson 2000b). These cortical efferent pathways are most likely glutamate-mediated (Potashner et al., 1988; Malmierca & Ryugo, 2010) and are believed to potentially have an inhibitory effect on the cochlear amplifier by exciting MOC neurons (Warren & Liberman, 1989). Nonetheless, it is thought that the auditory cortex can modulate auditory processing at any level of the auditory pathway through these efferent innervations (Coomes & Schofield 2007).

Several researchers have devoted their efforts to establish a physiological link between these auditory nuclei. For example, Leon et al. (2012) showed that the pharmacological deactivation of ongoing AC activity in chinchillas typically reduces the cochlear sensitivity (measured as a reduction in the amplitude of the CM) and this is accompanied by decreases in AN nerve activity (measured as the amplitude of the auditory nerve compound action potentials (CAPs) (gross AN activity). The authors concluded that cortico-olivocochlear circuits modulate both AN and cochlear activity by regulating the auditory efferent baseline activity. Moreover, Dragicevic et al. (2015) performed AC microstimulation in order to evaluate its effects on cochlear sensitivity. The importance of Dragicevic's work, is that the AC microstimulation was evaluated

simultaneously to the sound evoked MOC reflex. In this study, the authors found that microstimulation of the auditory cortex can produce two distinctive modulations of cochlear activity. Similar to Leon et al., modulation of CM and CAPs was evident but also an enhancement or suppression of the MOC reflex strength. Interestingly, in both studies, the authors suggested that corticofugal efferent effects on cochlear sensitivity may not only be explained via MOC activation but also potential AC and IC efferent recruitment of LOC neurons, therefore modulating both OHC and IHC activities (see Figure 1.3. right side).

In human models, Perrot et al. (2006) were able to establish the functional relationship between the AC and the periphery. In this innovative approach, the authors performed electrical stimulation of the AC and non-auditory cortical areas while simultaneously evoking TOAE during refractory epilepsy surgery (Perrot et al., 2006). They found that when the AC was electrically stimulated, it induced a decrease in TOAE amplitude in the contralateral ear. In addition, they found that when non-auditory cortical areas were electrically stimulated, no changes in TOAE magnitude were observed.

1.8 Attentional modulation of MOC effects

Attention has been defined as the brain's ability to prioritize different mental processes to optimize cognitive resources (Knudsen, 2007). The perceptual separation of different sound sources, known as auditory stream segregation, is presumably modulated by attention, yet its role has been extensively debated (Sussman et al, 1999; Carlyon et al, 2001; Macken et al, 2003). That is, it is still unclear to what extent "ignored" stimuli are deprived of attentional resources (Lavie, 2005, Lavie et al., 2014, Molloy et al., 2015).

In hearing research, auditory attention has been described as the ability to direct or focus perception towards an auditory stimulus of interest (Fritz et al., 2007). Auditory attention seems to be a combination of assessing incoming stimuli based on their

saliency (bottom-up) and the re-orientation of cortical processing towards the task (top-down) during scene analysis. Although it is generally well accepted that attention generates activity changes at cortical levels, it is still unclear whether these attentional effects reach lower auditory processing structures and even the auditory periphery.

Hernandez-Peon et al. (1956) described a reduction in CN electrical activity to click stimuli when a cat's attention was diverted to olfactory or visual stimuli. Although the methodology employed by Hernandez-Peon et al. (1956) was subsequently questioned; Oatman (1971, 1976) and Glenn & Oatman (1977) were able to successfully replicate Hernandez-Peon et al.'s findings. Oatman's group showed that AN responses (recorded with implanted round window electrodes) to the auditory stimuli decreased when animals were engaged in visual discrimination tasks.

In humans, Lukas (1980, 1981) hypothesized that attention may play a role in modulating AN responses through the auditory efferent system. His ABR experiments showed that visual attentional tasks can suppress wave I (AN activation) and it was suggested this could only occur through attention-driven efferent feedback from the olivocochlear nuclei. Other authors subsequently explored the potential effects of attention on OAE amplitude during visual tasks (Puel et al., 1988; Froehlich et al., 1990; Avan & Bonfils 1992; Meric & Collet 1994). They were able to observe OAE suppression during visual tasks, presumably caused by an increase in the MOC inhibitory control of the OHC. However, these effects varied across participants. More recently, Delano et al., (2007) measured sound-evoked AN CAPs and CMs in awake chinchillas while performing both visual and auditory tasks. They found decreases in cochlear sensitivity during the visual task but not during the auditory attention task. These results, ruled out potential arousal effects on the cochlear sensitivity and pointed towards visual attention-specific effects.

Studies in subjects with Meniere's disease who underwent vestibular neurectomy (which results in the absence of a functional auditory efferent system innervating the cochlea) have found that some, but not all, participants had difficulties detecting speech in noise (Zeng & Shannon 1994; Giraud et al. 1997; Zeng et al. 2000). Scharf et al. (1994, 1997) behaviourally tested subjects before and after vestibular neurectomy. In particular, Scharf et al., (1997) showed that after surgery, although no hearing impairment was evident either in quiet or noise, participants performed better in a two-interval detection forced choice task at unexpected frequencies (Greenberg & Larking 1968). The authors suspected a MOC-mediated frequency-specific attentional impairment given that improved performance could potentially reduce the focus on target frequency regions. An alternative explanation, given by Tan et al. (2008) who performed similar experiments in normal hearing subjects, was that the MOC effects may either: enhance target detection; help attenuate the non-target signal or even perform both effects simultaneously. In conclusion, these studies suggested that olivocochlear antimasking effects may attenuate the response of the auditory system to the non-target stimuli at their specific stimulus frequencies, leading the target stimulus to be unmasked and easily perceived.

The effects of auditory attention on the MOC inhibition of the cochlear gain in normal hearing subjects remain controversial. For instance, some authors have found decreases in the OAE level while participants actively attend to auditory stimuli (Maison et al., 2001; De Boer & Thornton 2007; Garinis et al., 2011; Smith et al.,2012) while others demonstrate increases in OAE magnitude (Maison et al., 1999; Ferber-Mart et al., 1995; Harkrider & Bowers, 2009). Mishra & Lutman (2014) speculated that these discrepancies may be due to large methodological differences among studies when measuring MOC inhibitory effects on OAEs.

These previous studies on attentional modulation of OAE amplitude by the MOC activity have provided only indirect data of the efferent connection between higher processing structures and the periphery in humans. Nevertheless, in 2001, Khalfa et al. provided functional evidence that cortex and MOC were connected. This study revealed that patients whose temporal brain structures involved in auditory processing had been resectioned show either a decrease in the MOC reflex in the ear contralateral to the temporal resection or an increase in the ear ipsilateral to the surgery.

More recently, Srinivasan et al. (2012), Smith et al. (2012), Walsh et al., 2014a, Walsh et al., 2014b, Wittekindt et al., (2014) and Walsh et al., (2015) explored how both auditory and visual attention may modulate cochlear sensitivity. For instance, Srinivasan et al., (2012) and Smith et al., (2012) found enhanced DPOAE magnitude when participants were attending to visual stimuli while DPOAE magnitude were suppressed when participants attended to an auditory signal. On the other hand, while Wittekindt et al., (2014) found no significant DPOAE magnitude changes during the auditory task, they found significant suppression of DPOAE magnitude during a visual task. As for SFOAEs, Walsh and colleagues (2014a) and (2014b) found that their magnitude is suppressed during both visual and auditory attention. In 2015, the same group measured non-linear SFOAEs during periods of selective auditory or visual attention as well as inattention periods (Walsh et al., 2015); once again, they found suppression of SFOAEs during periods of attention (both auditory and visual) relative to conditions with periods of inattention. Although the results across studies are inconsistent and differ for OAE types explored as well as the attentional tasks tested, they have revealed that cochlear amplification can be attention-modulated via the efferent system.

Interestingly, Wittekindt et al. (2014) also explored the potential interconnectivity between central processing and the auditory periphery. They made the first simultaneous EEG (Auditory Steady State Recordings (ASSRs) and alpha oscillations

(8-14Hz cortical oscillations)) and OAE (DPOAE) recordings during visual and auditory (non-speech) attentional tasks. In general, the authors were able to find effects of selective attention at all levels of the auditory pathway they explored. For instance, physiological responses to acoustic stimuli were enhanced both in the early sensory cortex (increases in ASSR magnitude) and higher cortical areas (increase in alpha power) during periods of auditory attention but were suppressed during visual attention. Although, DPOAE magnitude did not significantly change during auditory attention, it was significantly suppressed during visual attention, lending weight to the hypothesis that the distracting signal (auditory stimuli) is attenuated during visual attention.

In summary, while it has been demonstrated across multiple studies that both cochlear and auditory brainstem activities are reduced during visual attention tasks, the effects of auditory attention remain variable. As suggested by Mishra & Lutman (2014), auditory attention effects on the cochlear gain may depend both on the type of auditory task used as well as methodological flaws regarding, among other, MEMR and the SNRs of OAE recordings. Although several studies have used an auditory lexical decision task to evoke auditory attention (Garinis et al, 2011; Smith & Cone, 2015; Kalaiah et al., 2017a), they have also used white noise to evoke contralateral suppression of OAEs. In these cases, it might be difficult to disambiguate between the lexical decision task *per se* causing the attentional effects on OAEs magnitude and the noise automatically activating the MOC reflex. This could make it hard to disentangle the effects of pure auditory attention Vs. stimulus effects.

In addition, most studies exploring the effects of auditory attention on OAEs magnitude have not matched stimulus and arousal levels across active and passive listening conditions (Froehlich et al., 1990; Meric et al., 1994; Srinivasan et al., 2012). This means that, during the passive listening conditions, participants were either listening to a different stimulus than the one presented during the active listening

condition, and/or not doing any task at all, thereby again confounding the effects of attention, arousal and stimulus differences on the cochlear gain. As such, it is still necessary to disambiguate the effects of attention and stimulus on OAE amplitudes to understand how attention affects the auditory pathway.

1.9 Research Proposal

The main objective of this study is to address key methodological issues that have become apparent with previous attentional studies of OAEs. Moreover, here the relationship between speech intelligibility and the auditory efferent control of the cochlear gain is explored. To address this, speech intelligibility was degraded in three ways: 1) noise vocoding the speech signal; 2) adding babble noise (BN) to the speech signal at different SNRs or 3) adding speech-shaped noise (SSN) to the speech signal at different SNRs. The reason for using noise-vocoded speech while contralaterally recording CEOAEs is that it allowed speech intelligibility to be manipulated without increasing noise levels (a classical way of evoking auditory efferent effects (MOC reflex) (Berlin et al., 1993; Norman & Thornton 1993; Kalaiah et al., 2017b)). This avoided confounding CEOAE magnitude changes due to stimulus-driven MOC reflex activation with attention-driven MOC effects on CEOAE magnitudes. Moreover, because the level of the speech spectrum decreases with increasing frequency, white noise (which is the most commonly used stimulus to evoke MOC reflex in the literature) predominantly masks only the high frequency component of the speech signal, therefore it is not considered an efficient speech masker. However, BN (besides representing a more ethological auditory type of noise) and SSN (which is the spectrally matched long-term averaged of the speech signal) have the same long-term average spectrum as speech. Therefore, these noises were able to mask the speech signal equally across frequencies.

We also attempted to maintain performance (using three difficulty levels: high, medium, and low) across these three speech intelligibility manipulations to determine whether CEOAE suppression was influenced by task difficulty, noise type or a combination of these two factors. The hypothesis is that if suppression of CEOAEs was observed when speech was less intelligible (without increasing noise levels), then it is most likely linked to increases in attentional recruitment of corticofugal inputs rather than a stimulus-specific activation of the MOC reflex. This would allow the differentiation between attention-driven and stimulus-related effects on CEOAE amplitude. Additionally, an auditory lexical decision paradigm was used (similar to Garinis et al, (2011), Smith and Cone (2015) and Kalaiah et al., (2017a)) that required full lexical processing (Goldinger, 1996) and which should both evoke sustained attention and engage cortical speech networks (Hickok & Poeppel 2000; 2007). Attentional engagement was monitored both behaviourally and objectively by measuring task performance and speech-onset auditory cortical activity respectively.

The methodological concerns arising from previous studies of attentional effects on CEOAEs have not only prevented direct comparison of these studies but also distorted the interpretation of their results. Most have implemented paradigms with large differences in their arousal state (or alertness levels (Eysenck, 2012)) and stimulus type between the active auditory task (e.g. speech stimuli presented while CEOAEs are recorded) and passive listening conditions (no task, CEOAEs recorded during no-noise conditions or with-noise conditions) (Froehlich et al., 1990; Meric et al., 1994; Srinivasan et al., 2012). The experimental protocol design in this study will address this issue in three main ways: 1) utilisation of the same stimuli for both active and passive listening conditions; 2) utilisation of a controlled visual scene across the experimental sessions; and 3) attempting to control for differences in alertness during the passive condition by asking subjects to watch an engaging cartoon movie. The homogeneity of visual and

auditory scenes across the experiments allowed the effects of attending to the speech on CEOAE magnitude to be disentangled from the stimulus-driven effects.

To avoid confounding attentional effects with those of MEMR activation (triggered by loud sound levels), which can potentially create unpredictable changes in OAE magnitudes, auditory stimuli levels were kept below those for MEMR activation and speech intelligibility was manipulated without increasing overall stimulus levels (above the MEMR threshold). Moreover, a minimum OAE SNR was kept at 6 dB which allowed for more reliable and replicable OAE recordings (Goodman et al., 2013).

A simultaneous OAE/EEG recording setup was implemented similar to Wittekindt et al. (2014), ensuring that recordings from the periphery and central auditory pathway were performed concurrently. Because of the simultaneous nature of this setup, the same click stimuli that evoked OAEs could also evoke auditory brainstem responses (ABRs) and this was recorded with the EEG system. Here, it could be argued that if CEOAE suppression was observed, then a reduction in the ABR components should also be detected, since the cochlear gain to the click stimulus would be reduced across the entire auditory afferent pathway (which can be assessed using ABRs). Therefore, in the context of this thesis, the magnitude of click-evoked ABRs (i.e. their wave amplitudes) were also used as an independent measure to confirm auditory efferent suppression of the cochlear gain. Together, this experimental setup allowed us to disentangle auditory attention-driven efferent effects from automatic stimulus-specific effects along the auditory pathway during speech perception.

1.10 Research question, hypothesis and objectives

In this PhD project, we sought to explore how active listening to degraded speech modulates CEOAE magnitudes. In particular, unlike previously published studies, we assessed auditory changes objectively and subjectively as part of a highly controlled experimental paradigm, maintaining a constant performance across three experimental

manipulations of speech intelligibility as well as minimizing influences of MEMR activation and controlling for visual and auditory scenes.

Research question: Does auditory attention modulate cochlear gain, via the auditory efferent system, in a task-dependent manner?

Hypothesis: Decreases in speech intelligibility raise auditory attention and this reduces cochlear gain (measured using CEOAEs).

The objectives are:

1) to evaluate lexical decision performance with changes in speech intelligibility.
2) to analyze speech-onset event related potentials (ERPs) to confirm auditory cortex engagement.
3) to analyze click-evoked ABRs waves magnitude to confirm the effects of cochlear gain reduction in the afferent auditory pathway.
4) to determine the changes in CEOAE amplitude with decreases in speech intelligibility during active and passive listening.

To investigate whether attention to degraded speech influences the cochlear gain, auditory attention was modulated by manipulating task difficulty for a lexical decision task while simultaneously measuring OAEs and ABRs to click stimuli. Auditory attention was objectively monitored, by recording cortical speech-onset ERPs, and subjectively assessed by evaluating performance during an auditory lexical decision task. Changes in brainstem activity and the cochlear gain were measured by comparing ABR and OAE waveforms across active and passive conditions and across the varying task difficulties. All measures were simultaneously obtained.

This study was approved by the Human Research Ethics Committee of Macquarie University (ref: 5201500235), see Appendix section. Each participant signed a written informed consent form and was provided with a small financial remuneration for their time.

2.1 Hearing Screening

All 66 participants included in the study had normal pure tone thresholds (<20 dB HL tested at octave intervals between 0.5-8 kHz using a standard Hughson and Westlake technique (Hughson & Westlake 1944); Interacoustics Hearing Aid Fitting Analyzer Affinity 2.0 Audiometry) and normal middle ear function (assessed using otoscopy and standard 226 Hz tympanometry; Titan, Interacoustics). To ensure normal outer hair cell function, DPOAEs were demonstrated in all participants between 0.5-10 kHz (DPOAE properties: f1/f2 =1.2, f1 =65 dB SPL; f2 =55 dB SPL; classified by SNR>6 dB; minimum DPOAE level above -10 dB SPL; reliability >98%; Interacoustics Titan – DPOAE440). To confirm that the levels of speech and click stimuli used (75 dB SPL) would not elicit an MEMR, the MEMR threshold was assessed for steady-state BBNs both contralaterally and ipsilaterally (Titan, Interacoustics) and was found to be greater than 75 dB HL in all participants. According to the ANSI S3.6-1996 standard for the

conversion of dB SPL to dB HL, at least a 10 dB difference was kept between MEMR thresholds and our experimental stimulus level. Although we cannot fully rule out any subtle effects of the middle ear muscles to subthreshold stimulus levels, obtaining MEMR thresholds to BBN in all participants minimised the possibility that MEMR activation was a cause of any changes in OAE amplitudes.

2.2 Experimental Protocol

Participants were seated in a comfortable chair inside a sound-proof booth (ISO 8253-1:2010). The experimental protocol comprised two experimental conditions Passive and Active listening, presented in counterbalanced order across participants. In the Passive listening condition, subjects were requested to ignore the auditory stimuli presented and to watch a soundless movie. To ensure they had paid attention to the movie, participants were monitored during the experiment with a video camera located in the experimental booth and then, at the end of the experimental condition, they were asked questions about the movie's content (e.g. what happened in the movie? How many characters were present? Who was the main character?).

During the Active listening condition, the participants performed an auditory lexical decision task (McCusker et al., 1981; Marslen-Wilson, 1980), where they were asked to press the space key on a keyboard each time they heard a word that did not make sense (i.e. a non-word). Behavioural responses were obtained for each subject across all experimental conditions; these included the number of correct responses (hit rate (HR)), the number of incorrect responses (false alarm rate (FA)) and the reaction times to the correct (hits) and incorrect (false alarm) responses. The d prime, calculated as: Z(hit rate) – Z(false alarm rate), was used as a measure of response accuracy.

2.3 Speech Stimuli

All word items were acquired from Australian-English adapted versions of monosyllabic consonant-nucleus-consonant (CNC) word lists (Lehiste & Peterson, 1959; Henry et al., 1998) and were spoken by a female native, Australian-English speaker. The duration of all words ranged between 420-650ms. Non-word lists were selected from the Australian Research Council (ARC) non-word database (Rastle et al., 2002). The following criteria were used to form the non-word lists: 1) each non-word list matched the phoneme frequency and position (beginning, middle, or end of words) of a corresponding real-word list; 2) non-words were phonetically legal and 3) non-word lists had to match the lexical neighbourhood density of their corresponding real-word list (the frequency of real-words that are acoustically similar to a specific real- or non-word (Luce & Pisoni, 1998)).

Speech tokens were normalized to have the same root mean square (RMS) level. The presentation system was calibrated using a pink noise with the same RMS level as the experimental items. The ER-3C headphone was placed in an artificial ear (IEC 60711 Ear Simulator RA 0045) coupled to a sound level meter (B&K G-4 type 2250, amplifier type 4189). Speech stimuli were delivered using Presentation® software (Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com, version 18.1.03.31.15). All speech stimuli were digitized at 44.1 kHz, 16-bits.

2.4 EEG recording and analysis

The EEG data were acquired using 64 channels (Ag-AgCl electrodes) arranged according to the 10–20 system. Vertical eye movements were monitored by placing electrodes above and below the left eye. Horizontal eye movements were monitored by placing electrodes on the outer canthi of each eye. Impedance levels were kept below 5kΩ for all electrodes. Signals were sampled at a rate of 20 kHz in AC mode with a gain of 2010 and an accuracy of 0.15 nV/LSB (SynAmps2 amplifier, Compumedics Limited).

An online filtering for visualization purposes (high- and low-pass filters of 0.05 Hz-3.5 kHz respectively were applied).

2.4.1 Speech-onset ERPs recording and analysis

The off-line analysis of EEG recordings was performed using Fieldtrip (open source Matlab toolbox developed by Oostenveld, et al. 2010). The data were re-referenced to the average of the left (M1) and the right (M2) mastoids. The EEG recordings were then divided into trials/epochs that started 200 ms prior to the speech tokens onset (both words and non-words) and ended 1.2 sec after the sound onset (-200 ms to 1.2 sec). The components visually identified as eye blinks and horizontal eye movement were excluded from the data. In order to remove noisy trials, those trials between -200 ms and 1.2 sec from sound onset which had absolute amplitude values higher than 75 μV were excluded from further analysis, leaving between 60-80% of total trials per condition.

The accepted trials were then band-pass filtered with a frequency cut-off between 0.5 to 30 Hz and a transition band roll-off of 12 dB/octave. Band-pass filtering between those frequencies can help to remove both common artefacts such as line noise 'humming' of the electric power supply as well non-neural and irrelevant information, therefore improving the SNR (Luck, 2005; Widmann et al., 2014; Acunzo et al., 2012). The resulting output signals were baseline-corrected using the mean value of a window between -200 ms and 0 ms from the onset of the speech tokens. The baseline-corrected trials were averaged to obtain the ERP waveforms. Rather than using the absolute magnitude of each ERP component at its maximum, the mean values for fixed-latency, 'analysis windows' centred on these maximums were selected (Woldorff et al., 1993; Luck, 2004; Woodman, 2010). These 'analysis windows' were visually inspected and chosen based on the grand average ERP waveforms for each condition. Because attentional effects had been previously reported on the early ERP components (Woldorf

et al., 1987, 1993; Karns & Knight, 2009; Saupe et al., 2009), the components analysed were: P1 ('analysis window': 90-130 ms), N1 ('analysis window':130-170 ms), P2 ('analysis window': 220-260 ms). Here, only the mean amplitude was analysed because it has been reported to be the most reliable measurement when analysing cognitive effects on ERPs (Clayson et al., 2013; Keil et al., 2014; Luck & Gaspelin 2017).

In addition, the auditory lexical decision task requires semantic processing (Goldinger, 1996) and the N400 (550-590 ms) component has been associated with the processing of meaning (Kutas & Federmeier, 2011): therefore, this component was also examined. Then the mean amplitude for each component between the given latencies was calculated using a Fieldtrip-based script. These were then generated for each participant and experimental condition for posterior statistical analysis.

2.4.2 ABRs recording and analysis

The offline analysis of ABR signals was performed using the same Fieldtrip toolbox. The data were re-referenced to the mastoid with the smallest noise magnitude (left mastoid (M1) or right mastoid (M2)). Only ABR signals from central electrodes (FZ, FCZ, CZ) were analysed given their close-to-vertex position where auditory responses have larger magnitude (Jewett et al., 1970; Vaughan & Ritter, 1970). The EEG signal was then divided into trials/epochs that started 5 ms prior to the click stimuli and ended 10 ms after the click onsets for a total analysis window of 15 ms. All trials were band-pass filtered with a frequency cut-off between 200-3000 Hz. These cut-off filters have been shown to provide clean ABRs waveforms and do not distort the different wave amplitudes or latency measurements (Pratt & Sohmer, 1976; Boston & Ainslie 1980). In addition, Bidelman et al., 2014 reported that filters with cut-off frequencies higher than 200 Hz may minimize post-auricular muscle artefact (O'Beirne & Patuzzi, 1999). The processed traces were then averaged using a weighted averaging method (Don & Elberling 1994; Silva 2009) to give a final averaged ABR waveform per participant and

condition consisting of 19200 trials (which corresponds to the click rate of 32 Hz (for Experiments 1-3) across 10 minutes of stimulation for all conditions). Notice that these same clicks allowed us to obtain CEOAEs. The weighting averaged method in this study is based on Bayesian inference and weights the individual sweeps proportionally to their estimated precision calculated as the residual noise for every 32 sweeps (Box & Tiao, 1973). All sweeps were then rescaled with a factor inversely proportional to this variance or residual noise (Elberling & Don, 1984). This helps preserve all trials without any rejection required for artefacts. All ABR components analysed were visually selected only when they had a significantly higher magnitude than the residual noise calculated for each epoch (positive SNR). ABR waveforms with residual noise equal to or higher than the averaged signal were discarded, see Appendix section Tables 1-4.

Due to restrictions in click-stimulus level (lower or equal to 75 p-p dB SPL to avoid activation of the MEMR), it was not possible to extract wave I from the EEG residual noise, see Figure 2.1. Instead, higher amplitude waves associated with activity in the SOC (wave III) and IC (wave V) only were analysed, see Figure 2.1. Amplitude and latency of waves III and V were visually determined for each subject across blocks and conditions when observable with a positive SNR (waves' amplitudes above the dotted red lines in Figure 2.1). In addition, ABR waves V-III amplitude ratios and latency differences were calculated. These components allowed the observation of potential changes at the level of the SOC (wave III) and IC (wave V) (Jewett & Williston, 1971; Pratt & Sohmer, 1976; Suzuki et al., 1986) during Active and Passive listening of speech.

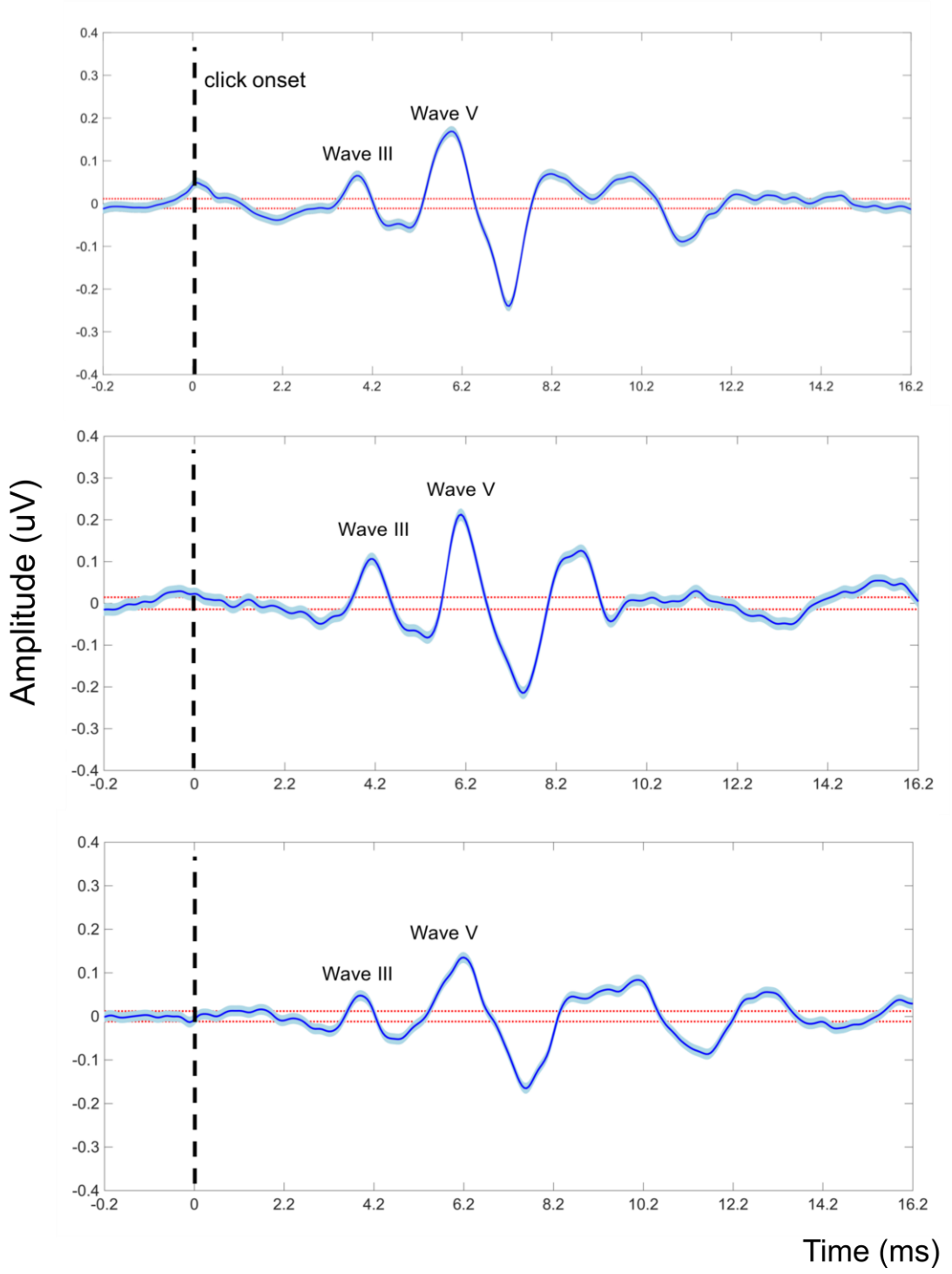

Figure 2.1. Example of an averaged ABR waveform from three participants during the Passive listening of Natural speech. The black dashed line represents onset of the click stimulus. The solid dark blue line corresponds to the ABR waveform while the variance is represented by the shaded blue area. In addition, the red dotted lines exemplify the (±) residual noise calculated across 19200 trials. Only the waves that showed positive SNRs with respect to the residual noise were selected for statistical analysis.

2.5 OAE recordings and analysis

Non-filtered click stimuli, with a positive polarity and 83 μsec of duration were digitally generated using RecordAppX (Advanced Medical Diagnostic Systems) software. The presentation rate was 42 Hz in the pilot experiment for Experiment 1 and 32 Hz in Experiments 1-3. The linear method was used for click presentation because it allows the linear and non-linear components of MOC effects to be maintained in the OAE magnitudes (Guinan, 2006; Mishra & Lutman 2014). Both the generation of clicks and the recording of OAEs were controlled via an RME UCX soundcard (RME, Haimhausen, Germany), and delivered/collected to and from the ear canal through an Etymotic ER-10B probe connected to ER-2 insert earphones with pre-amplifier gain set at 20 dB (Etymotic Research, Elk Grove Village, IL).

Stimulus calibration is a critical step in measuring MOC effects on CEOAE magnitude as: 1) efferent effects on OAE magnitude are level dependent (Collet et al., 1990; Guinan, 2006) and 2) they only produce small magnitude changes of around 1-3 dB (Goodman et al., 2013): therefore, by calibrating the stimulus, a study can ensure accurate measurements of OAE magnitude changes without the confounding effects of stimulus level differences caused by differing earphone insertion depth within and across participants. However, traditional calibration of signals, performed in standardized couplers (ANSI, 1995; International Organization for Standardization (ISO), 1997; ISO, 2006) typically does not account for individual differences in ear canal anatomy and geometry as the couplers are considered as surrogates for the average ear canal thanks to their similar terminal impedances. In those studies, where stimuli have been calibrated for individual ear canals, it has generally been performed by adjusting the voltage drive to the speakers based on the pressure response detected by the probe microphone; therefore, relying on the basic assumption of equivalent sound

pressure at all frequencies between the probe microphone and the surface of interest – the ear drum. However, as a result of sounds reflections in the external ear canals, this can underestimate the pressure at the tympanic membrane by up to 2-3 dB for low frequencies (1-2 kHz) and 20 dB at frequencies greater than 3-4 kHz (Siegel 1994; Souza et al., 2014). To avoid these issues in this study, stimuli used to evoke OAEs were calibrated in forward equivalent pressure level (the pressure travelling towards the tympanic membrane: FPL) determined using Thévenin-equivalent source parameters of the probe that has been shown to decrease the insertion dependent effects on both DPOAEs (Scheperle et al., 2008) and CEOAEs (Scheperle et al., 2011).

Stimulus levels were verified using a sound level meter (B&K G4) and microphone IEC 60711 Ear Simulator RA 0045. This setup was used to calibrate the click stimuli following BS EN 60645-3:2007. In-ear OAE probe calibration was completed at the beginning of every experimental block. In addition, if during the experimental blocks participants moved or touched the OAE probe, the block was stopped, and the probe was re-positioned and re-calibrated.

Off-line processing with a Matlab-based script allowed for signal reliability and artefact removal to be checked. The signal was band-pass filtered between 750-6000 Hz with 4^{th} and 10^{th} order Butterworth filters respectively. Epochs with RMS levels two standard deviations above or below the mean RMS amplitude of all epochs were discarded. The reliability of the CEOAE recordings was assessed by correlating consecutive epochs (odd and even epochs). Only epochs with a correlation higher or equal to 95% were accepted as valid signals for further analysis. The noise floor was estimated by subtracting the RMS difference between the grand response mean (which included means from both odd and even epochs) from even and odd epochs independently (Schimmel, 1967; Boothalingam et al., 2015).

The CEOAE levels were calculated as the mean RMS amplitudes in a time window 10-23 ms after click onsets; the initial 0-9 ms time window after the click stimulus being rejected due to cochlear ringing, see Figure 2.2. Data were binned in one-minute steps by averaging epochs contained within that period. The averaged RMS magnitudes of CEOAE signals between 1-2 kHz were analysed due to the known MOC effects being maximal in this frequency band (Collet et al., 1990; Hood et al. 1996; Goodman et al. 2013). Only binned data for averaged CEOAEs displaying a SNR ≥ 6 dB (Mishra & Lutman 2013; Mertes & Goodman 2016) and with > 80% of epochs retained (i.e. had RMS levels within the two standard deviations limit), were selected as valid signals for further analysis.

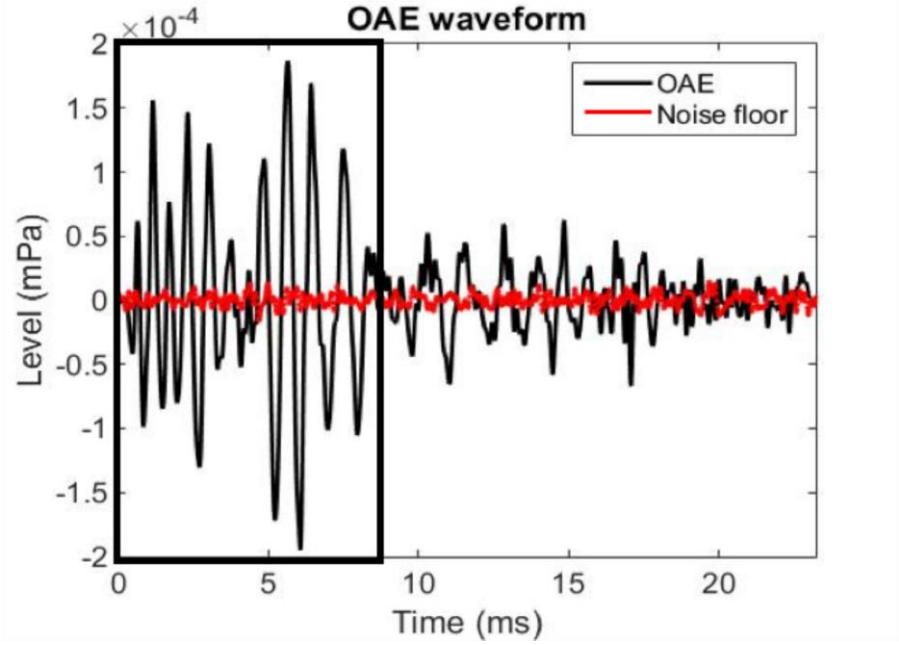

Figure 2.2. Time course of the CEOAE analysis window. The black trace represents the CEOAE signal while the red trace corresponds to the noise floor. The click stimulus ringing duration was 9 ms (shown in the black rectangle). This time window was removed prior to CEOAE offline analysis to avoid stimulus contamination of the CEOAE signal.

2.6 Synchronization Equipment

Presentation software and the OAE setup were synchronized by Stimtracker (Cedrus) which converted the click stimulus into Transistor–Transistor Logic signals (TTLs) that were recognized by Presentation software as an external input. After receiving the first TTL (associated with click stimulus number 1), Presentation software initiated the experimental paradigm. Presentation software and Neuroscan system were synchronized through a DB-25 parallel port connection. This setup allowed the simultaneous recording of CEOAEs, ABRs, ERPs and behavioural responses, see Figure 2.3.

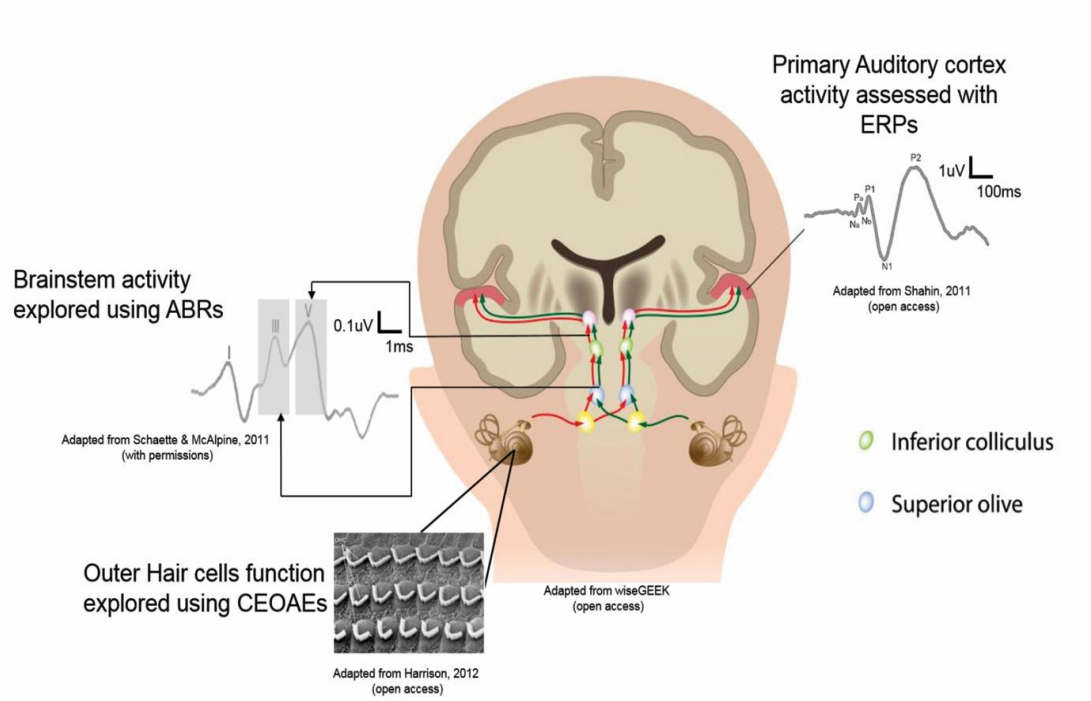

Figure 2.3. Schematic representation of the simultaneous recording of the physiological measurements: CEOAEs, ABRs and ERPs.

2.7 Specific Methods

Manipulations of the intelligibility of the speech stimuli were performed to increase task difficulty and modulate attention. Three stimulus manipulations were made: 1)

channel vocoding of the speech signal, 2) adding BN to the Natural spoken speech stimuli and 3) adding SSN to the Natural spoken speech stimuli. BN was filtered to match the envelope of the SSN and to match the International Long-Term Average Speech Spectrum (ILTASS) (Byrne et al., 1994). The ear receiving either the click or speech stimulus was randomized across participants.

2.7.1 Experiment 1. Noise Vocoded Speech: Pilot study

To determine the noise vocoded speech levels required to create significantly different levels of task difficulty, a pilot study was conducted with 5 native speakers of Australian-English with normal hearing (3 females; mean age = 30 ± 3 years old). The pilot study also enabled useful identification of potentially problematic speech tokens, as well an evaluation of the feasibility of simultaneous recording of CEOAEs, ABRs, ERPs and behavioural responses. Participants followed the **Experimental Protocol** procedure explained above. In the Passive listening conditions, participants watched a movie of their choice (subtitled); while in the Active listening conditions, participants were requested to watch a fixation cross displayed on a screen located 2 meters away from them.

The pilot study comprised 4 blocks per condition (each block with a duration of 6 min, see Figure 2.4): Block 1: Natural spoken tokens and Blocks 2, 3 and 4 corresponded to speech with different number of vocoded channels. Following the method of Shannon et al. (1995), the noise vocoded speech technique was employed. The speech tokens were band-pass filtered using a zero-phase, sixth-order Butterworth filter and their presentation constituted 5 minutes of each block duration, see Figure 2.4. The centre frequencies of the channels were equally spaced on a Greenwood scale from 200 to 8000 Hz (Greenwood, 1990). Speech envelopes were extracted by half-wave rectifying the filter outputs and low-pass filtering with a zero-phase, second-order Butterworth filter with a cut-off frequency equal to 300 Hz or half of the channel

bandwidth (whichever was lower). The extracted envelopes were used to modulate noise bands produced by filtering white noise with the same band-pass filter. The modulated bands were summed to give the final vocoded signal.

The order of experimental conditions was counter-balanced among subjects (i.e. Active and Passive listening conditions were alternated). The presentation of the blocks was also randomized as were the auditory stimuli within, while ensuring that the latter were only repeated ≤ 3 times in each Experimental session. These controls were not only applied to the Pilot studies but also to the main Experiments, making it less likely that our results were biased due to presentation order or training effects, (Quinn & Keough, 2002; Kirk, 2007). Simultaneous to the unilateral presentation of word/non-word, CEOAEs were recorded continuously in the contralateral ear (see Figure 2.4). The total duration of the pilot experimental session was 48 min. Participants were allowed to take short breaks between block presentations.

Based on previous behavioural studies (Shannon et al., 1995; Shannon et al., 1998), three vocoded conditions (16-, 12- and 8-channels: Voc16, Voc12 and Voc8 respectively) were piloted to represent three degrees of speech intelligibility (i.e. task difficulty; see Table 2.1). The HR – the proportion of non-words correctly recognised- decreased with the number of vocoded channels; while the false alarm rate - the proportion of words incorrectly identified as non-words increased as the number of vocoded channels decreased. Moreover, performance (d prime = zHR-zFA) was poorer as the number of noise vocoded channels decreased. Although d prime values did not appear different between Voc16 and Voc12, HR and FA decreased between Voc16 and Voc12, see Table 2.1. Therefore, the four task difficulty levels were selected for inclusion in Experiment 1.

CEOAEs were continuously recorded throughout each 6-minute block. The initial and final 30 secs of CEOAE recordings of each block, in which no speech tokens were

presented, see Figure 2.4, were taken as a baseline. During the middle five minutes, 88 words and 88 non-words were presented, and subjects made their lexical decisions, see Figure 2.4. Although, a drift in CEOAE magnitudes over time has been described, making MOC effects appear more or less significant (Goodman et al., 2013), it has been suggested that randomly interleaving noise and no-noise conditions are an important control for drifts (Guinan 2017). Here the contralateral speech stimulation was not continuous because a 2 second inter-stimulus interval was established to accommodate participants' behavioural response. Moreover, it has been reported that these drifts may cancel each other out in group data (Goodman et al., 2013); therefore, it is suggested that the influence of drift in this data set would be small or non-existent.

CEOAE magnitude change was considered as:

$$\text{CEOAE suppression} = \text{CEOAE}_{\text{speech presentation (average across minutes)}} - \text{CEOAE}_{\text{baseline (first 30 sec)}}$$

The suppression of CEOAE magnitude relative to the baseline allowed us to observe potential auditory efferent effects on the cochlear gain via MOC activation during the Active and Passive listening of speech.

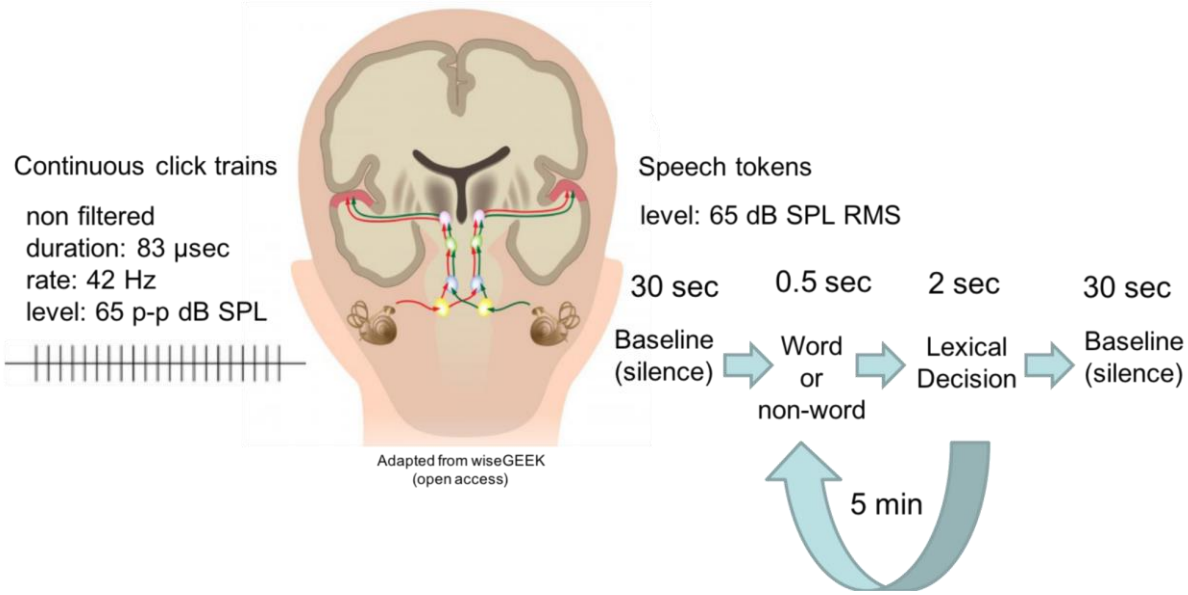

Figure 2.4. Schematic experimental design of the stimulus presentation for the pilot studies. Note that in one of the ears the click stimuli were continuously presented in order to evoke CEOAEs while in the other ear words and non-words were delivered. The differences between pilot experiments resided in the speech intelligibility manipulation explored (**Experiment 1**: Noise Vocoded Speech, **Experiment 2**: Babble Noise and **Experiment 3**: Speech-Shaped Noise.

Table 2.1. Pilot Experiment 1. Noise Vocoded. Behavioural data

Stimuli	Hit Rate	False Alarm Rate	d prime
Natural	0.88±0.06	0.05±0.03	2.93±0.29
Voc16	0.81±0.10	0.14±0.06	2.02±0.35
Voc12	0.78±0.09	0.11±0.05	2.04±0.42
Voc8	0.72±0.19	0.17±0.08	1.63±0.28

Problematic speech tokens which were identified in the pilot were removed (i.e. words with low frequency based on English Subtitles Database: SUBTLEX-UK (Van Heuven et al., 2013) and non-words with ambiguous pronunciation or pronunciation similar to words).

Although both CEOAEs and ABRs were successfully obtained for click stimuli at 65 p-p dB SPL and 42 Hz stimulus rate, the SNR for both signals was small. Therefore, the stimulation level of the click stimuli was increased by 10 dB and the stimulation rate was decreased by 10 Hz to obtain a clearer ABR waveform (Jewett & Williston, 1971; Pratt & Sohmer, 1976; Suzuki et al., 1986). The speech stimulus level was also increased by 10 dB accordingly**.**

2.7.2 Experiment 1. Noise Vocoded: Main study

Thirty participants were recruited to participate in the study. However, three participants were withdrawn from the final dataset: one due to inattention during the lexical decision task; one due to technical difficulties and the third one chose not to complete the experimental session. Twenty-seven native speakers of Australian-English (17 females, 25 right handed) comprised the final sample. Their ages ranged between 18-35 years old (mean age = 23 ± 5 years old). From a total of 500 words and 442 non-words, 423 words and 328 non-words tokens were selected as appropriate for this study. Consequently, 200 words and 100 non-words were randomly presented in each block across 10 minutes, see Figure 2.5. All the tokens were compensated in each block for stop and non-stop initial consonants: 100 words stop / 100 words non-stop consonants; 50 non-words stop / 50 non-words non-stop consonants. The same speech tokens were used for Experiments 1, 2 and 3.

The same four blocks piloted were presented per experimental condition: Block1: Natural spoken tokens; Block 2: Voc16; Block 3: Voc12 and Block 4: Voc8. Each block lasted for 12 minutes and the block's presentation order was randomized. Simultaneous to the unilateral presentation of words/non-words, CEOAEs were recorded continuously in the contralateral ear, see Figure 2.5, where the first 60 secs were taken as CEOAE baseline. During the Passive listening conditions, participants were asked to pay attention to a non-subtitled and soundless video (Shawn the Sheep seasons 1 and 2: https://shaunthesheep.com/); while in the Active listening condition, the video continued to be played, but participants were instructed to perform the lexical decision task. In contrast to the experimental design for Experiment 1 Pilot, this allowed a consistent visual scene across experimental conditions and without the confounding factor of reading processing occurring concurrently to the auditory processing. Eight blocks (4 blocks in the Active and 4 blocks in the Passive listening condition) of 12 min duration

were presented to the 27 subjects who participated in Experiments 1 (Natural, Voc16, Voc12 and Voc8). All blocks were presented in a unique experimental session that lasted 2.5 hours (including hearing screening and EEG cap set-up: 1 hour). Participants were allowed to take short breaks between block presentations.

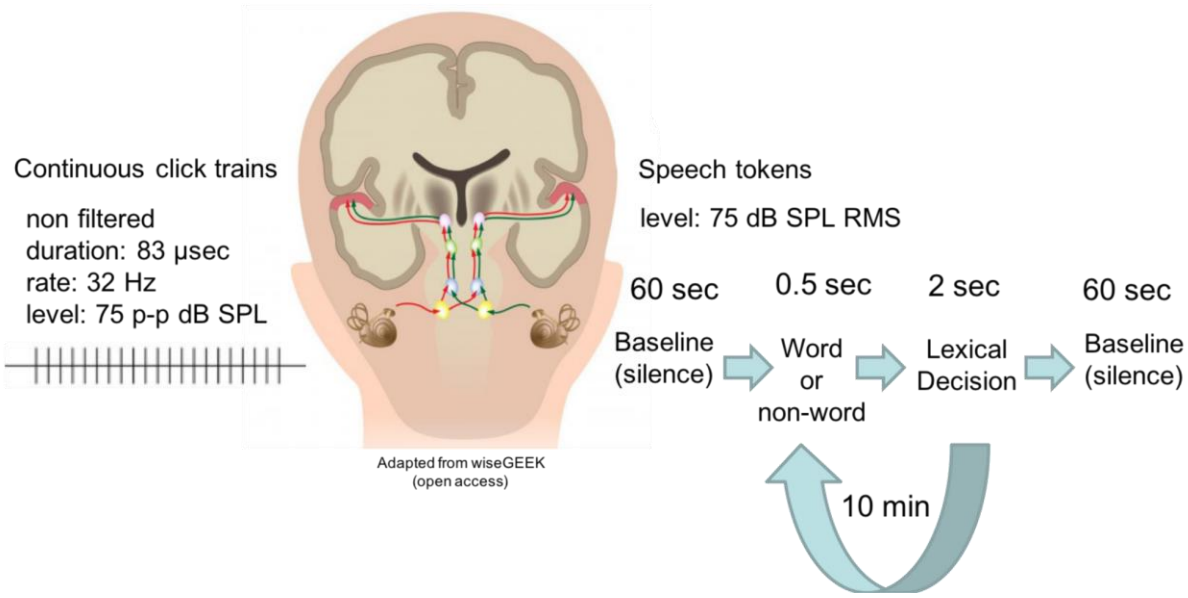

Figure 2.5. Schematic representation of the experimental design in Experiments 1, 2 and 3. Note that in one of the ears, the click stimuli were continuously presented in order to evoke CEOAEs while in the other the speech tokens were delivered. The differences between experiments resided in the speech intelligibility manipulations explored (**Experiment 1**: Noise Vocoded Speech, **Experiment 2**: Babble Noise and **Experiment 3**: Speech-Shaped Noise.

2.7.3 Experiment 2. Babble Noise: Pilot and main study

The BN used consisted of four females and four male talkers speaking simultaneously and recorded in an anechoic chamber (Keidser et al., 2002). The relative intensities of each one third octave frequency of the speech signals matched the average relative intensities of speech produced by male and female talkers. Random

segments from a 2-min recording of BN were randomly selected in order to create the speech tokens in BN. The speech tokens and the BN had the same durations.

Three different speech in BN ratios were piloted (speech in BN +10 dB SNR (BN10); speech in BN +7 dB SNR (BN7) and speech in BN +5 dB SNR (BN5), see Table 2.2). Six normal hearing, Australian-English speakers (five of whom participated in Experiment 1 Pilot study; 3 females; mean age = 33 ± 5 years old), were tested to select SNRs where the behavioural results could be comparable with the lexical decision performance in Experiment 1. The pilot experimental protocol was similar to the protocol described for Experiment 1.

Table 2.2. Pilot Experiment 2. Babble Noise. Behavioural data

Stimuli	Hit Rate	False Alarm Rate	d prime
Natural	0.81±0.14	0.07 ±0.04	2.53±0.54
BN10	0.59± 0.14	0.09 ±0.04	1.62±0.44
BN7	0.57 ± 0.16	0.09±0.06	1.57±0.48
BN5	0.54 ± 0.15	0.13±0.06	1.28±0.46

The SNRs for BN that produced behavioural data comparable to the data obtained for Voc8 speech and the averaged behavioural data for Voc16 and Voc12 were BN5 and BN10 respectively. Therefore, three blocks presented in the Active and Passive listening conditions were: Block1: Natural spoken tokens; Block 2: BN10 and Block 3: BN5.

A new sample of thirty participants were recruited to participate in the main Experiments 2 and 3. However, due to incomplete data collection in one participant, twenty-nine native speakers of Australian-English (19 females, 28 right handed) participated in these experiments. Their ages ranged between 20-35 years old (mean age = 26 ± 9 years old). Because Natural speech tokens (Block 1) for Active and Passive listening were shared blocks in both Experiments 2 and 3, they were only presented once to participants involved in these experiments.

2.7.4 Experiment 3. SSN: Pilot and main study

SSN has been used as a successful masker in studies of speech perception (Nelson et al., 2003; Qin & Oxenham, 2003). To produce the SSN, the spectrally matched long-term average of all the Natural speech tokens (words and non-words) was obtained. Segments from the 2 min SSN were randomly selected and added to the speech tokens at the desired SNR to create the stimuli for the SSN conditions. The speech tokens and the SSN had the same durations. All the stimuli were digitalized at 44.1 kHz, 16-bits and RMS level-normalized across all tokens.

Several speech in SSN ratios were piloted (speech in SSN +8 dB SNR (SSN8); speech in SSN +5 dB SNR (SSN5) and speech in SSN +3 dB SNR (SSN3), see Table 2.3) in the same group of 6 normal hearing, Australian-English speakers as piloted in Experiment 2. This was in order to select an SNR that provided comparable behaviour results with that obtained in Experiment 1. The pilot experimental protocol was similar to the protocol described for Experiment 1.

Table 2.3. Pilot Experiment 3. Speech-Shaped Noise. Behavioural data

Stimuli	**Hit Rate**	**False Alarm Rate**	**d prime**

Natural	0.81±0.14	0.07 ±0.04	2.53±0.54
SSN8	0.62±0.16	0.09 ±0.04	1.74 ± 0.40
SSN5	0.50 ± 0.15	0.10 ±0.06	1.32 ± 0.34
SSN3	0.52 ± 0.13	0.15 ±0.07	1.22 ± 0.34

The SNR for SSN that produced behavioural data comparable to the data obtained for Voc8 speech and averaged behavioural data for Voc16 and Voc12 channels speech was SSN3 and SSN8 respectively**.** Because Experiments 2 and 3 were performed in the same experimental session, Block1: Natural spoken tokens were the same for both experiments. Experiment 3 also consisted of: Block 2: SSN8 and Block 3: SSN3 presented in the Active and Passive listening conditions.

Similar to Experiment 1, the order of the conditions was counterbalanced, and, within each condition, the blocks' order was randomized. Ten blocks (5 blocks in the Active and 5 blocks in the Passive listening condition) of 12 min duration were presented to the 29 subjects who participated in Experiments 2 and 3 (Natural, BBN10, BBN5, SSN8 and SSN3). All blocks were presented in a unique experimental session that lasted 3 hours (including hearing screening and EEG cap set-up: 1 hour). Participants were allowed to take short breaks between block presentations.

2.8 Statistical Analysis

All behavioural, CEOAE and electrophysiological data were tested for normality using Shapiro-Wilk test. A parametric (ANOVA and paired t-test) or non- parametric analysis (Kruskall-Wallis and Mann-Whitney U test) was performed accordingly in order to compare the effects between experimental conditions (alpha=0.05, with Bonferroni

corrections for multiple comparisons). Nevertheless, the specific analysis performed in each section will be detailed in the Results chapter. The effects of attention were assessed by comparing Active Vs. Passive listening conditions. The effects of speech intelligibility manipulation were evaluated by comparing within each experiment the variables obtained for each stimulus type. Effects of task difficulty were calculated by observing the interactions between the effects of attention and stimulus type.

Chapter 3 Results

In this chapter, the results and statistical analyses for behavioural and physiological measurements (performed across all three experiments) are presented. This includes examination of subjective (performance) and objective (cortical speech onset-ERPs) variables, which were used to confirm sustained auditory attention across the study. Moreover, the effects of attention and the methods used to degrade speech intelligibility were also examined at the auditory brainstem level (click-evoked ABRs and OAEs).

3.1 Performance across 3 different speech intelligibility manipulations

Speech stimuli were piloted (see Section 2.7.1) and selected to ensure that significant differences existed in lexical decision task performance and attentional allocation across all conditions: Natural speech, 16 channels (Voc16), 12 channels (Voc12) and 8 channels (Voc8) noise vocoded speech. Mean (+/- standard deviation; see Table 3.1) results for 27 normal hearing participants for lexical decision task and all stimulus conditions were calculated including HR, FA, d prime, and reaction time for the hits (RT Hit) and false alarm (RT FA) responses.

Table 3.1. Experiment 1 (Noise Vocoded). Lexical decision task behavioural results (mean +/- standard deviation)

Stimuli	Hit Rate	False Alarm Rate	d prime	RT Hit (sec)	RT FA (sec)
Natural	0.76 ± 0.18	0.05 ± 0.03	2.58 ± 0.60	1.15 ± 0.08	1.12 ± 0.16
Voc16	0.66 ± 0.20	0.10 ± 0.07	1.87 ± 0.65	1.18 ± 0.09	1.14 ± 0.15
Voc12	0.63 ± 0.18	0.11 ± 0.05	1.66 ± 0.57	1.19 ± 0.09	1.15 ± 0.12
Voc8	0.55 ± 0.19	0.13 ± 0.08	1.30 ± 0.58	1.19 ± 0.09	1.17 ± 0.12

Overall, all HR residuals were normally distributed (Shapiro-Wilk test of normality, $p>0.05$). A one-way ANOVA revealed significant main effects: $F(3, 78) = 35.81$, $p = 0.0001$. A post-hoc t-test pairwise comparison with Bonferroni correction showed that all HR comparisons between the Natural stimuli and the vocoded stimuli were significantly different: HR Natural Vs. HR Voc16 ($p=0.0001$); HR Natural Vs. HR Voc12 ($p=0.0001$); HR Natural Vs. HR Voc8 ($p=0.0001$). While no statistical differences were found between HR Voc16 and HR Voc12: ($p=0.61$), HR Voc16 and HR Voc12 were significantly higher than HR Voc8 ($p=0.0001$ and ($p=0.001$), respectively. This suggests that greater intelligibility of the speech stimuli resulted in significantly better classification of non-words for all conditions, except for Voc16/Voc12.

The residuals of the false-alarm rate data were tested for normality and most variables did not follow a normal distribution (Shapiro-Wilk test, $p<0.05$). In consequence, a non-parametric Kruskal-Wallis test was performed revealing significant main effects: $H_{(3)} =37.34$, $p=0.0001$). The post-hoc pairwise Mann-Whitney comparisons with Bonferroni correction showed that FA rate for the Natural condition was statistically smaller than for all the vocoded conditions: FA Natural Vs. FA Voc16: $p=0.007$; FA Natural Vs. FA Voc12 ($p=0.0001$) and FA Natural Vs. FA Voc8 ($p=0.0001$). On the other hand, the FA Voc16 was not significantly different from the FA Voc12 ($p=0.81$). FA rate for Voc16 was not significantly different from the FA Voc8 $p=0.08$. Finally, FA rate for Voc12 was not statistically different than FA Voc8 ($p=1.00$). Therefore, these results suggest that participants made fewer mistakes in the more intelligible speech condition (Natural speech).

The d prime was calculated as (zHR-zFA) and was used as a measure of sensitivity of performance. The d prime values were normally distributed (Shapiro-Wilk test, $p>0.05$). The one-way ANOVA showed significant main effects: $F(3, 78) = 70.92$, $p = 0.0001$. The post-hoc t-test pairwise comparison with Bonferroni correction showed that

d prime for Natural speech was significantly higher than all the vocoded conditions: d prime Natural Vs. d prime (p=0.0001); d prime Natural Vs. d prime Voc12 (p=0.0001) and d prime Natural Vs. d prime Voc8: (p=0.0001). In addition, d prime for Voc16 was also significantly higher than for Voc8 (p=0.0001), but not significantly different than the d prime obtained for Voc12 speech (p=0.11). Finally, d prime obtained for Voc12 was significantly higher than the values for the Voc8 (p=0.0001). This suggests that three significantly different levels of task sensitivity were observed: high intelligibility (Natural speech); moderate intelligibility (Voc16 and Voc12 speech) and low intelligibility (Voc8).

Reaction times to hit and false-alarm responses were analysed to explore whether task difficulty was reflected in these variables (similar to Kalaiah et al., 2017a). These were normally distributed (Shapiro-Wilk test, $p>0.05$), therefore a one-way ANOVA was also performed here. Reaction times to correctly classified non-words (hits) showed a significant main effect of conditions: $F(3, 78) = 4.05$, $p=0.01$. However, the post-hoc t-test pairwise analysis with Bonferroni correction only showed significantly smaller reaction times for Natural speech when compared to Voc12 ($p=0.001$). Reaction times to false alarm responses were also analysed but did not show any significant main effect: $F(3, 78) = 1.16$, $p=0.33$. These results suggested that reaction times were not a sensitive measure for discriminating between task difficulty within the current study.

Although differences in lexical decision making were not consistently observed across all the behavioural variables, d prime showed three significantly different levels in participants' performance: high intelligibility (Natural), moderate intelligibility (Voc16 and Voc12) and low intelligibility (Voc8). Therefore, d prime was used as a measurement of performance for Experiment 2 BN and Experiment 3 SSN.

For Experiment 2 (BN), three SNR levels were selected for comparison with Experiment 1 (Noise Vocoded): high intelligibility (Natural speech); moderate intelligibility (BN10); and low intelligibility (BN5), to manipulate the task difficulty as well

as auditory attention, see Section 2.7.3. The behavioural responses of twenty-nine normal hearing, young adults were obtained for the lexical decision (mean (+/- standard deviation, see Table 3.2). Because no task difficulty effects were observed in the reaction times to hits or false alarm responses in the Noise Vocoded experiment, these variables were not analysed in the BN or SSN experiments. Although hit rate and false alarm rate are shown in Table 3.2, statistical analysis was only performed for the d prime data.

Table 3.2. Experiment 2 (Babble Noise). Lexical decision task behavioural results (mean +/- standard deviation)

Stimuli	Hit Rate	False Alarm Rate	d prime
Natural	0.80±0.13	0.06±0.03	2.61±0.54
BN10	0.63±0.15	0.09±0.05	1.78±0.53
BN5	0.57±0.14	0.13±0.06	1.40±0.38

The d prime residuals were normally distributed (Shapiro-Wilk test, $p>0.05$) for all conditions. The one-way ANOVA showed significant main effects: $F(2, 56) = 110.76$, $p = 0.0001$. The pairwise t-test comparisons with Bonferroni corrections showed that the following conditions were significantly different: d prime Natural Vs. d prime BN10 ($p=0.0001$); d prime Natural Vs. d prime BN5 ($p=0.0001$) and d prime BN10 Vs. d prime BN5 ($p=0.0001$). Participants consistently showed better performance in the most intelligible conditions. In addition, d prime data revealed significant differences in response to the three levels of task difficulty selected *a priori* (the three levels: high (Natural), moderate (BN10) and low intelligibility (BN5)), demonstrating that we were able to manipulate performance and, presumably, auditory attention allocation in this way.

In Experiment 3 (SSN), the SSN conditions were piloted (see Section 2.7.4) and selected to ensure that significant differences existed in lexical decision performance similar to Experiment 1 (Noise Vocoded) and Experiment 2 (BN). Mean behavioural results (+/- standard deviation; see Table 3.3) were calculated for the same 29 normal hearing participants as in Experiment 2 (BN). Reaction times to hit and false alarm responses were not analysed here for SSN as they did not consistently reflect task difficulty in the Noise Vocoded experiment. Similar to the BN experiment, hit rate and false alarm rate are shown in Table 3.3, however, the statistical analysis was only performed for the d prime data.

Table 3.3. Experiment 3 (Speech-Shaped Noise). Lexical decision task behavioural results (mean +/- standard deviation)

Stimuli	Hit Rate	False Alarm Rate	d prime
Natural	0.80±0.13	0.05±0.03	2.61±0.54
SSN8	0.63±0.16	0.09±0.06	1.79±0.45
SSN3	0.55±0.15	0.12±0.06	1.36±0.37

The d prime residuals were normally distributed (Shapiro-Wilk test, $p>0.05$) for all conditions. The one-way ANOVA showed significant main effects: $F(2, 56) = 86.23$, $p = 0.0001$. The pairwise t-test comparisons with Bonferroni corrections showed that d prime values were statistically significant across all conditions: Natural Vs. SSN8 ($p=0.001$); Natural Vs. SSN3 ($p=0.001$) and SSN8 Vs. SSN3 ($p=0.001$). This result suggests that participants always performed better in more intelligible conditions. Therefore, the three levels of task difficulty were achieved: high (Natural), moderate (SSN8) and low intelligibility (SSN3).

Moreover, in order to corroborate that the three levels of performance were comparable between experiments, a planned pairwise comparison was performed. The d prime values when speech intelligibility was high were not significantly different between speech manipulations: Experiment 1 Natural Vs. Experiment (2 and 3 Natural): t (26) =-0.15, p=0.88. The d prime when speech intelligibility was moderate were not significantly different across experiments: Voc16 Vs BN10: t (26) =0.21, p=0.84; Voc16 Vs. SSN8: t (26) =0.35, p=0.73; Voc12 Vs. BN10; t (26) =-1.21, p=0.24; Voc12 Vs. SSN8: t (26) =-1.02, p=0.32; BN10 Vs. SSN8: t (28) =-0.09, p=0.93. Finally, performance at low levels of speech intelligibility were not statistically different across experiments: Voc8 Vs. BN5: t (28) =18.52, p=0.001; Voc8: t (28) =18.52, p=0.001 Vs. SSN3; BN5 Vs. SSN3: t (28) =18.52, p=0.001. Therefore, it can be assumed that performance at high, moderate and low speech intelligibility was comparable across experiments.

3.2 Attentional effects in speech-onset ERPs

The effects of auditory attention on speech-onset ERPs were assessed to confirm attentional differences between the Active and Passive listening conditions as well as the engagement of auditory cortex in the task. Early ERP components (P1, N1, P2) have shown attentional effects in previous reports (Woldorf et al., 1987, 1993; Karns & Knight, 2009; Saupe et al., 2009) and they are largely evoked in the auditory cortex (Da Silva 2009; Beres 2017). Moreover, the N400 component was also analysed because it has been previously associated with the processing of meaning (Kutas & Federmeier, 2011).

3.2.1 Attentional effects in speech-onset ERPs in Experiment 1 (Noise Vocoded)

The two-way Repeated Measured Analysis of Variance (RMANOVA) test performed allowed the investigation of attentional effects and potential stimulus-type effects. Attentional effects were assessed by the Conditions factor which had two levels that corresponded to: Active and Passive listening conditions. Stimulus-type effects were identified by the Stimuli factor which was the 2nd factor in the RMANOVA design

with 4 levels: stimuli 1(Natural), stimuli 2 (Voc16), stimuli 3 (Voc12) and stimuli 4 (Voc8) speech. A significant main effect of Conditions represents gross differences between Active and Passive listening while a significant main effect of Stimuli pointed to general stimulus-specific statistical differences. A significant interaction Conditions x Stimuli, on the other hand, represented potential task difficulty effects that were confirmed with a planned pairwise comparison within the Active listening conditions. It is worth mentioning that this design was maintained across all analyses performed in this study.

Event related activity, obtained from 64 electrodes, was measured in 27 participants; however, only 13 electrodes were statistically analysed due to their relevance in attentional and language brain activity related networks: (FZ (1), F3(2), F4(3), CZ (4), C3(5), C4(6), TP7(7), TP8(8), T7(9), T8(10), PZ (11), P3(12), P4(13). Grand averages of the activity recorded in the 13 selected electrodes are shown in Figure 3.1.

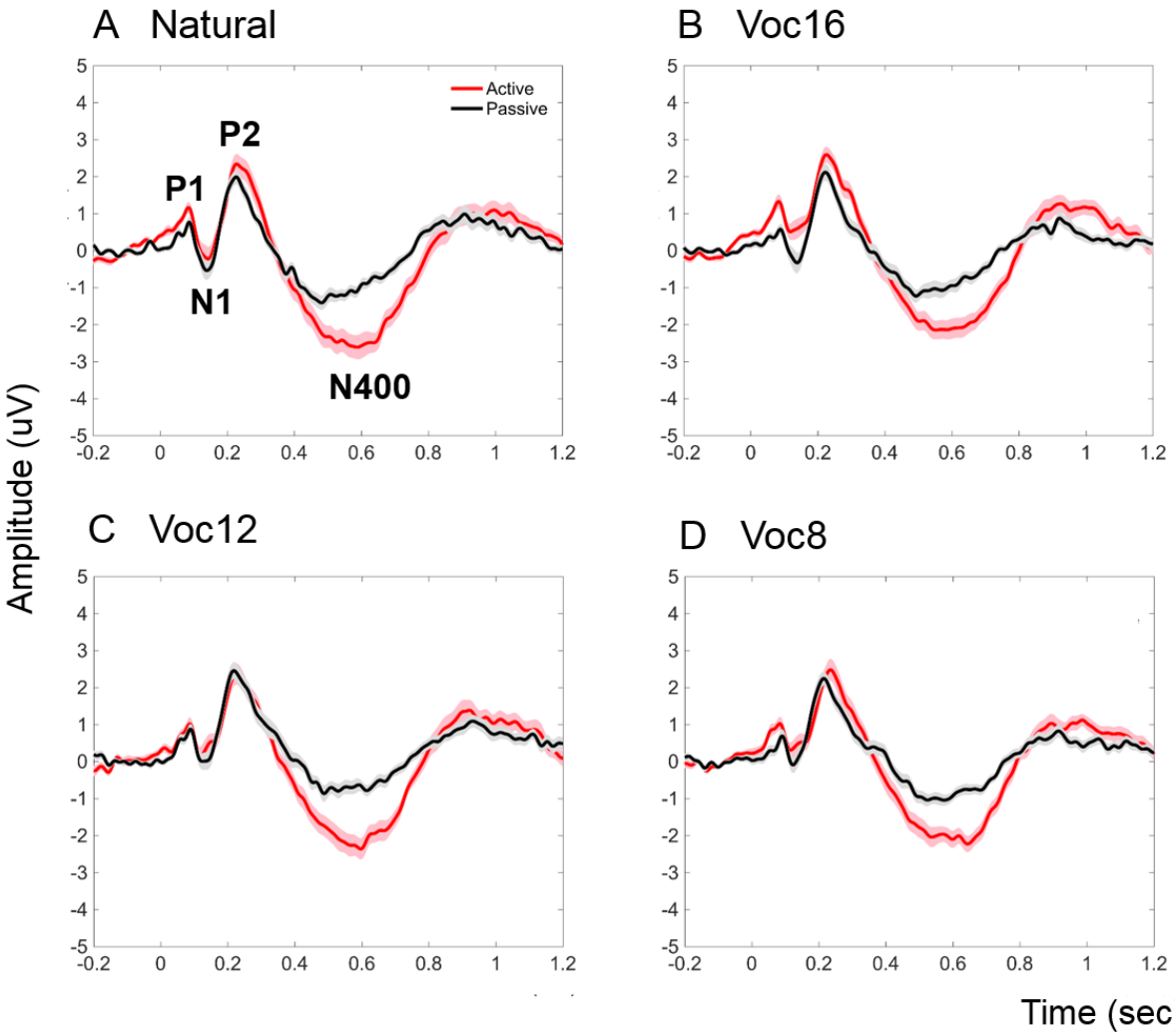

Figure 3.1. Grand average ERP components (P1, N1, P2 and N400) to the onset of words and non-words in the Noise Vocoded experiment for Active (red line) and Passive (black line) conditions. The shaded areas correspond to the SE of each waveform. Grand average ERP waveforms elicited during: **A.** Natural speech; **B.** Voc16 channels speech; **C.** Voc12 channels speech; and **D.** Voc8 channels speech.

The amplitudes of the ERP components: P1, N1, P2 and N400 were analysed, see Table 3.4 for mean and standard deviation values. The components were also tested for normality and they were normally distributed overall (Shapiro-Wilk test, $p>0.05$).

Table 3.4. Experiment 1 (Noise Vocoded). Event Related Potentials (uV) descriptive data (mean +/- standard deviation).

	Active Listening				Passive Listening			
	Natural	**Voc16**	**Voc12**	**Voc8**	**Natural**	**Voc16**	**Voc12**	**Voc8**
P1	0.60±0.90	0.91±0.70	0.66±0.68	0.64±0.86	0.13±0.96	0.26±0.75	0.46±0.73	0.23±0.75
N1	0.20±1.44	0.84±1.18	0.72±0.99	0.70±1.11	-0.16±1.12	0.07±0.98	0.39±1.18	0.47±0.88
P2	2.38±1.47	2.53±1.06	2.30±1.34	2.45±1.56	1.85±0.98	1.86±1.06	2.26±1.23	1.82±0.97
N400	-2.64±1.68	-2.21±1.32	-2.33±1.48	-2.06±1.13	-1.05±0.64	-1.04±0.98	-0.67±0.95	-1.03±0.75

An RMANOVA for each of the averaged ERP components for the 13 electrodes of interest was performed. The RMANOVA for the P1 component amplitude showed a significant main effect of Conditions: $F(1, 23) = 19.34$, $p = 0.0001$, where the P1 component obtained during Active listening had a significantly higher magnitude (0.71 ± 0.12 uV) than in Passive listening (0.27 ± 0.13 uV).

A significant, main effect of Conditions: $F(1, 24) = 15.32$, $p = 0.001$, and Stimuli: $F(3, 72) = 6.31$, $p = 0.001$, for the N1 component was observed. A *post hoc* pairwise comparison showed that N1 amplitudes were more positive in Active listening (0.62 ± 0.21 uV) than in the Passive listening conditions (0.19 ± 0.17 uV). The *post hoc* analysis of the Stimuli factor showed that the N1 component for Natural speech (0.02 ± 0.24 uV) was significantly more negative than the N1 component obtained for vocoded conditions: Voc16 (0.46 ± 0.21 uV), $p=0.012$; Voc12 (0.56 ± 0.19 uV), $p=0.006$ and Voc8 (0.59 ± 0.18 uV), $p=0.04$.

The repeated measures ANOVA performed for the P2 component indicated only a significant main effect of Conditions: $F(1, 25) = 11.22$, $p = 0.003$, where P2 amplitudes obtained during Active listening were significantly higher (2.42 ± 0.24 uV) than in the Passive listening conditions (1.95± 0.18 uV).

Finally, the RMANOVA for the N400 component revealed a significant main effect of both Conditions: $F(1, 24) = 32.84$, $p = 0.0001$). N400 amplitudes obtained during Active listening were statistically more negative (-2.31 ± 0.24 uV) than those in Passive listening (-0.95 ± 0.10 uV).

3.2.2 Attentional effects in speech-onset ERPs in Experiment 2 (Babble Noise)

Similar to Experiment 1, an RMANOVA analysis was performed. This RMANOVA maintained the Conditions factor with 2 levels that corresponded to Active and Passive listening conditions. The type of stimuli was the 2nd factor in the RMANOVA design with 3 levels: stimulus 1(Natural), stimulus 2 (BN10) and stimulus 3 (BN5). ERP components, obtained in 29 participants, were analysed to explore changes in the magnitude of the ERP components due to attention or stimulus-type delivered. The same 13 electrodes analysed in Experiment 1 were taken into consideration for grand average responses, see Figure 3.2

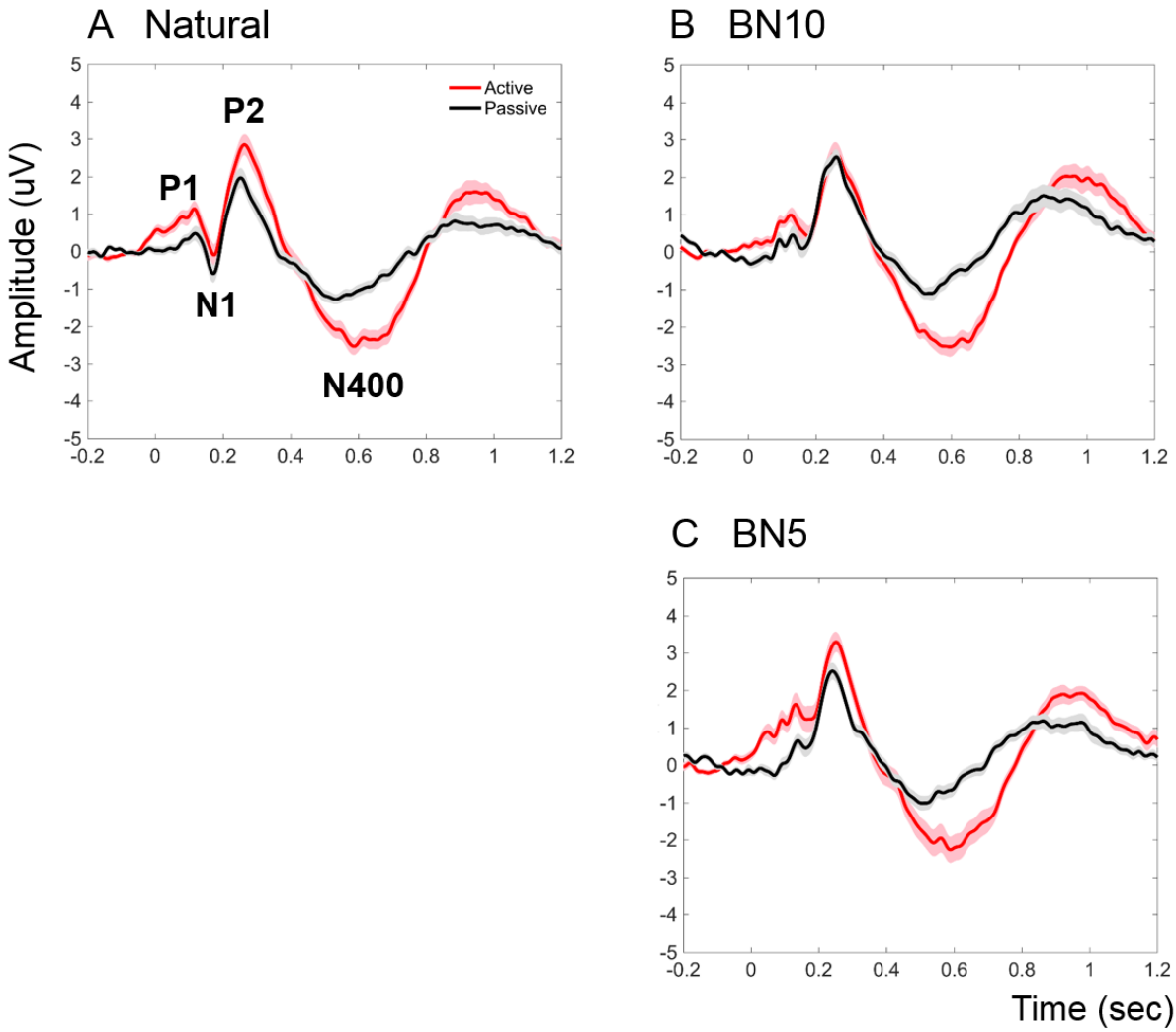

Figure 3.2. Grand Average of onset ERP components (P1, N1, P2 and N400) to words and non-words in the BN experiment for Active (red line) and Passive (black line) listening conditions. The shaded areas correspond to the SE of each waveform. Grand average ERP waveforms obtained during: **A.** Natural speech; **B.** Speech in BN10 and **C.** Speech in BN5.

The magnitudes of the ERP components (P1, N1, P2 and N400) were analysed, see Table 3.5 for mean and standard deviation values. The data residuals were normally distributed (Shapiro-Wilk test), $p>0.05$).

Table 3.5. Experiment 2 (Babble Noise). Event Related Potentials (uV) descriptive data (mean +/- standard deviation).

	Active listening			Passive listening		
	Natural	**BN10**	**BN5**	**Natural**	**BN10**	**BN5**
P1	0.91±0.86	0.87±0.76	1.02±0.63	0.48±0.98	0.24±0.69	0.27±0.88
N1	0.39±1.46	0.74±1.38	1.45±1.75	-0.12±1.26	0.23±1.28	0.57±1.11
P2	2.57±1.59	2.47±1.61	3.23±1.54	1.86±1.42	2.45±1.17	2.55±1.19
N400	-2.50±1.29	-2.51±1.30	-2.13±1.99	- 1.20±0.79	-0.91±0.99	-0.74±1.00

P1 magnitude showed a significant main effect of Conditions: (F (1, 25) = 22.53, p = 0.0001); the P1 component obtained during Active listening had a significantly higher magnitude (0.93 ± 0.12 uV) than in Passive listening (0.33 ± 0.14 uV).

The N1 component displayed a significant, main effect of Conditions: F (1, 28) = 12.06, p = 0.002. The *post hoc* pairwise comparison showed that the N1 amplitudes had a more positive magnitude (0.86. ± 0.24 uV) in the Active listening than in the Passive listening conditions (0.23 ± 0.19 uV). In addition, a main effect of Stimuli was also significant: F (2, 56) = 10.20, p = 0.0001. The *post hoc* analysis of the Stimuli factor showed that the N1 component for Natural speech was statistically more negative (0.13 ± 0.22 uV) than the N1 component for BN5 (1.01 ± 0.23 uV), p = 0.0001.

The RMANOVA performed for the P2 component showed a significant, main effect of Conditions: F (1, 28) = 7.31, p = 0.01; Stimuli: F (2, 56 = 8.01, p = 0.001 and a significant Interaction Conditions x Stimuli: F (2, 56) = 3.66, p = 0.032. During the Active listening conditions, the P2 component was significantly higher (2.76 ± 0.26 uV) than in the Passive listening conditions (2.29 ± 0.19 uV). For the Stimuli Factor, the pairwise comparison, showed only significant smaller P2 amplitudes during Natural speech (2.22 ± 0.25 uV) when compared to BN5 (2.89 ± 0.22 uV), p = 0.004. Finally, the *post hoc* analysis of the interaction showed that the P2 component was significantly higher in Active listening (2.57 ± 0.29 uV) than Passive listening (1.86 ± 0.26 uV) for Natural

speech (p=0.005). A similar relationship was also observed for BN5 stimuli: Active listening (3.23 ± 0.28 uV) Vs Passive listening: (2.55 ± 0.22 uV), p=0.013. Moreover, within the Active listening conditions, *post hoc* analysis confirmed that P2 magnitudes for BN5 (the less intelligible condition) (3.23 ± 0.28 uV) was significantly higher than the magnitudes obtained during Natural speech (2.57 ± 0.29 uV), p=0.008 and BN10 condition (2.47 ± 0.30 uV), p=0.010.

Additionally, the N400 component showed a significant main effect of Conditions: F (1, 28) = 43.75, p = 0.0001. N400 amplitudes during Active listening were statistically more negative (-2.38 ± 0.24 uV) than those during Passive listening (-0.95 ± 0.13 uV).

3.2.3 Attentional effects in speech-onset ERPs in Experiment 3 (SSN)

Similar to Experiment 2, the averaged EEG activity was obtained from the 13 electrodes of interest in the same 29 participants, see Figure 3.3. All ERP components' residuals followed a normal distribution across all conditions (Shapiro-Wilk, p>0.05). Means and standard deviations for all ERP components are shown in Table 3.6.

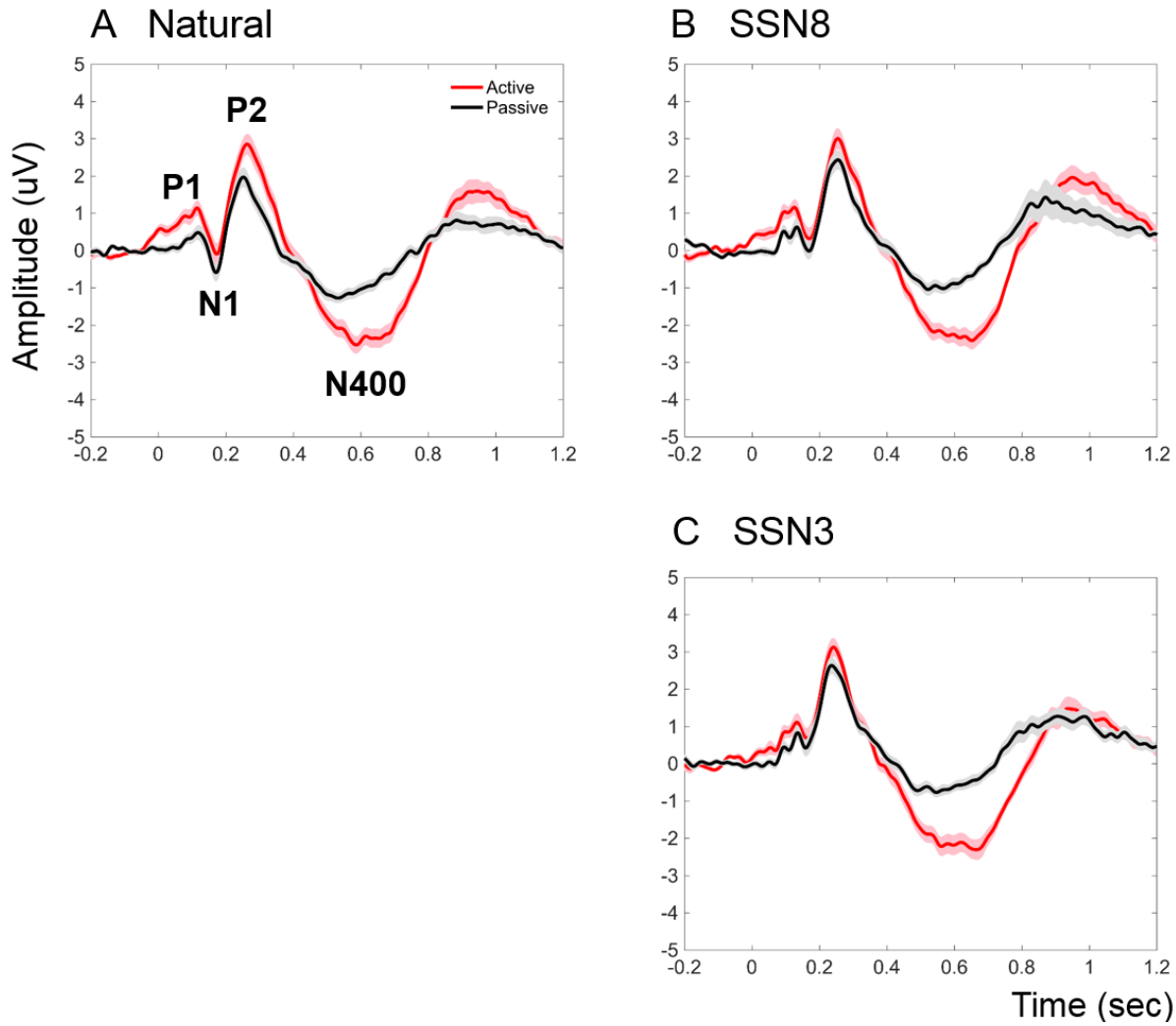

Figure 3.3. Grand Average of onset ERP components (P1, N1, P2 and N400) to words and non-words in the Speech-Shaped Noise experiment for Active (red line) and Passive (black line) listening conditions. The shaded areas correspond to the SE of each waveform. Grand average ERP waveforms obtained during: **A.** Natural speech; **B.** Speech in SSN8 and **C.** Speech in SSN3.

Table 3.6. Experiment 3 (SSN). Event Related Potentials (uV) descriptive data (mean +/- standard deviation).

	Active listening			Passive listening		
	Natural	**SSN8**	**SSN3**	**Natural**	**SSN8**	**SSN3**
P1	0.95±0.89	1.09±0.95	0.88±0.88	0.48±0.94	0.47±1.01	0.49±0.80

N1	0.24±1.25	0.66±1.37	0.78±1.17	-0.10±1.28	0.29±1.24	0.59±1.01
P2	2.57±1.59	2.83±1.58	3.16±1.39	1.86±1.42	2.40 ±1.43	2.67±1.14
N400	-2.50±1.29	-2.24±1.094	-2.28±1.47	- 1.20±0.79	-1.02±0.77	-0.73±0.71

P1 component amplitudes showed only a significant main effect of Conditions: F (1, 28) = 23.49, p = 0.0001, where P1 amplitudes during Active listening were significantly higher (0.97 ± 0.15 uV) than in Passive listening (0.48 ±0.14).

The RMANOVA performed for the N1 component showed a significant main effect of Conditions: F (1, 27) = 5.63, p = 0.025 and Stimuli: F (2, 54) = 7.43, p = 0.001. N1 amplitudes were more negative during Passive listening conditions (0.26 ± 0.19 uV) than during Active listening conditions (0.56 ± 0.21 uV). The *post hoc* pairwise comparison for the Stimuli factor showed that N1 amplitudes during Natural speech were significantly more negative (0.07 ± 0.22 uV) than those obtained during SSN8 (0.48 ± 0.22 uV), p = 0.018 and SSN3 (0.68 ± 0.19 uV), p = 0.005.

The RMANOVA showed a significant main effect of Conditions for P2 component magnitudes: F (1, 28) = 13.00, p = 0.001) and Stimuli: F (2, 56) = 8.32, p = 0.001). A *post hoc* analysis indicated that P2 amplitudes during Active listening were significantly higher (2.86 ± 0.26 uV) than those during Passive listening conditions (2.31± 0.20 uV). In addition, the pairwise comparison for Stimuli factor showed that P2 amplitudes obtained for Natural speech were significantly smaller (2.22 ± 0.25 uV) than those obtained for SSN3 (2.92 ± 0.21 uV), p = 0.0001.

The N400 component revealed a significant main effect of Conditions: F (1, 28) = 52.40, p = 0.0001. N400 amplitudes obtained during Active listening were statistically more negative (-2.34 ± 0.21 uV) than the magnitudes obtained during Passive listening (-0.98 ± 0.11 uV).

As shown for Experiments 1 (Noise vocoded speech), 2 (BN) and 3 (SSN), all ERP components were enhanced during Active listening compared to Passive listening conditions. This appears to confirm both auditory cortex engagement and the manipulation of auditory attention.

3.3 Click-evoked ABRs and OAEs in Experiment 1 (Noise Vocoded)

To determine the effects of attention and stimuli manipulations on the auditory brainstem and the periphery, ABRs and OAEs evoked by click stimulus were obtained. ABR components were analysed as well as CEOAE changes relative to the baseline.

3.3.1 ABRs in Experiment 1 (Noise Vocoded)

Click-evoked ABRs components: wave III (approximately corresponding to the SOC) and wave V (approximately corresponding to the IC) were evaluated (Møller & Jannetta, 1985). Amplitudes and latencies of waves III and V were analysed for each subject across blocks, conditions and experiments where possible (ABR components with positive SNR: see Table 1 in the Appendix section). The means and standard deviations for all the ABR variables are shown in Table 3.7. A Shapiro-Wilk test showed that ABR variables were normally distributed overall $p>0.05$.

Table 3.7. Experiment 1 (Noise Vocoded). Auditory brainstem responses descriptive data (mean +/- standard deviation).

	Active listening				**Passive listening**			
	Natural	**Voc16**	**Voc12**	**Voc8**	**Natural**	**Voc16**	**Voc12**	**Voc8**
Wave III Latency (ms)	4.03±0.24	4.03±0.24	4.05±0.23	4.02±0.26	4.01±0.24	4.03±0.25	4.02±0.25	4.01±0.24
Wave III Amplitude (uV)	0.03±0.01	0.03±0.01	0.03±0.02	0.04±0.01	0.04±0.02	0.03±0.01	0.03±0.02	0.04±0.02

Wave V Latency (ms)	6.09±0.24	6.04±0.22	6.07±0.26	6.05±0.25	6.04±0.23	6.03±0.24	6.07±0.22	6.02±0.21
Wave V Amplitude (uV)	0.11±0.03	0.11±0.03	0.11±0.03	0.12±0.03	0.11±0.02	0.11±0.02	0.11±0.02	0.11±0.02
Latency Difference	1.94±0.23	2.00±0.21	1.96±0.16	1.98±0.17	2.00±0.17	1.97±0.19	2.01±0.22	1.99±0.15
Amplitude ratio	2.89±0.82	3.12±0.76	3.42±1.76	2.86±1.20	2.55±1.02	3.06±1.36	2.80±1.06	2.68±1.13

An RMANOVA was performed for the latencies and amplitudes of wave III. This analysis showed that no significant main effects or interactions ($p > 0.05$) were observed for wave III latencies. Wave V was also analysed in terms of its latencies and amplitudes and the RMANOVA showed no significant main effects or interactions ($p > 0.05$). The latency differences between waves V and III were also calculated; the means and standard deviations for both the amplitude ratio and latency differences are also shown in Table 3.7. No main effects or significant interactions were found for the latency differences between wave V and wave III ($p > 0.05$). However, a significant main effect of Conditions was found for the wave V/III amplitude ratio: it was significantly higher (3.01 ± 0.23) during Active listening compared to Passive listening (2.77 ± 0.23): $F(1, 12) = 6.07$, $p < 0.05$. No significant effects of Stimuli or significant interaction Conditions x Stimuli were found ($p>0.05$).

To disentangle whether one or both of waves III and V were driving the significant main effects of Conditions for the ABR ratio, an RMANOVA for wave V and III amplitude was performed (only in the group of subjects who had valid data across conditions for the amplitude ratio). As shown in Figure 3.4, wave V amplitude remained constant between Active and Passive listening ($p<0.05$) while wave III was significantly smaller during Active Listening (0.04 ± 0.01) compared to Passive listening (0.05 ± 0.01): $F(1, 12) = 6.87$ $p = 0.02$.

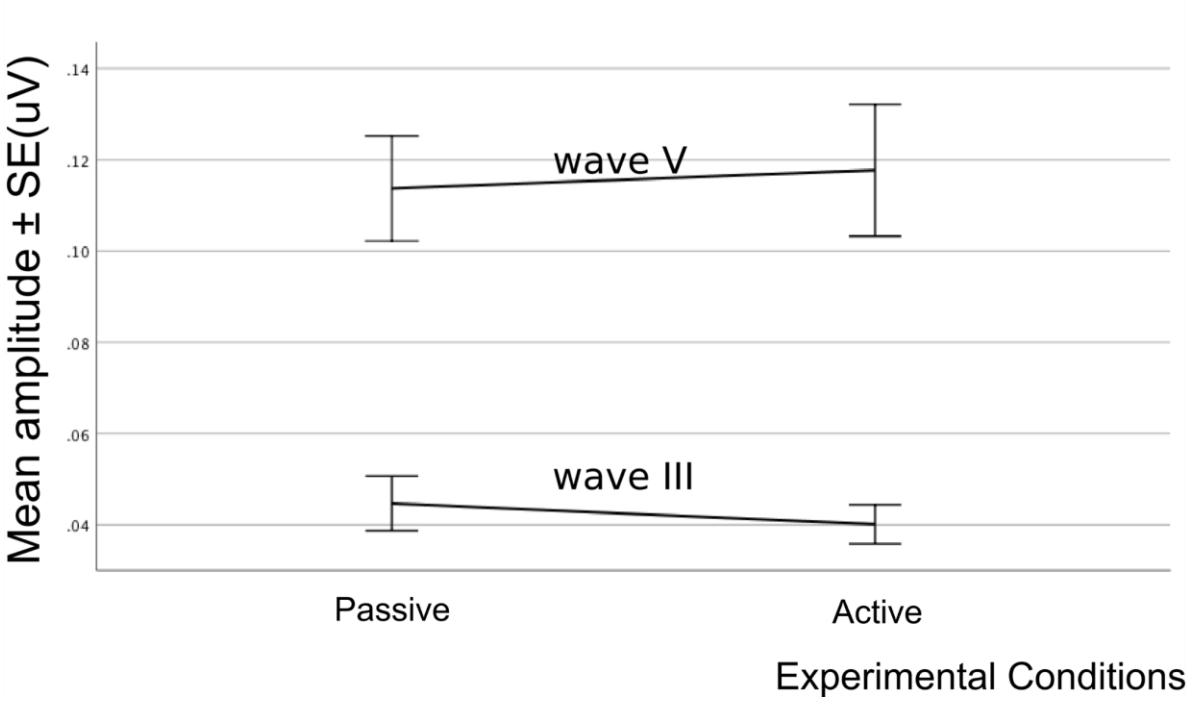

Figure 3.4. Wave V and wave III mean amplitudes during Active and Passive listening conditions. Note that the relative difference between the two waves decreased in the Passive listening compared to the Active listening.

In summary, most of the ABR variables were not sensitive to manipulations in either auditory attention, stimuli presented or task difficulty. Only, the magnitude of the ratio between the wave V and wave III showed significant variations due to auditory attention effects.

3.3.2 CEOAEs in Experiment 1 (Noise Vocoded)

The analysis of CEOAE magnitudes allowed us to explore the effects of attention, task difficulty as well as stimulus type on the cochlear gain via auditory efferent system activation. An RMANOVA analysis of all baselines (first minute of click stimulation) was performed in order to evaluate CEOAE magnitude effects.

CEOAEs from minute 1 (baseline) were tested for potential differences between experimental conditions that could potentially affect the absolute magnitude change.

Baselines from each condition were tested for normality and proved to be normally distributed overall (Shapiro-Wilk test, $p>0.05$). No significant main effects or interaction were observed for Conditions: $F(1, 23) = 1.13$, $p = 0.30$; Stimuli: $F(3, 69) = 1.57$, $p > 0.05$ or Interaction (Conditions x Stimuli): $F(3, 69) = 1.52$, $p > 0.05$. Because no significant differences were observed between CEOAE magnitudes across the different experimental conditions, an average baseline was calculated across all conditions. This allowed us to include data from participants with unreliable CEOAE baselines for some experimental conditions but not others in all subsequent analysis.

The absolute change in magnitude was obtained by subtracting CEOAE amplitudes measured in min 1 (Baseline) from those collected in the remaining 10 minutes of stimulus presentation. The CEOAE magnitude change relative to the baseline was the variable that informed about differential auditory efferent drive effects among experimental conditions.

The mean CEOAE magnitude change relative to the baseline is shown in Figure 3.5 while changes in individual CEOAEs are shown in Appendix Figure 1. The normality tests indicated that overall all the variables followed a normal distribution (Shapiro-Wilk test, $p>0.05$).

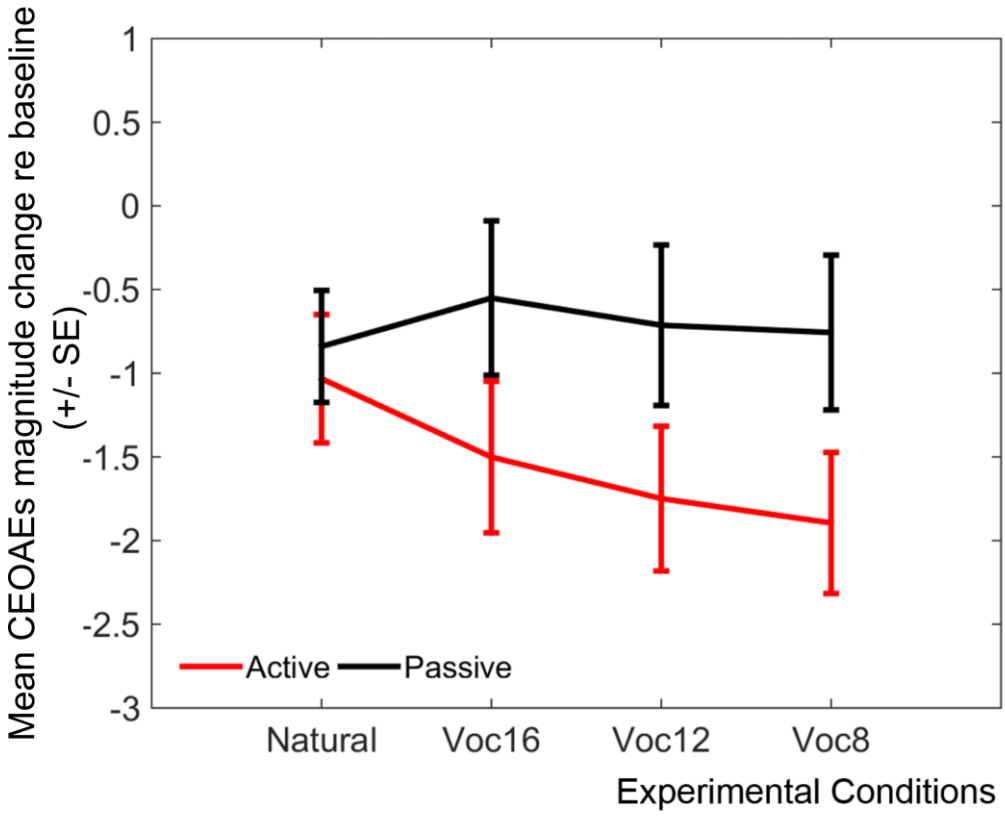

Figure 3.5. Mean CEOAE magnitude changes relative to the baseline in the Noise Vocoded experiment for Active (red line) and Passive (black line) listening conditions. The stimuli type presented are also shown: Natural, Voc16, Voc12 and Voc8 speech.

The RMANOVA showed a significant main effect of Conditions: CEOAEs during Active listening were more suppressed (-1.01±0.18 dB) than during Passive listening (-0.38 ± 0.18): $F(1, 22) = 8.49$, $p < 0.05$, however no significant main effect of Stimuli was observed ($p > 0.05$). In addition, a significant interaction of Conditions x Stimuli was also found: $F(3, 66) = 2.80$, $p < 0.05$. A *post hoc* pairwise comparison indicated there were no significant differences between CEOAEs for Natural speech during Active or Passive conditions: $t(24) = 0.62$, $p=0.54$. However, for all the vocoded conditions, CEOAEs were significantly more supressed in Active compared to Passive listening conditions: Voc16: $t(23) = -2.16$, $p=0.04$; Voc12: $t(24) = -2.19$, $p=0.038$ and Voc8: $t(25) = -3.51$, $p=0.002$. To test the effect of task difficulty, a planned pairwise comparison was performed within

the Active listening conditions which showed that CEOAE magnitudes were only significantly more suppressed for Voc8 compared to Natural speech: t (23) = 2.69, p=0.01. These results indicated that CEOAE magnitude changes relative to the baseline were sensitive to both auditory attention and task difficulty.

Finally, a t-test was performed to compare CEOAE magnitude changes relative to the baseline, thus reporting potential CEOAE magnitude changes relative to the baseline for any given experimental condition. Results suggest that during Active listening, CEOAEs for all the conditions were significantly smaller than the baseline (i.e. they showed significant levels of suppression): Natural: t (24) = -2.33, p=0.03; Voc16: t (24) = -3.43, p=0.002; Voc12: t (24) = -3.99, p=0.001 and Voc8: t (25) = -5.14, p=0.001. Only CEOAEs obtained during the Passive listening of Natural speech were significantly different from the baseline: t (25) = -2.29, p=0.03. This result shows that both auditory attention and stimulus-specific, auditory efferent effects can influence CEOAE magnitude when Natural speech is presented.

In this experiment, it was also shown that auditory attention modulated both the auditory brainstem and the cochlear gain. The ratio between waves V and III showed that, during Active listening, this component is enhanced compared to Passive listening conditions. On the other hand, the cochlear gain was suppressed during Active listening when compared to Passive listening; this suppression was stronger in the moderate (Voc16-Voc12) and low intelligibility (Voc8) conditions.

3.4 Click-Evoked ABRs and OAEs responses in Experiment 2 (Babble Noise)

3.4.1 ABRs in Experiment 2 (Babble Noise)

Both ABR and CEOAE measurements were analysed in a similar manner to Experiment 1. Table 3.8 shows means and standard deviations of all ABR variables analysed for the 29 subjects across all experimental conditions. Although not all

participants yielded valid ABRs recordings (see Table 2 in the Appendix section), a Shapiro-Wilk test showed that ABR variables were normally distributed overall, p>0.05.

Table 3.8. Experiment 2 (Babble Noise). Auditory brainstem response descriptive data (mean +/- standard deviation).

	Active listening			Passive listening		
	Natural	**BN10**	**BN5**	**Natural**	**BN10**	**BN5**
Wave III Latency (ms)	4.04±0.25	4.03±0.22	4.03±0.27	4.04±0.24	4.05±0.27	4.03±0.23
Wave III Amplitude (uV)	0.03±0.02	0.04±0.01	0.03±0.01	0.03±0.02	0.03±0.01	0.03±0.01
Wave V Latency (ms)	6.02±0.34	6.05±0.29	5.99±0.27	6.03±0.28	6.05±0.29	6.03±0.30
Wave V Amplitude (uV)	0.11±0.03	0.11±0.03	0.12±0.03	0.11±0.03	0.11±0.03	0.11±0.04
Latency Difference	1.92±0.25	1.98±0.20	1.93±0.17	1.96±0.17	1.96±0.20	1.97±0.21
Amplitude ratio	3.19±1.42	2.76±1.12	3.33±1.55	3.01±1.17	3.25±1.44	2.95±1.11

The RMANOVA performed for the variables, wave III latency and amplitude, showed no significant main effects or interactions (p>0.05). However, wave V latency showed a significant main effect of Stimuli: $F(2, 56) = 3.29$, $p < 0.05$; but no main effect of Conditions: $F(1, 28) = 1.29$, $p > 0.05$; or interaction Conditions x Stimuli: $F(2, 56) = 0.57$, $p > 0.05$. Although a significant main effect of Stimuli was observed, the pairwise comparison did not show any significant differences after Bonferroni correction. On the other hand, wave V amplitudes displayed a significant main effect of Conditions: $F(1, 27) = 5.67$, $p < 0.05$. Wave V magnitudes were significantly higher in Active listening (0.12±0.01 uV) than in Passive listening (0.11±0.01 uV). Neither the latency difference nor amplitude ratio showed significant main effects or interactions (p>0.05).

Although most of the ABRs variables did not show sensitivity to manipulations in auditory attention, task difficulty or stimulus presented; wave V magnitude showed significant variations due to auditory attentional effects.

3.4.2 CEOAEs in Experiment 2 (Babble Noise)

In order to determine whether attentional allocation, type of stimulus used or task difficulty had an effect on the way the auditory efferent system modulates the cochlear gain, CEOAE magnitudes were analysed. The residuals of CEOAE baselines were tested across all conditions for both normality and potential differences. A Shapiro-Wilk showed the data to be normally distributed ($p>0.05$). The RMANOVA analysis showed no significant, main effects: Conditions: $F(1, 26) = 0.22$, $p = 0.64$; Stimuli: $F(2, 54) = 0.54$, $p > 0.05$ or Interaction (Conditions x Stimuli): $F(2, 54) = 0.73$, $p > 0.05$. Because no effects of attention, type of stimulus used, or task difficulty were observed, an average baseline across all conditions was obtained which allowed for the inclusion of subjects with invalid baselines for a given condition.

Mean CEOAE magnitude changes relative to the baseline (averaged CEOAEs across 10 min) is shown in Figure 3.6 and its magnitude represents potential auditory efferent effects on the cochlear gain. Individual CEOAE changes relative to the baseline can be observed in Appendix Figure 2. Overall, these data's residuals followed a normal distribution (Shapiro-Wilk test, $p>0.05$). The RMANOVA showed no significant, main effects: Conditions: $F(1, 25) = 1.21$, $p = 0.30$; Stimuli: $F(2, 50) = 0.05$, $p > 0.05$ or interactions: Conditions x Stimuli: $F(2, 50) = 1.67$, $p > 0.05$. These results showed opposite findings to Experiment 1 (Noise Vocoded), here no change relative to the baseline in CEOAEs was observed due to manipulations of attention, type of stimulus used or task difficulty.

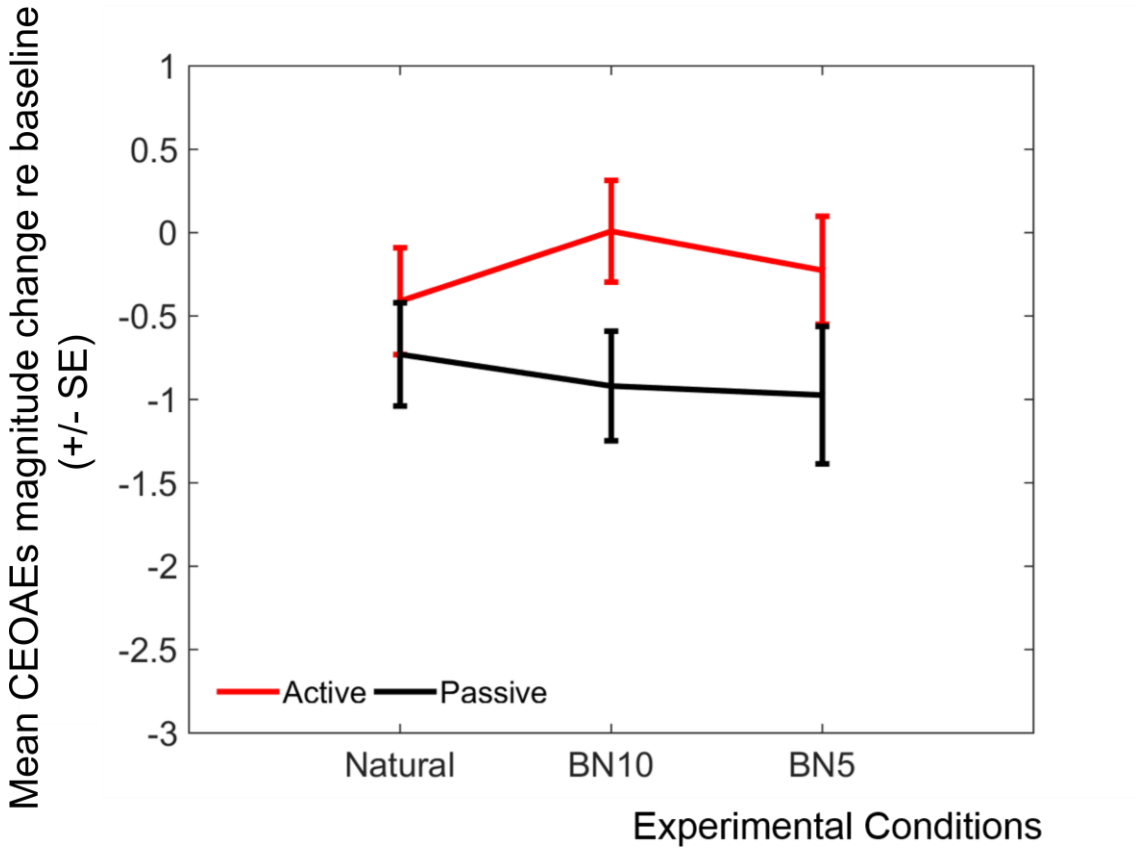

Figure 3.6. Mean CEOAE magnitude changes relative to the baseline in the Babble Noise experiment for Active (red line) and Passive (black line) conditions. The stimuli type presented are also shown: Natural, BN10, BN5.

Nevertheless, a t-test for CEOAE magnitude changes relative to the baseline was performed to explore whether CEOAE magnitudes were suppressed in any particular experimental condition. CEOAEs obtained during Passive listening conditions were significantly smaller than the baseline: Natural: $t(26) = -2.17$, $p=0.04$; BN10: $t(28) = -2.80$, $p=0.009$ and BN5: $t(28) = -2.36$, $p=0.02$. CEOAEs obtained during Active listening were not significantly different from the baseline ($p>0.05$). This result indicates that although CEOAE magnitude changes obtained during Active listening were not statistically different from the CEOAE amplitude changes in Passive listening, the individual experimental conditions in Passive listening were statistically smaller than the CEOAE baseline.

When participants actively listened to BN there was a significant enhancement of the ABR wave V amplitude in the Active listening compared to the Passive listening conditions, but this was not accompanied by a change in the cochlear gain. Nevertheless, CEOAEs were statistically different from the baseline during the Passive listening which suggests there was a significant suppression of CEOAE magnitudes.

3.5 Click-Evoked ABRs and OAEs responses in Experiment 3 (Speech-Shaped Noise)

3.5.1 ABRs in Experiment 3 (Speech-Shaped Noise)

Both the latencies and amplitudes of waves III and V were analysed for each subject across all experimental conditions. In addition, the amplitude ratio between waves V/III and the latency difference between waves V-III were also calculated. However, not all participants yielded valid ABRs data, see Table 3 in the Appendix section. Nevertheless, a Shapiro-Wilk test showed that residuals for all ABR variables were normally distributed overall, $p>0.05$. Table 3.9 shows the means and standard deviations for all ABR variables across experimental conditions.

Table 3.9. Experiment 3 (Speech-Shaped Noise). Auditory brainstem responses descriptive data (mean +/- standard deviation)

	Active listening			**Passive Listening**		
	Natural	**SSN8**	**SSN3**	**Natural**	**SSN8**	**SSN3**
Wave III Latency (ms)	4.08±0.28	4.07±0.27	4.06±0.29	4.05±0.24	4.06±0.25	4.09±0.28
Wave III Amplitude (uV)	0.03±0.02	0.04±0.02	0.03±0.02	0.03±0.02	0.03±0.02	0.02±0.01
Wave V Latency (ms)	6.02±0.34	6.03±0.31	6.07±0.29	6.03±0.28	6.05±0.32	6.02±0.35
Wave V Amplitude (uV)	0.11±0.03	0.11±0.03	0.11±0.03	0.11±0.03	0.11±0.03	0.10±0.03

Latency Difference	1.90±0.24	1.92±0.17	2.00±0.18	1.97±0.17	1.99±0.25	1.92±0.18
Amplitude ratio	3.03±1.25	2.59±1.06	3.75±1.38	3.05±1.38	2.80±1.01	3.33±1.42

The RMANOVA performed for wave III amplitude and latency showed no significant main effects or interactions ($p>0.05$). Similarly, the RMANOVA for the same variables of wave V indicated no significant main effects for wave V latency ($p>0.05$), but it did show a significant, main effect of conditions for its amplitude: the Active listening condition leading to higher wave V amplitudes (0.11±0.01 uV) than in the Passive listening condition (0.10±0.01 uV): $F(1, 25) = 8.91$, $p < 0.05$.

Latency differences between wave V-III did not show any significant main effects; however, the interaction Conditions x Stimuli was statistically significant: $F(2, 46) = 5.87$, $p < 0.05$. *Post hoc* analysis showed that only SSN3 produced a significantly smaller latency difference in Passive listening than in Active listening: $t(26) = 2.20$, $p=0.037$.

Although, the RMANOVA performed for the wave V-III amplitude ratio showed a significant main effect of Stimuli: $F(2, 36) = 2.97$, $p < 0.05$, the *post hoc* pairwise comparison did not display any significant effect ($p>0.05$).

3.5.2 CEOAEs in Experiment 3 (Speech-Shaped Noise)

CEOAE baselines were analysed here as in Experiments 1 and 2. The residuals were normally distributed (Shapiro-Wilk test, $p>0.05$). The RMANOVA analysis of all baselines showed that no main effects or interaction: Conditions: $F(1, 26) = 2.75$, $p>0.05$; Stimuli: $F(2, 52) = 0.83$, $p > 0.05$ or Interaction (Conditions x Stimuli): $F(2, 52) = 0.58$, $p > 0.05$. Therefore, an average baseline across all conditions was calculated.

Mean CEOAE magnitude changes relative to the baseline are shown in Figure 3.7. Individual CEOAE changes relative to the baseline can be observed in the Appendix

Figure 3. The data followed a normal distribution (Shapiro-Wilk test, $p>0.05$). An RMANOVA was performed showing a significant main effect of Conditions: $F(1, 24) = 4.44$, $p<0.05$. CEOAE suppression during the Passive listening condition (-0.83±0.26) was significantly higher than that for Active listening (-0.16±0.21). No significant main effect of Stimuli was observed ($F(2, 48) = 0.18$, $p>0.05$), but a significant Conditions x Stimuli interaction was reported: $F(2, 48) = 4.67$, $p<0.05$. The *post hoc* pairwise comparison indicated no significant differences between CEOAEs for Natural speech during Active or Passive conditions: $t(25) = -0.05$, $p=0.96$. However, for all the SSN conditions, CEOAEs were significantly more supressed in Passive listening than in Active listening: SSN8: $t(27) = 2.71$, $p=0.01$; SSN3: $t(28) = 2.67$, $p=0.012$. As this significant interaction showed a potential stimulus-specific effect during Passive listening, a planned pairwise comparison was performed within the Passive listening conditions. This showed no significant differences between CEOAE magnitudes in the Passive conditions, therefore CEOAEs did not significantly decrease with decreases with SNR.

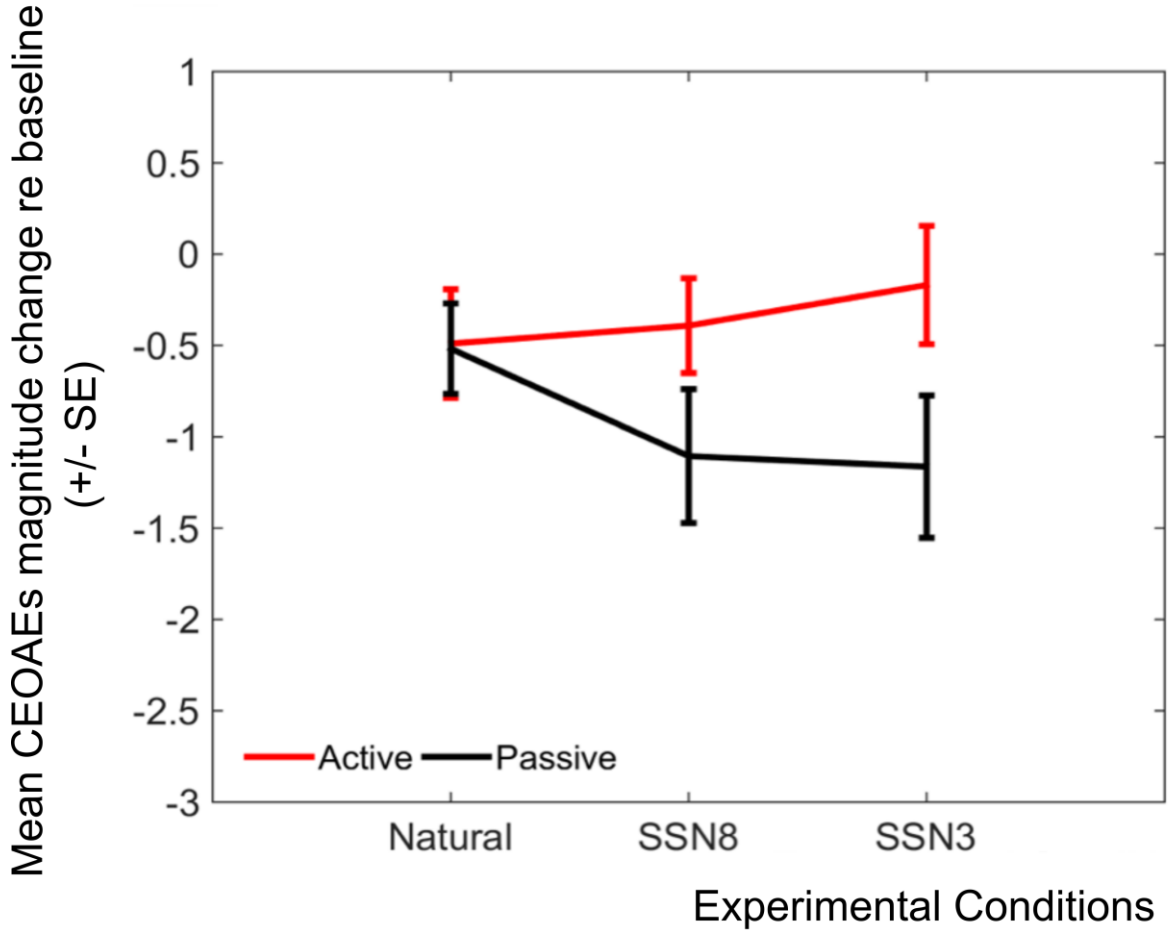

Figure 3.7. Mean CEOAE magnitude changes relative to the baseline in the SSN experiment, for Active (red line) and Passive (black line) conditions. The stimuli presented are also shown: Natural, SSN8, SSN3.

Furthermore, a t-test was performed in order to report CEOAE changes relative to the baseline. Note that similar to Experiment 2, only CEOAEs recorded during Passive conditions were significantly smaller than at baseline (therefore CEOAE magnitudes were significantly suppressed) Natural: $t(26) = -2.17$, $p=0.04$; SSN8: $t(28) = -3.37$, $p=0.002$ and SSN3: $t(28) = -3.50$, $p=0.002$. CEOAEs obtained during the Active listening conditions were not different from the baseline ($p>0.05$).

The results in this experiment were similar to the results in Experiment 2 (BN) but completely opposite to the findings in Experiment 1 (Noise Vocoded). There was a significant enhancement of the ABR wave V amplitude in the Active listening compared

to the Passive listening conditions. However, here the cochlear gain was suppressed in the Passive listening relative to the Active listening conditions. Like the results for the BN experiment, CEOAE magnitudes were only different from the baseline during Passive listening, which suggests there was only significant suppression of CEOAE magnitudes during this condition. Therefore, it is concluded that, as for the BN experiment, only during Active listening was the cochlear gain not significantly modulated relative to the baseline.

3.6 Exploring potential differences between populations tested

Although the populations tested across the three experiments were all normal hearing adults assessed under the same normal hearing criteria, age- and gender matched and recruited from undergraduate courses at Macquarie University (therefore presumably possessing similar literacy), potential differences due to intrinsic dissimilarities in the populations tested had to be excluded. Therefore, common variables across experiments were tested for potential differences such as ABR components and CEOAE magnitude changes relative to the baseline during the Active and Passive listening of Natural speech.

During Active or Passive listening of Natural speech, no statistical differences between the wave V-III amplitude ratio in Experiment 1 and Experiment (2-3) were observed: Active Natural: $t(12) = 0.90$, $p=0.39$; Passive Natural: $t(23) = 1.58$, $p=0.13$. Wave V amplitude during the Active or Passive listening of Natural speech was not statistically different between Experiment 1 and Experiment 2-3: (Active Natural: $t(23) = 0.09$, $p=0.93$; Passive Natural: $t(24) = -0.24$, $p=0.81$). Moreover, no differences were observed in CEOAE magnitude changes relative to the baseline during the Active listening of Natural speech: Experiment 1 Vs. Experiment (2-3): $t(23) = -0.21$, $p=0.83$; or during the Passive listening of Natural speech: Experiment 1 Vs. Experiment (2-3): $t(24) = -0.36$, $p=0.72$. These results ruled out the possibility that the attentional effects

observed across the three experiments were due to intrinsic differences between the populations tested.

Chapter 4 General discussion

It is widely accepted that sustained attention influences auditory cortical function (for examples see Näätänen & Picton, 1987; Karns & Knight, 2009; Saupe et al., 2009); however, its effects on lower levels of the auditory pathway remain unclear. Therefore, in this study, our main aim was to determine whether sustained attention to speech tokens (monosyllabic words) during an auditory lexical decision task could modulate cochlear activity via the auditory efferent system, while monitoring performance and auditory cortical activity. A secondary aim was to establish the sensitivity of such changes to increased levels of task difficulty (comparing across different levels of speech intelligibility) and to different manipulations of speech intelligibility (noise vocoded speech, speech in BN and SSN).

Our results showed that performance was successfully manipulated in the lexical decision task by implementing three manipulations of speech intelligibility. In each manipulation, three levels of task difficulty were achieved which were also comparable across experiments. These tasks engaged the auditory cortex and attention as demonstrated by the enhancement of cortical ERP components (P1, N1, P2 and N400 components) during the lexical decision task compared to Passive listening.

Our main finding was that there was a significant difference in CEOAE suppression between Active and Passive listening conditions for both Noise Vocoded and SSN experiments, which was not evident in the BN experiment. Despite these significant results for Noise Vocoded and SSN, the direction of the effects was opposite. Nevertheless, whenever CEOAE suppression was observed, it was accompanied by a reduction in ABR components, confirming that a cochlear gain reduction could also be observed in the activity of the auditory afferent pathway.

Interestingly, there was significant suppression of CEOAEs during Passive conditions for Natural speech across the three experimental manipulations of speech

intelligibility. While no significant difference in CEOAE suppression between the Active and Passive conditions was found for the BN experiment, the magnitude of the CEOAE suppression was significantly different from the baseline in the Passive condition, similar to the SSN experiment. This suggests that the MOC inhibitory drive was not active during the Active listening of speech in BN or SSN, whereas, during the Active listening of noise vocoded speech, it was.

For the Noise Vocoded Experiment, CEOAE suppression increased with task difficulty, therefore, the effect of modulated attention (varying speech intelligibility) on CEOAEs was only observed in this experiment. This suggests that noise vocoded speech may be a good experimental manipulation to better understand the effects of auditory attention on OAEs.

4.1 Subjective and objective measurements of auditory attention

Individuals are able to selectively direct their attentional resources to a signal of interest while reducing their allocation to the irrelevant stimulus in the scene. It has been described that attention can be top-down (a voluntary or task-dependent process) or bottom-up (saliency-based sound processes) (Fritz et al., 2007). Top-down attention is likely to operate via a gain control mechanism by which signals of interest are perceptually enhanced while the background scene is suppressed to maintain goal-directed behaviour (Ahveninen et al., 2006; Johnson & Zatorre 2006; Knudsen 2007). In each of the experiments performed in the current study, participants' attention was directed towards the speech signal while they performed the lexical decision task (Active listening conditions) and away from the speech signal while they watched a movie (Passive listening condition).

In addition, speech intelligibility was modulated using three manipulations: noise vocoded speech, speech in BN and SSN. Here it was assumed that decreases in speech intelligibility which caused increases in task difficulty, increased the allocation of auditory

attention towards the task (Kahneman, 1973). In line with this, behavioural results showed three levels of speech intelligibility: high, moderate and low, see Figure 4.1, which corresponded to high, moderate and low lexical decision performance. Behavioural performance on the auditory lexical decision task was significantly different across each intelligibility level, but performance was the same for each level of intelligibility across the three experimental manipulations (e.g., comparing Voc16 and Voc12, BN10 and SSN8).

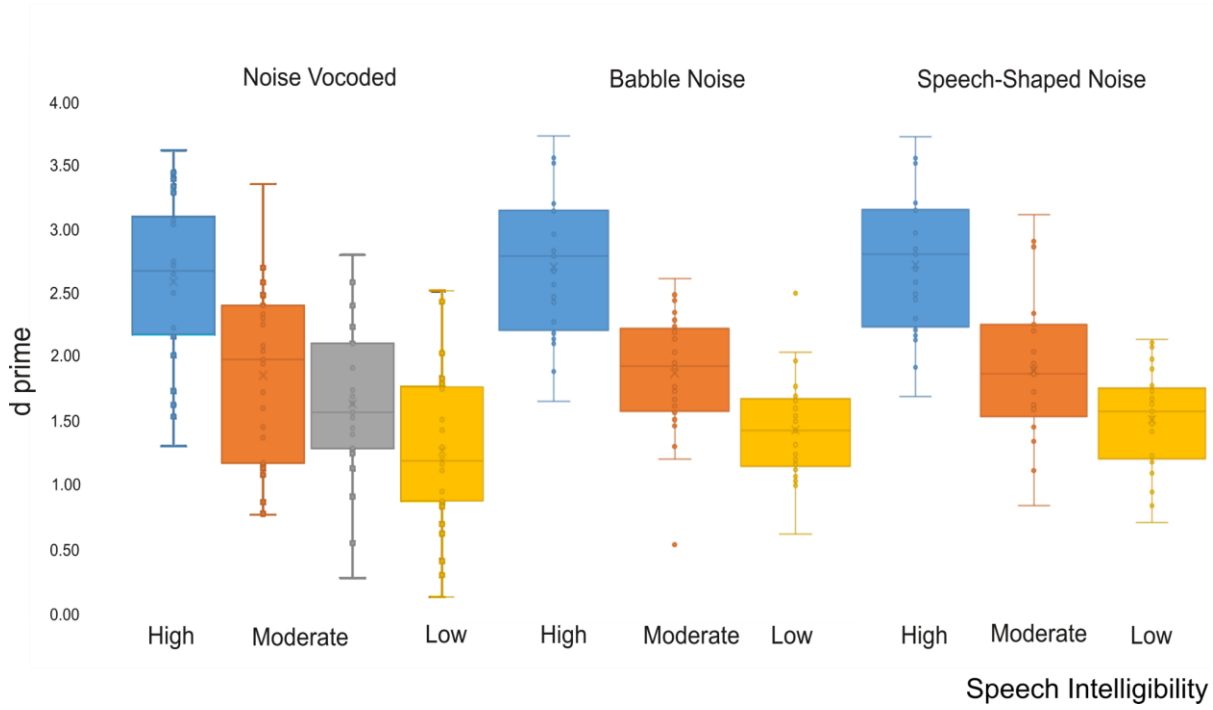

Figure 4.1. Performance (shown as d prime on the vertical axis) across the three experimental manipulations of speech intelligibility (Noise Vocoded, Babble Noise and Speech-Shaped Noise experiments) for three levels of intelligibility (high, moderate and low). Notice that for Noise Vocoded, two conditions showed performance corresponding to the moderate level of speech intelligibility manipulation.

While performance was significantly different across each of the three levels of intelligibility (high, moderate and low), this was generally not observed within the ERPs components, except for the P2 component in the BN experiment.

Since 1959, it has been known that neural activity in the auditory cortex can be modified by attention (Hubel et al., 1959). ERP components are mainly the result of synaptic activation, specifically postsynaptic potentials mediated by a number of neurotransmitter systems (Da Silva 2009; Beres 2017). It has been described that ERP components are larger in attended conditions compared to unattended conditions (Hillyard et al., 1973; Woldorf et al., 1987, 1993; Fujiwara et al., 1998). In line with this, across the three experiments in this study, statistically significant differences were consistently observed between the Active and Passive listening conditions (ERP components' magnitudes enhanced during the Active listening).

It has been also described that early ERP components can reflect stimulus characteristics while later components represent conscious cognitive processing and can be elicited in certain experimental conditions (Beres 2017). Consistent with this, significant differences were found mainly between Natural speech and the degraded speech stimuli across each of the three experiments. The only evidence of sensitivity to task difficulty was found for the P2 component in the BN experiment alone, which suggests that early positive components might be able to reveal decreases in speech intelligibility in certain auditory scenes such as multi-talker environments. Nevertheless, the ERP components analysed here were mainly sensitive to attentional changes (i.e. differences between Active and Passive conditions).

N400 is a later ERP component that peaks around 400 ms after the target word onset (Kutas & Hillyard 1980). It can be elicited by both visual and auditory words and non-words (Holcomb & Neville 1993; Bentin et al., 1993; Friedrich et al., 2006, Leinonen et al., 2009); there is general agreement that this component reflects the processing of

meaning (Deacon & Shelley-Tremblay 2000; Kutas & Federmeier 2011). Moreover, previous reports have found that the N400 magnitude is attenuated when participant's attention is diverted to non-target speech stimuli during either dichotic listening tasks or visual selective paradigms (Bentin et al., 1993; McCarthy & Nobre 1998; Okita & Jibu 1998). In line with these reports, the N400 component across the three experiments performed here showed larger magnitudes when participants were performing the lexical decision task (i.e. when participants were assigning a meaning to the speech material heard) compared to when participants focused their attention on the visual scene.

4.2 Influence of Active and Passive listening of speech on click-evoked ABRs

Attentional effects on shorter-latency potentials, such as ABRs are not consistently found in the literature. That is, whereas some ABRs studies have shown that the early auditory subcortical stages are not sensitive to attentional processes (Picton et al., 1971; Connolly et al., 1989; Hackley et al., 1990; Hirschhorn & Michie, 1990), others have found positive attentional effects of ABRs. For instance, Lukas (1980) tested 16 normal hearing adults and asked participants either to attend to 1 kHz tones or to mentally count flashed letters on a screen while recording ABRs to tone pips. The author found a decrease in magnitude and an increase in the latency of wave V (IC activity) when participants attended to the visual stimuli. Moreover, this same researcher tested 16 participants in 1981 using a similar paradigm that now involved either counting target tone pips or target letters while recording ABRs to both target and non-target tone pips. Here the author found a decrease in the magnitude and increase in the latency of wave I (auditory nerve component) for target tone pips when participants were performing the visual task. In line with Lukas's results, Brix (1984) recorded click-evoked ABRs in 100 normal hearing subjects while the participants either had to attend to the click stimulus (by attempting to count them), were distracted from it (participants were asked to ignore

the click stimulus) or read a newspaper (visual attention task). The author here found a shorter inter-peak latency between wave I and wave V when participants performed the auditory task than when they were asked to ignore the auditory stimuli.

In addition, Papanicolaou et al. (1986) compared ABR wave V amplitudes to click stimuli in 14 normal hearing adults who were instructed to relax and fixate their gaze. Participants had to either: repeat voiced or whispered speech passages; repeat speech in silent articulation; mentally recite speech or quietly attend to the binaurally presented click stimulus. The authors showed a significant reduction in wave V amplitude for the voiced and whispered conditions, suggesting that click stimuli (a stimulus irrelevant to the speech task) were suppressed at the IC level.

In line with the previously mentioned studies, wave V amplitudes in this study were slightly but significantly smaller during the Passive listening conditions in both BN and SSN experiments compared to the Active listening conditions. Moreover, during the SSN experiment, the latency difference between waves V-III was significantly greater during the Passive listening for the condition with lowest SNR (SSN3). These results suggest that the irrelevant and distracting auditory information was less efficiently processed at the brainstem and IC levels, most likely to facilitate focused attentional resources on the specific task (speech production or perception) or the visual scene.

In the Noise Vocoded experiment, during Active listening an increased waves V-III ratio was observed compared to Passive listening. This result was driven by a decrease in wave III amplitude (SOC activity) during Active listening compared to Passive listening while wave V (IC activity) remained constant across conditions.

Overall, these results suggest that the activity of nuclei such as those in the SOC and IC is able to reflect attentional states, consistent with described enhancements of auditory cortical activity to an attended auditory stimulus and suppression of auditory cortical activity to the ignored auditory stimulus (Woodruff et al., 1996; Zatorre et al.,

1999; Laurienti et al., 2002; Shomstein & Yantis 2004). Although authors such as Lukas (1980, 1981) and Brix (1994) have suggested that irrelevant auditory stimuli could be attenuated during visual tasks at a peripheral level due to the effects of the MOC on OHC electromotility (decreasing the input to the auditory brainstem), these authors could not confirm this hypothesis in their studies owing to the lack of MOC activity measurements. As ABRs and CEOAEs were simultaneously recorded in the current study, we were able to confirm whether the origin of the attentional effects along the auditory pathway were due to MOC effects on the cochlear amplifier or not; this will be further discussed in relation to the attentional effects observed in ABR variables across the three experiments performed here.

4.3 Effects of listening to speech on CEOAE magnitude.

In this thesis, two main effects were observed. The first one relates to the MOC activation during the Passive listening of Natural speech, Natural speech in BN and Natural speech in SSN (this did not occur in the noise vocoded conditions). The second aspect is the differential MOC activation during the Active listening conditions across the three experiments performed. During the Active listening to noise vocoded speech, CEOAE magnitude suppression was stronger with decreasing speech intelligibility; whereas Active listening in BN and SSN did not cause any changes to CEOAE magnitude compared to the baseline. In the next sections, these two main results will be discussed in detail.

4.3.1 Influence of Active listening to speech on CEOAE magnitude.

The effect of actively listening to speech on the cochlear gain has been studied for several years. However, the relationship between efferent control of the cochlear gain and speech processing remains unclear. Interest in this topic emerged from physiological reports of an antimasking effect (Wiederhold, 1970; Winslow & Sachs 1987; Guinan & Gifford 1988; Kawase et al., 1993) generated by the activation of the

MOC reflex when a transient signal is presented in noise. In this case, activation of the MOC reflex caused a suppression of OHC electromotility in response to the noise, while partially restoring the AN-responses to the transient signal (Kawase et al., 1993; Kawase & Liberman 1993; Guinan 2006) (see Chapter 1 for more details). This physiological evidence from animal models led to the hypothesis that the MOC reflex may play an important role in aiding humans to perceive speech in noise (Giraud et al., 1997; Liberman & Guinan 1998). The strict non-invasive nature of human experiments has made measuring MOC effects on OAE amplitudes the main technique for testing the antimasking hypothesis during speech processing. Although studies have explored the MOC effects on both TOAEs and DPOAEs during speech processing, its effects on DPOAEs are rather complicated to interpret because they do not always cause a decrease in DPOAE magnitude, most likely due to the mechanical origins of this OAE type (Siegel & Kim 1982; Wagner et al., 2007). Therefore, we will focus this discussion section on reports of TOAEs (CEOAEs) during speech processing which can be directly compared to the data obtained in this thesis.

Several studies have explored the relationship between the amount of CEOAE suppression and speech performance in noise (De Boer & Thornton 2008; Garinis et al., 2011; Smith & Cone, 2015; Kalaiah et al., 2017a). Garinis et al. (2011) found that when participants attended to words embedded in BBN and classified them as either food or animal items, CEOAEs were more suppressed than when backward words were presented in BBN or BBN was presented by itself.

Another similar study by Smith and Cone (2015) was successful in showing behavioural differences between the performance of an Easy Stroop task (listen to monosyllabic words and identify speaker gender) Vs. a Hard Stroop task (listen to monosyllabic words and identify the gender of the word spoken (gender-meaning specific: e.g., queen (female)) in young children. However, these increases in task

difficulty did not lead to increases in CEOAE suppression. More recently, Kalaiah et al., (2017a) asked their participants to categorize incoming words into two groups (e.g., animal or vehicle) in three white noise conditions (+3, -3 and -9 dB SNR). Although the authors were able to find task difficulty effects in the reaction times (quicker reaction times for the +3 dB SNR condition, higher speech intelligibility condition), these task difficulty effects did not translate into increases in the suppression of CEOAE magnitudes. Although similar to the experimental design used within this thesis, all the aforementioned studies used BBN as a speech masker. BBN is well known to maximally evoke MOC reflex (Lilaonitkul & Guinan 2009) but it is not the most efficient masker of speech because the level of the speech spectrum decreases with increasing frequency. Therefore, BBN (with its flat spectrum) predominantly masks the high frequency components of the speech. A question arising from these studies was whether the effects on the cochlear gain during Active speech processing would be seen when speech intelligibility was manipulated either in the absence of additional noise (i.e. by vocoding the speech) or in the presence of noises spectrally matched to the speech tokens such as BN and SSN.

In the Noise Vocoded experiment, a clear task difficulty effect was observed. This means that CEOAE suppression was significantly higher when the stimuli were less intelligible (see Figure 4.2 Noise Vocoded: Moderate and Low speech intelligibility levels). However, in Experiments 2 (BN) and 3 (SSN) no CEOAE magnitude changes relative to the baseline were observed during the lexical decision task, see Figure 4.2. This suggests that auditory attention can either suppress or does not affect CEOAE magnitude, therefore activating or inhibiting MOC-mediated effects in the cochlear gain depending upon which type of auditory stimuli is being processed.

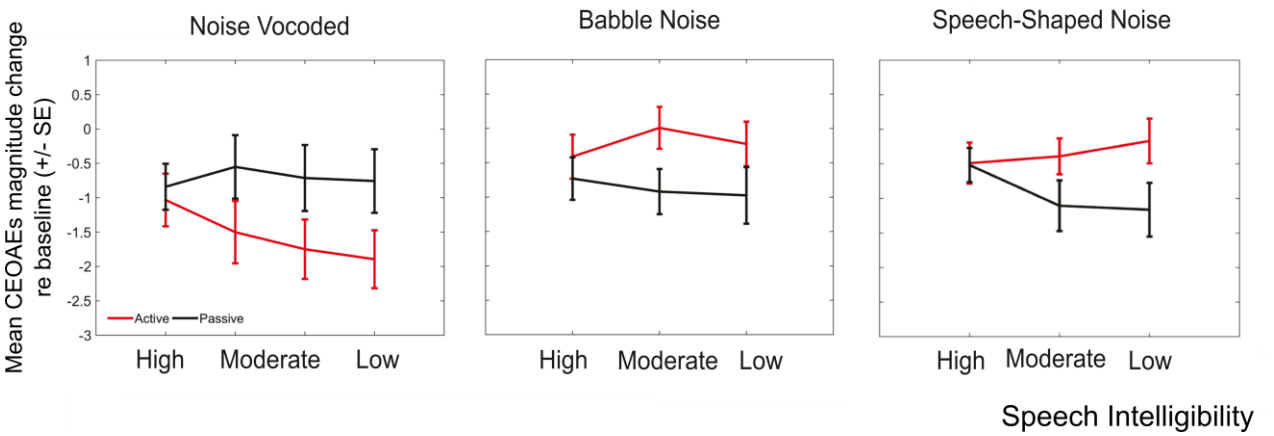

Figure 4.2. Click-Evoked OAEs suppression across the three experiments. The mean CEOAE magnitude changes relative to the baseline and the standard error are shown for Noise Vocoded, Babble Noise and Speech-Shaped Noise experiments.

Recently, Strauss and Francis (2017) provided a taxonomic description of attentional mechanisms during effortful listening based on a model of external and internal attention described by Chun et al. (2011). The internally-directed attentional mechanism is driven by a central, endogenous representation of the stimulus, while the externally-directed mechanism requires the selection of perceptual information by extraction of cues from the auditory scene. For example, a single degraded stream (such as noise vocoded speech) relies on endogenous, *a priori*, representations of the stimulus (conveyed by cognitive processes such as long-term and working memory) (Fougnie 2008; Chun et al., 2011; Strauss & Francis 2017), because no other cues are available in the scene that can contribute to speech perception. Here, we speculate that the cochlear gain reduction, observed in the contralateral ear to the noise vocoded speech during the lexical decision task, might have facilitated access to the internal representation of the speech tokens.

With regards to the externally-directed attentional mechanism as applied to effortful listening by Strauss and Francis (2017), one could argue that it might facilitate the

detection of speech-in-noise given the necessary assessment of all an incoming stimulus' physical properties to achieve speech perception (Chun et al., 2011; Strauss & Francis 2017). A caveat to be noted, however, is that in contrast to Garinis et al. (2011); Smith and Cone (2015) and Kalaiah et al. (2017), who observed changes in CEOAE magnitudes to negative SNRs during Active listening to speech; here positive SNRs between the target speech and the background noises were used in order to keep performance comparable across experiments. It is arguable that these SNRs may have been too high to obtain any benefits from the MOC antimasking effects.

In summary, listening to noise vocoded speech (one single degraded speech stream) or speech-in-noise (speech in BN or SSN), where two or more streams are competing (although comparable in terms of task difficulty level and auditory cortex generated ERPs components), may mobilize different neurophysiological mechanisms in the central nervous system (CNS) to achieve the same lexical classification. Evidence that single-speech streams are differentially processed compared to masked speech has been shown in the activity of the autonomic nervous system (Francis et al., 2016). The authors of this study used a very similar acoustic paradigm to the current one (synthesized speech, speech in BN or SSN) but measured skin conductance, pulse-amplitude and rate. Despite maintaining similar task difficult across conditions, the masked speech elicited stronger physiological reactions than the single-speech stream. In our study, although the Active listening of noise vocoded speech (single stream) generated strong CEOAE suppression while speech in BN or SSN did not, it suggests that different physiological mechanisms are recruited depending on the characteristics of the auditory scene.

4.3.2 Influence of Passive listening to speech on CEOAE magnitude.

During Passive listening to Natural speech, CEOAE magnitudes were always significantly smaller than the baseline (CEOAEs were suppressed). In addition, CEOAE

magnitudes were significantly smaller than the baseline in all Passive listening conditions during the BN and SSN experiments, see Figure 4.2. This means that whenever Natural speech was being presented (with or without noise) while participants watched the cartoon movie, the cochlear gain was being suppressed. However, for the Noise Vocoded speech, CEOAE magnitude remained similar to the baseline.

An MOC-mediated reduction of the cochlear gain during top-down visual attention has been described in animal models and humans (Puel et al., 1988; Maison et al., 2001; Delano et al., 2007; Wittekindt et al., 2014). These authors have speculated that this generalized reduction in the cochlear gain during visual attention most likely helps to suppress irrelevant auditory information during visual tasks. However, auditory stimuli may capture attention in a different manner depending on how easy they are to detect: i.e., they are saliency-based (bottom-up processes) (Kayser et al., 2005). However ultimately both types of processes interact in order to optimize performance (Egeth & Yantis, 1997; Sarter et al., 2001; Drover et al., 2018). For instance, it has been previously reported that the pitch carried in the voice or the fundamental frequency F0 is a very salient cue that plays an important role in the perceptual segregation of speech sources (Terhardt et al., 1982; Bregman, 1990; Darwin & Carlyon, 1995). Studies have shown that, when speech is noise vocoded, the pitch saliency carried by the envelope periodicity of the speech is diminished (Burns & Viemeister, 1976; 1981; Shackleton & Carlyon, 1994; Bernstein & Oxenham, 2003). In the context of our experiments, it is possible that noise vocoded speech was not salient (or distracting) enough to require MOC-mediated reduction of the cochlear gain to suppress irrelevant auditory information during visual attention. On the other hand, when Natural speech was presented either by itself or in the presence of BN or SSN, the voice pitch was preserved therefore the stimuli may have remained salient. In this case, the reduction of the cochlear gain via

MOC activation could have been beneficial in supressing irrelevant auditory information, therefore helping to focus attentional resources on the visual scene.

Although CEOAE magnitudes were statistically smaller than the baseline for both BN and SSN experiments, the strength of this suppression was larger for the latter. Both noises are considered wide-band stimuli and, as shown by Lilaonitkul and Guinan (2009), wide-band stimuli are more effective elicitors of the MOC reflex than narrower-band counterparts. However, although BN has been shown to elicit MOC reflex activations (Smith et al., 2001, Mishra & Lutman 2014; Kalaiah et al., 2017b), the magnitude of the CEOAE suppression reported was weaker than the suppression generated by stationary noises such as white noise (Kalaiah et al., 2017b). SSN, on the other hand, is not only wide-band but also has similar stationary characteristics (statistics over time) to white noise. In line with this, speech in SSN evoked significantly stronger CEOAE suppression during Passive listening compared to Active listening (the CEOAE suppression was stronger as the stimuli decreased in SNR (i.e., became noisier)). In summary, we argue here that during Passive listening conditions, the characteristics of the auditory stimulus such as its saliency and whether or not it strongly evokes the MOC reflex determines its effects on CEOAE magnitudes.

4.5 Contributions and limitations of this study to understand attentional effects in higher auditory centres

In this study, auditory lexical performance ensured participants' attentional engagement increased with task difficulty. Moreover, speech-onset ERPs served as a proxy to objectively confirm participants' attentional engagement in the task as well as auditory cortical involvement. As shown previously, performance was comparable between the three experiments and the auditory attentional effects on the ERPs were consistent across experiments. Moreover, whenever CEOAE magnitudes were statistically different from the relative baseline, a reduction in ABR components was also

observed. It has been previously pointed out that ABRs mainly register afferent activity from the auditory pathway (Moore; 1987; Meric & Collet 1992). For some time, it has been well known that the electrical stimulation of MOC fibres reduces the amplitude of AN compound action potential responses to clicks (Galambos, 1956; Wiederhold & Peake, 1966; Gifford & Guinan, 1987). Moreover, Guinan et al. (2005), reported that MOC efferent neurons inhibit single AN fibre's responses to click stimuli. Although, in this study, wave I components (thought to be generated by the AN) could not be extracted from the ABRs (because of stimulus level limitations), we speculate that a decrease in the AN-responses due to MOC reflex activation, could potentially also be reflected in the reduced magnitude of higher ABR components across the auditory brainstem. In line with this, a reduction in brainstem activity related to the click stimuli was consistently observed in this study whenever CEOAE magnitude was significantly suppressed. For instance, in the Noise Vocoded experiment where a significant MOC inhibitory effect on CEOAE magnitude was observed, an increase in the brainstem gain (ratio between wave V and III) was observed during the lexical decision task. As shown in Chapter 3, this higher wave V-III magnitude ratio during the Active listening was a result of decreases in wave III magnitude (SOC activity); this confirms that when CEOAE magnitudes to a click stimulus were suppressed, the associated reduction in the cochlear gain was evident along the afferent auditory pathway. Moreover, when a significant CEOAE suppression was observed during the Passive listening conditions in both BN and SSN experiments, a reduced wave V magnitude (IC activity) was also observed. These results again support the hypothesis that when the MOC inhibits the cochlear gain, these effects can be observed along the auditory efferent pathway. However, the auditory efferent system comprises multiple descending loops that can be modulated by central influences at any stage. In particular, IC receives direct efferent input from the AC (Bajo & Moore 2005) and even in animals with impaired efferent

synapse in the OHC (absence of alpha-9/10 nicotinic receptors), wave V can still be modulated when AC is microstimulated (Aedo et al., 2016). Because our findings were restricted to effects in waves III (SOC) and V (IC), we cannot rule out that the changes observed in wave V were also generated by efferent AC input to the IC. Nevertheless, the simultaneous recording of the cochlear gain modulation and brainstem activity in this study does confirm that the MOC efferent affects the auditory afferent pathway.

One aspect that cannot be underestimated is individual differences: not only the different effects of MOC reflex on CEOAEs among participants (Mertes & Goodman 2016), but also the variety of potential AC effects on the MOC reflex among individuals (Dragicevic et al., 2015). Dragicevic et al. (2015) showed that electrically stimulating the AC of chinchillas produced enhancement, suppression or no change of the cochlear gain depending on the strength of the individual MOC reflex. Although the approach described within this thesis attempted to find auditory efferent strategies during speech perception at a population level while both monitoring performance and confirming AC involvement in the task (by monitoring speech-onset ERPs), it is possible that different participants utilize different perceptual and cognitive strategies/resources in order to achieve the task (Motowildo et al., 1997; Barret et al., 2004; Eysenck et al., 2007). Although highly relevant, this was not analysed here because it is beyond the scope of this thesis and would need to be explored further in future studies.

Moreover, there are well-known gender-related differences in CEOAEs, ABRs and ERPs components. For example, several studies have described how CEOAE magnitude in humans are stronger in females than in males (Bilger et al., 1990; Talmadge et al., 1993; McFadden & Shubel, 2003; McFadden et al., 2009). In addition, females have shown earlier click-ABRs latencies (Jerger & Hall, 1980) and ERPs components (such as N400: higher order semantic processing) (Wirth et al., 2007). Although exploring gender-differences was not a main objective of the present study,

the sample across and within experiments was gender-balanced therefore we do not foresee any potential bias in our results. However, this could be investigated in future studies.

We also acknowledge that it is still uncertain which stimulus features are enhanced by the MOC activity in order to aid speech perception. It is possible that a more frequency-specific OAE measurement such as SFOAEs may help to better understand the MOC effects on the cochlear gain. Moreover, the parallel analysis of both the amplitude and the phase of OAEs when the MOC is being activated, might also help to disentangle not only individual differences but also attentional-specific effects in the cochlear gain.

In addition, our results are confined to CEOAE changes in the contralateral ear that received the speech stimuli. Therefore, we cannot exclude that the ear in which the CEOAEs were not recorded was not also affected by MOC activation. Although, it has been anatomically described that MOC neurons receive descending cortical (Mulders & Robertson 2000b) and IC input (Faye-Lund 1986; Thompson & Thompson 1993; Vetter et al., 1993), the potential ipsi-contra or bilateral activation profiles that these connections produce in the MOC activity has not yet been explored.

Nevertheless, we can conclude that for the sample of subjects tested here in the Noise Vocoded experiment, contralateral MOC-mediated mechanisms were most likely recruited via auditory attention and the corticofugal pathway. However, for the BN and SSN experiments, auditory attention inhibited the contralateral cochlear gain suppression via the MOC pathway. Therefore, we speculate that auditory attention acted adaptively as a goal-oriented mechanism that dynamically recruited the auditory efferent control of the cochlear gain depending upon the characteristics of auditory scene being analysed.

Chapter 5. Implications for future studies and Conclusions.

5.1 Implications for future studies

The classical view that only neurons above the thalamus can be actively modulated by attending to auditory streams has been challenged since the anatomical discovery of auditory efferent connections from the brainstem to the cochlea (Rasmussen, 1946). In 1956, Hernandez-Peon (1956) showed how CN activity to click stimuli were suppressed when cats were presented with olfactory or visual stimuli which diverted the animal's attention away from the auditory stimulus. Although the methodology of this study was challenged, the finding was replicated with adequate controls by Oatman (1971, 1976) and Glenn and Oatman (1977), who showed inhibitory effects on the AN in ABRs when animals were engaged in visual discrimination tasks. The hypothesis that attentional states could modulate activity in the brainstem and periphery gained more attention after direct inputs from the auditory cortex were described by Mulders and Roberson (2000b). Yet most human studies of attention maintain a cortically-centred view of attentional effects. A more integrative view of how the afferent and efferent pathways interact during attentional states may facilitate a better understanding of attentional disorders in populations with learning difficulties (Veuillet et al., 1999, 2007; Hoen et al., 2008), auditory processing disorders (Muchnik et al.,2004; Sanches & Carvallo, 2006) and autism (Collet et al., 1993; Danesh & Kaf 2012; Wilson et al., 2017), where auditory efferent control of the cochlear gain has been reported to behave abnormally.

Nobre and Katsner (2014) have described attention as a process that allows the selection and integration of information across time and space by dynamically tracking events in a scene. This usually involves complex interactions of selection, attenuations and enhancements across and within multiple sensory modalities (Karns & Knight 2009). Therefore, the initial hypotheses that an MOC-driven, cochlear gain suppression is larger during auditory attention (in order to improve the SNR) or during visual attention

(in order to decrease sensitivity to irrelevant stimulus) seems somewhat simplistic and does not reflect the way that the brain processes multi-sensory naturalistic scenes. Studies are needed where the cochlear gain is monitored during tasks where both auditory and visual information must be integrated to build up the scene. This will contribute to the better understanding of the likely role that the auditory efferent system performs in integrating information across sensory domains and in object formation.

CEOAE studies have typically found that an MOC-driven effect on the cochlear gain will be activated/beneficial for in signals in the presence of white noise but, as shown in this thesis, less so in BN or SSN. The lack of an effect in the presence of these different noise types may be due to the high SNR used in this study. Potentially, the anti-masking effect is less beneficial at such SNRs, given that maximal CEOAE suppression has been found for negative SNRs between speech and white noise (-3 dB) (Garinis et al., 2011; Smith & Cone 2015; Kalaiah et al. (2017a).

Several studies have implemented the MOC efferent reflex in auditory speech-processing models (Ghitza, 1988; Messing et al., 2009; Brown et al., 2010; Clark et al., 2012) and have shown better prediction of human performance than standard afferent models. Recently, Lopez-Poveda et al. (2016) proposed a sound coding strategy with the MOC reflex implemented for binaural cochlear implants. Three binaural cochlear implant participants and two single-sided CI users were asked to report sentences presented in spatially-separated SSN. The authors found that speech reception thresholds were better with the MOC reflex algorithm than with the standard cochlear implant processing. However, most CEOAE studies, including this thesis, have not been performed in binaural conditions as in the Lopez-Poveda et al. (2016) study. Therefore, it is possible that under certain spatial configurations, the MOC drive actually improves speech perception by affecting AN output both contralaterally (as explored here in this thesis), ipsilaterally or even bilaterally to the speech stimuli. These experimental

differences must be reconciled in future studies with larger study populations where not only attentional states are controlled but also their individual differences (e.g. MOC reflex strength, motivation, listening effort).

Finally, the relationship between speech processing and modulation of the cochlear gain must be better understood. Contralateral CEOAE suppression (MOC - mediated, cochlear gain reduction) has previously been correlated with different variables of speech in noise performances such as: phoneme perception in noise, (Giraud et al., 1997; De Boer & Thornton 2008); monosyllabic words perception in noise (Kumar & Vanaja, 2004) and sentence recognition in noise (Bidelman & Bhagat, 2015). However, other studies of contralateral OAE suppression have found no correlation (Mukari & Mamat, 2008; Wagner et al., 2008; Stuart & Butler, 2012) or even a negative correlation (participants with stronger efferent reflex had poorer speech-in-noise scores (De Boer et al., 2012)). In consequence, no strong hypothesis linking peripheral gain control and the processing of continuous speech has been made. Low-frequency brain activity has been linked to fluctuations in the speech envelope and seems to be crucial for speech intelligibility (Rosen, 1992; Greenberg & Ainsworth, 2006, Riecke et al., 2017). While continuous natural speech has been successfully probed in the cortex with neurophysiological tools such as EEG and Magneto Encephalography (MEG), ABRs to continuous speech have been recorded successfully only recently (Reichenbach et al., 2016; Maddox & Lee 2017). Moreover, Forte et al. (2017) showed that brainstem responses to continuous speech are consistently modulated by attention in a two-competing speakers paradigm. The challenge is to develop a non-invasive method to explore the modulation of the cochlear gain to continuous speech. The non-invasive nature of OAE recordings makes them a perfect candidate; however, as of now, OAE recordings to continuous speech have yet to be reported. Comprehending the way in which continuous speech is processed along the entire auditory pathway might help

“close the gap” between speech and auditory fields, leading towards a more integrative view of how attentional mechanisms select and integrate information across time and space in order to achieve speech comprehension.

5.2 Conclusions

This research contributes to the growing body of literature that seeks to better understand the effects of attention on neural activity in the CNS. Performance in a lexical decision task was combined with cortical electrophysiological measurements to confirm the manipulation of auditory attention and the engagement of the auditory cortex. Indeed, it was found that relative activity in cortex, midbrain and brainstem was a suitable metric for distinguishing between passive and active listening conditions in the population of subjects tested. However auditory attention would appear to modulate the cochlear gain in a stimulus-specific manner (i.e. when its activation presumably benefits target perception). The cochlear gain was reduced when task difficulty increased with noise vocoded speech stimuli; however, no changes in the cochlear gain were observed when speech was presented in noise. Passively listening to speech during visual attention seems to effectively reduce the cochlear gain. The simultaneous monitoring of the auditory periphery, brainstem and cortex while participants were engaged in ecologically valid tasks helped identify how broadly the attentional effects on the cochlear gain spread along the auditory afferent pathway. Finally, this study advances our understanding of how attention acts adaptively as a goal-oriented mechanism to achieve perception by dynamically recruiting auditory efferent control of the cochlear gain.

References

Acunzo, D. J., MacKenzie, G., & van Rossum, M. C. (2012). Systematic biases in early ERP and ERF components as a result of high-pass filtering. Journal of neuroscience methods, 209(1), 212-218.

Aedo, C., Terreros, G., León, A., & Delano, P. H. (2016). The corticofugal effects of auditory cortex microstimulation on auditory nerve and superior olivary complex responses are mediated via alpha-9 nicotinic receptor subunit. *PloS one*, *11*(5), e0155991.

Alho K, Vorobyev VA, Medvedev SV, Pakhomov SV, Starchenko MG, Tervaniemi M, Naatanen R: Selective attention to human voice enhances brain activity bilaterally in the superior temporal sulcus. Brain Res 2006, 1075:142-150.

Alho K, Vorobyev VA: Brain activity during selective listening to natural speech (review). Front Biosci 2007, 12:3167-3176.

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., & Witzel, T. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. Proceedings of the National Academy of Sciences, 103(39), 14608-14613.

American National Standards Institute (1995). S3.7-1995 (R2008), Method for Coupler Calibration of Earphones (American National Standards Institute, New York).

American National Standards Institute. (1996). American National Standard Specification for Audiometers. (ANSI S3.6-1996). New York: ANSI.

Arnesen, A. R. (1984). Fibre population of the vestibulocochlear anastomosis in humans. Acta oto-laryngologica, 98(5-6), 501-518.

Ashmore, J., Avan, P., Brownell, W. E., Dallos, P., Dierkes, K., Fettiplace, R., & Canlon, B. (2010). The remarkable cochlear amplifier. Hearing research, 266(1), 1-17.

Avan, P., & Bonfils, P. (1992). Analysis of possible interactions of an attentional task with cochlear micromechanics. Hearing research, 57(2), 269-275.

Azeredo, W. J., Kliment, M. L., Morley, B. J., Relkin, E., Slepecky, N. B., Sterns, A., & Woods, C. I. (1999). Olivocochlear neurons in the chinchilla: a retrograde fluorescent labelling study. Hearing research, 134(1), 57-70.

Backus, B. C. (2007). Bias due to noise in otoacoustic emission measurements. *The Journal of the Acoustical Society of America*, *121*(3), 1588-1603.

Bajo, V. M., & Moore, D. R. (2005). Descending projections from the auditory cortex to the inferior colliculus in the gerbil, Meriones unguiculatus. Journal of Comparative Neurology, 486(2), 101-116.

Bentin, S., Kutas, M., & Hillyard, S. A. (1993). Electrophysiological evidence for task effects on semantic priming in auditory word processing. Psychophysiology, 30(2), 161-169.

Bentin, S., Kutas, M., & Hillyard, S. A. (1995). Semantic processing and memory for attended and unattended words in dichotic listening: Behavioural and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(1), 54.

Beres, A. M. (2017). Time is of the Essence: A Review of Electroencephalography (EEG) and Event-Related Brain Potentials (ERPs) in Language Research. Applied Psychophysiology and Biofeedback, 42(4), 247-255.

Berlin, C. I., Hood, L. J., Wen, H., Szabo, P., Cecola, R. P., Rigby, P., & Jackson, D. F. (1993). Contralateral suppression of non-linear click-evoked otoacoustic emissions. Hearing research, 71(1), 1-11.

Bernstein, J. G., and Oxenham, A. J. 2003. “Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number?” J. Acoust. Soc. Am. 113, 3323–3334.

Barrett, L. F., Tugade, M. M., & Engle, R. W. (2004). Individual differences in working memory capacity and dual-process theories of the mind. *Psychological bulletin*, *130*(4), 553.

Bidelman, G. M., & Khaja, A. S. (2014). Spectrotemporal resolution tradeoff in auditory processing as revealed by human auditory brainstem responses and psychophysical indices. Neuroscience letters, 572, 53-57.

Bidelman, G. M., & Bhagat, S. P. (2015). Right-ear advantage drives the link between olivocochlear efferent ‘antimasking’and speech-in-noise listening benefits. Neuroreport, 26(8), 483-487.

Bidet-Caulet A, Bertrand O: Dynamics of a temporo-frontoparietal network during sustained spatial or spectral auditory processing. J Cogn Neurosci 2005, 17:1691-1703.

Bilger R., Matthies M. L., Hammel D. R., and Demorest M. E. (1990). “Genetic implications of gender differences in the prevalence of spontaneous otoacoustic emissions,” J. Speech Lang. Hear. Res. 33, 418–432.

Boothalingam, S., & Purcell, D. W. (2015). Influence of the stimulus presentation rate on medial olivocochlear system assays. *The Journal of the Acoustical Society of America*, *137*(2), 724-732.

Box, G. E. P., & Tiao, G. C. (1973). Bayesian inference in statistical analysis. AddisonWesley Ed. *Reading, Massachusetts*.

Bredberg, G., Ades, H. W., & Engström, H. (1972). Scanning electron microscopy of the normal and pathologically altered organ of Corti.*Acta Oto-Laryngologica*, *73* (S301), 3-48.

Bregman, A. S. (1994). *Auditory scene analysis: The perceptual organization of sound*. MIT press.

Brix, R. (1984). The influence of attention on the auditory brain stem evoked responses preliminary report. *Acta oto-laryngologica*, *98*(1-2), 89-92.

Brown MC (1994) Antidromic responses of single units from the spiral ganglion. J Neurophysiol, **71**:1835–1847.

Brown, G. J., Ferry, R. T., & Meddis, R. (2010). A computer model of auditory efferent suppression: implications for the recognition of speech in noise. The Journal of the Acoustical Society of America, 127(2), 943-954.

Brown, M. C. (1987). Morphology of labeled efferent fibers in the guinea pig cochlea. *Journal of Comparative Neurology*, *260*(4), 605-618.

Brown, M. C. (1989). Morphology and response properties of single olivocochlear fibres in the guinea pig. Hearing research, 40(1), 93-109.

Brown, M. C. (2014). Single-unit labeling of medial olivocochlear neurons: the cochlear frequency map for efferent axons. Journal of neurophysiology, 111(11), 2177-2186.

Brown, M. C., & Levine, J. L. (2008). Dendrites of medial olivocochlear neurons in mouse. Neuroscience, 154(1), 147-159.

Brown, M. C., & Vetter, D. E. (2009). Olivocochlear neuron central anatomy is normal in α9 knockout mice. Journal of the Association for Research in Otolaryngology, 10(1), 64-75.

Brown, M. C., De Venecia, R. K., & Guinan Jr, J. J. (2003). Responses of medial olivocochlear neurons. Experimental brain research, 153(4), 491-498.

Brownell, W. E., Bader, C. R., Bertrand, D., & De Ribaupierre, Y. (1985). Evoked mechanical responses of isolated cochlear outer hair cells. Science, 227(4683), 194-196.

Brownell, W. E. (1990). Outer hair cell electromotility and otoacoustic emissions. Ear and hearing, 11(2), 82.

Brugge, J. F., & Howard, M. A. (2002). Hearing. *Encyclopedia of the Human Brain*, *2*, 429-448.

Burns, E. M., and Viemeister, N. F. (1976). "Nonspectral pitch," J. Acoust. Soc. Am. 60, 863–869.

Burns, E. M., and Viemeister, N. F. (1981). "Played again SAM: Further observations on the pitch of amplitude-modulated noise," J. Acoust. Soc. Am. 70, 1655–1660.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., & Kiessling, J. (1994). An international comparison of long-term average speech spectra. The Journal of the Acoustical Society of America, 96(4), 2108-2120.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. Journal of Experimental Psychology: Human Perception and Performance, 27(1), 115.

Chadwick, R. S., Dimitriadis, E. K., & Iwasa, K. H. (1996). Active control of waves in a cochlear model with subpartitions. Proceedings of the National Academy of Sciences, 93(6), 2564-2569.

Chan, D.K. and Hudspeth, A.J. (2005). Ca (2+) current-driven nonlinear amplification by the mammalian cochlea in vitro. Nat Neurosci 8 (2), 149-155.

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. The Journal of the Acoustical Society of America, 25(5), 975-979.

Chun, M., Golomb, J. D., & Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. Annual Review of Psychology, 62, 73–101.

Clark, N. R., Brown, G. J., Jürgens, T., & Meddis, R. (2012). A frequency-selective feedback model of auditory efferent suppression and its implications for the recognition of speech in noise. The Journal of the Acoustical Society of America, 132(3), 1535-1541.

Clayson, P. E., Baldwin, S. A., & Larson, M. J. (2013). How does noise affect amplitude and latency measurement of event-related potentials (ERPs)? A methodological critique and simulation study. Psychophysiology, 50(2), 174-186.

Collet, L., Kemp, D. T., Veuillet, E., Duclaux, R., Moulin, A., & Morgon, A. (1990). Effect of contralateral auditory stimuli on active cochlear micro-mechanical properties in human subjects. Hearing research, 43(2), 251-261.

Collet, L., Roge, B., Descouens, D., Moron, P., Duverdy, F., & Urgell, H. (1993). Objective auditory dysfunction in infantile autism. *The Lancet*, *342*(8876), 923-924.

Connolly, J. F., Aubry, K., McGillivary, N., & Scott, D. W. (1989). Human brainstem auditory evoked potentials fail to provide evidence of efferent modulation of auditory input during attentional tasks. Psychophysiology, 26(3), 292-303.

Coomes Peterson, D., & Schofield, B. R. (2007). Projections from auditory cortex contact ascending pathways that originate in the superior olive and inferior colliculus. Hearing research, 232(1), 67-77.

Cooper, N. P., & Guinan, J. J. (2003). Separate mechanical processes underlie fast and slow effects of medial olivocochlear efferent activity. The Journal of physiology, 548(1), 307-312.

Cooper, N. P., & Guinan, J. J. (2006). Efferent-mediated control of basilar membrane motion. *The Journal of physiology*, *576*(1), 49-54.

Da Silva, F. L. (2009). EEG: Origin and measurement. In EEG-fMRI (pp. 19-38). Springer Berlin Heidelberg.

Danesh, A. A., & Kaf, W. A. (2012). DPOAEs and contralateral acoustic stimulation and their link to sound hypersensitivity in children with autism. International journal of audiology, 51(4), 345-352.

Darrow, K. N., Benson, T. E., & Brown, M. C. (2012). Planar multipolar cells in the cochlear nucleus project to medial olivocochlear neurons in mouse. *Journal of Comparative Neurology*, *520*(7), 1365-1375.

Darwin, C. J., and Carlyon, R. P. (1995). “Auditory Grouping,” in Hearing, edited by B. C. J. Moore Academic, San Diego.

Davis, M. H., & Johnsrude, I. S. (2007). Hearing speech sounds: top-down influences on the interface between audition and speech perception. *Hearing research*, *229*(1-2), 132-147.

De Boer, J., & Thornton, A. R. D. (2007). Effect of subject task on contralateral suppression of click evoked otoacoustic emissions. Hearing research, 233(1), 117-123.

De Boer, J., & Thornton, A. R. D. (2008). Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. *Journal of Neuroscience*, *28*(19), 4929-4937.

De Boer, J., Thornton, A. R. D., & Krumbholz, K. (2012). What is the role of the medial olivocochlear system in speech-in-noise processing? Journal of neurophysiology, 107(5), 1301-1312.

De Venecia, R. K., Liberman, M. C., Guinan, J. J., & Brown, M. C. (2005). Medial olivocochlear reflex interneurons are located in the posteroventral cochlear nucleus: a kainic acid lesion study in guinea pigs. Journal of Comparative Neurology, 487(4), 345-360.

De Venecia, R. K., Liberman, M. C., Guinan, J. J., & Brown, M. C. (2005). Medial olivocochlear reflex interneurons are located in the posteroventral cochlear nucleus: a kainic acid lesion study in guinea pigs. *Journal of Comparative Neurology*, *487*(4), 345-360.

Deacon, D., & Shelley-Tremblay, J. (2000). How automatically is meaning accessed: a review of the effects of attention on semantic processing. *Frontiers in Bioscience*, *5*(Part E), 82-94.

Delano, P. H., Elgueda, D., Hamame, C. M., & Robles, L. (2007). Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas. The Journal of Neuroscience, 27(15), 4146-4153.

Don, M., & Elberling, C. (1994). Evaluating residual background noise in human auditory brain-stem responses. The Journal of the Acoustical Society of America, 96(5), 2746-2757.

Doucet, J. R., & Ryugo, D. K. (2003). Axonal pathways to the lateral superior olive labeled with biotinylated dextran amine injections in the dorsal cochlear nucleus of rats. Journal of Comparative Neurology, 461(4), 452-465.

Doucet, J. R., Rose, L., & Ryugo, D. K. (2002). The cellular origin of corticofugal projections to the superior olivary complex in the rat. Brain research, 925(1), 28-41.

Dragicevic, C. D., Aedo, C., León, A., Bowen, M., Jara, N., Terreros, G., Delano, P. H. (2015). The Olivocochlear Reflex Strength and Cochlear Sensitivity are Independently Modulated by Auditory Cortex Microstimulation. Journal of the Association for Research in Otolaryngology, 16(2), 223-240.

Drover, W., Wood, M. S., & Corbett, A. C. (2018). Toward a cognitive view of signalling theory: individual attention and signal set interpretation. *Journal of Management Studies*, *55*(2), 209-231.

Edwards, D. H. (2006). Neural Sciences Neuroscience. Edited by Dale Purves, George J Augustine, David Fitzpatrick, William C Hall, Anthony-Samuel LaMantia, James O McNamara, and S Mark Williams. Sunderland (Massachusetts): Sinauer Associates.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual review of psychology*, *48*(1), 269-297.

Elberling, C., & Don, M. (1984). Quality estimation of averaged auditory brainstem responses. *Scandinavian audiology*, *13*(3), 187-197.

Elgoyhen, A. B., Johnson, D. S., Boulter, J., Vetter, D. E., & Heinemann, S. (1994). α9: an acetylcholine receptor with novel pharmacological properties expressed in rat cochlear hair cells. Cell, 79(4), 705-715.

Elgoyhen, A. B., Vetter, D. E., Katz, E., Rothlin, C. V., Heinemann, S. F., & Boulter, J. (2001). α10: a determinant of nicotinic cholinergic receptor function in mammalian vestibular and cochlear mechanosensory hair cells. Proceedings of the National Academy of Sciences, 98(6), 3501-3506.

Elgueda, D., Delano, P. H., & Robles, L. (2011). Effects of electrical stimulation of olivocochlear fibers in cochlear potentials in the chinchilla. *Journal of the Association for Research in Otolaryngology*, *12*(3), 317-327.

Emadi, G., C.-P. Richter, et al. (2004). "Stiffness of the gerbil basilar membrane: radial and longitudinal variations." Journal of neurophysiology 91(1): 474-488.

Engström, H., Ades, H. W., & Andersson, A. (1966). Structural pattern of the organ of Corti: a systematic mapping of sensory cells and neural elements. Stockholm: Almqvist & Wiksell.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, *7*(2), 336.

Eysenck, M. (2012). *Attention and arousal: Cognition and performance*. Springer Science & Business Media.

Faye-Lund, H. (1986). Projection from the inferior colliculus to the superior olivary complex in the albino rat. Anatomy and embryology, 175(1), 35-52.

Feeney, M. P., Keefe, D. H., & Marryott, L. P. (2003). Contralateral acoustic reflex thresholds for tonal activators using wideband energy reflectance and admittance. *Journal of Speech, Language, and Hearing Research*, *46*(1), 128-136.

Ferber-Mart, C., Duclaux, R., Collet, L., & Guyonnard, F. (1995). Influence of auditory stimulation and visual attention on otoacoustic emissions. Physiology & behavior, 57(6), 1075-1079.

Forte, A. E., Etard, O., & Reichenbach, T. (2017). The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention. *eLife*, *6*.

Fougnie, D. (2008). The relationship between attention and working memory. *New research on short-term memory*, *1*, 45.

Fox, S. I. (2006). Human Physiology 9th Editon.

Francis, A. L., MacPherson, M. K., Chandrasekaran, B., & Alvar, A. M. (2016). Autonomic nervous system responses during perception of masked speech may reflect constructs other than subjective listening effort. *Frontiers in psychology*, *7*, 263.

Friedrich, M., & Friederici, A. D. (2006). Early N400 development and later language acquisition. *Psychophysiology*, *43*(1), 1-12.

Fritz, J. B., Elhilali, M., David, S. V., & Shamma, S. A. (2007). Auditory attention—focusing the searchlight on sound. Current opinion in neurobiology, 17(4), 437-455.

Froehlich, P., Collet, L., Chanal, J. M., & Morgon, A. (1990). Variability of the influence of a visual task on the active micromechanical properties of the cochlea. Brain research, 508(2), 286-288.

Froud, K. E., Wong, A. C. Y., Cederholm, J. M., Klugmann, M., Sandow, S. L., Julien, J. P., ... & Housley, G. D. (2015). Type II spiral ganglion afferent neurons drive medial olivocochlear reflex suppression of the cochlear amplifier. *Nature communications*, *6*, 7115.

Fuchs, P. A. (1996). Synaptic transmission at vertebrate hair cells. Current opinion in neurobiology, 6(4), 514-519.

Fujino, K., & Oertel, D. (2001). Cholinergic modulation of stellate cells in the mammalian ventral cochlear nucleus. Journal of Neuroscience, 21(18), 7372-7383.

Fujiwara, N., Nagamine, T., Imai, M., Tanaka, T., & Shibasaki, H. (1998). Role of the primary auditory cortex in auditory selective attention studied by whole-head neuromagnetometer. *Cognitive brain research*, *7*(2), 99-109.

Galambos, R. (1956). Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea. *Journal of neurophysiology*, *19*(5), 424-437.

Garinis, A. C., Glattke, T., & Cone, B. K. (2011). The MOC reflex during active listening to speech. Journal of Speech, Language, and Hearing Research, 54(5), 1464-1476.

Ghitza, O. (1988, April). Auditory neural feedback as a basis for speech processing. In Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on (pp. 91-94). IEEE.

Giard, M. H., Fort, A., Mouchetant-Rostaing, Y., & Pernier, J. (2000). Neurophysiological mechanisms of auditory selective attention in humans. Frontiers in bioscience, 5(1), 84-94.

Gifford ML, Guinan JJ Jr (1983) Effects of crossed-olivocochlear bundle stimulation on cat auditory nerve fibre responses to tones. J Acoust Sot Am 74: 115-123.

Gifford, M. L., & Guinan Jr, J. J. (1987). Effects of electrical stimulation of medial olivocochlear neurons on ipsilateral and contralateral cochlear responses. *Hearing research*, *29*(2), 179-194.

Giraud, A. L., Garnier, S., Micheyl, C., Lina, G., Chays, A., & Chéry-Croze, S. (1997). Auditory efferents involved in speech-in-noise intelligibility. Neuroreport, 8(7), 1779-1783.

Glenn, J. F., & Oatman, L. C. (1977). Effects of visual attention on the latency of auditory evoked potentials. Experimental neurology, 57(1), 34-40.

Gold, T., & Pumphrey, R. J. (1948). Hearing. I. The cochlea as a frequency analyzer. *Proceedings of the Royal Society of London B: Biological Sciences*, *135*(881), 462-491.

Gold, T. (1948). Hearing. II. The physical basis of the action of the cochlea. Proceedings of the Royal Society of London. Series B-Biological Sciences, 135(881), 492-498.

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, *11*(6), 559-568.

Goodman, S. S., Mertes, I. B., Lewis, J. D., & Weissbeck, D. K. (2013). Medial olivocochlear-induced transient-evoked otoacoustic emission amplitude shifts in individual subjects. Journal of the Association for Research in Otolaryngology, 14(6), 829-842.

Greenberg, G. Z., & Larkin, W. D. (1968). Frequency-Response Characteristic of Auditory Observers Detecting Signals of a Single Frequency in Noise: The Probe-Signal Method. *The Journal of the Acoustical Society of America*, *44*(6), 1513-1523.

Greenberg, S., & Ainsworth, W. A. (Eds.). (2006). *Listening to speech: an auditory perspective*. Psychology Press.

Greenwood, D. D. (1990). A cochlear frequency-position function for several species—29 years later. *The Journal of the Acoustical Society of America*, *87*(6), 2592-2605.

Guinan, J. J., Warr, W. B., & Norris, B. E. (1983). Differential olivocochlear projections from lateral versus medial zones of the superior olivary complex. Journal of Comparative Neurology, 221(3), 358-370.

Guinan Jr, J. J., & Gifford, M. L. (1988). Effects of electrical stimulation of efferent olivocochlear neurons on cat auditory-nerve fibres. I. Rate-level functions. Hearing research, 33(2), 97-113.

Guinan JJ Jr (1996) The physiology of olivocochlear efferents. In: Dallos PJ, Popper AN, Fay RR (eds) The cochlea. Springer, New York, pp 435–502.

Guinan, J. J., Lin, T., & Cheng, H. (2005). Medial-olivocochlear-efferent inhibition of the first peak of auditory-nerve responses: Evidence for a new motion within the cochlea. The Journal of the Acoustical Society of America, 118(4), 2421–2433.

Guinan Jr, J. J. (2006). Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and hearing*,*27*(6), 589-607.

Guinan, J. J. (2011). Physiology of the medial and lateral olivocochlear systems. In Auditory and vestibular efferents (pp. 39-81). Springer New York.

Guinan, J. J. (2017). Olivocochlear Efferents: Their action, effects, measurement and uses, and the impact of the new conception of cochlear mechanical responses. Hearing Research.

Hackley, S. A., Woldorff, M., & Hillyard, S. A. (1990). Cross-Modal Selective Attention Effects on Retinal, Myogenic, Brainstem, and Cerebral Evoked Potentials. *Psychophysiology*, *27*(2), 195-208.

Happe, H. K., & Morley, B. J. (1998). Nicotinic acetylcholine receptors in rat cochlear nucleus: [125I]-α-bungarotoxin receptor autoradiography and in situ hybridization of α7 nAChR subunit mRNA. Journal of Comparative Neurology, 397(2), 163-180.

Harkrider, A. W., & Smith, S. B. (2005). Acceptable noise level, phoneme recognition in noise, and measures of auditory efferent activity. *Journal of the American Academy of Audiology*, *16*(8), 530-545.

Harkrider, A. W., & Bowers, C. D. (2009). Evidence for a cortically mediated release from inhibition in the human cochlea. Journal of the American Academy of Audiology, 20(3), 208-215.

Harrison, R. V. (2012). The prevention of noise induced hearing loss in children. International journal of pediatrics, 2012.

Henry, B. A., McDermott, H. J., McKay, C. M., James, C. J., & Clark, G. M. (1998). A frequency importance function for a new monosyllabic word test. Aust. J. Audiol, 20, 79-86.

Hernandez-Peon, R., Scherrer, H., & Jouvet, M. (1956). Modification of electric activity in cochlear nucleus during" attention" in unanesthetized cats. Science, 1(956), 123.

Hickok, G., & Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. Trends in cognitive sciences, 4(4), 131-138.

Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. Nature Reviews Neuroscience, 8(5), 393-402.

Hillyard SA, Hink RF, Schwent VL, Picton TW (1973) Electrical signs of selective attention in the human brain. Science 182: 177–180.

Hirschhorn, T. N., & Michie, P. T. (1990). Brainstem auditory evoked potentials (BAEPs) and selective attention revisited. Psychophysiology, 27(5), 495-512.

Hoen, M., Grataloup, C., Veuillet, E., Thai-Van, H., Collet, L., & Meunier, F. (2008). Evaluation of auditory processing disorder and auditory efferent system functionality in adult dyslexics: towards a unification theory of auditory-language processing impairments. The Journal of the Acoustical Society of America, 123(5), 3566-3566.

Holcomb, P. J., & Neville, H. J. (1990). Auditory and visual semantic priming in lexical decision: A comparison using evoked potentials. *Language and Cognitive Processes*, *5*, 281–312.

Hood, L. J., Berlin, C. I., Hurley, A., Cecola, R. P., & Bell, B. (1996). Contralateral suppression of transient-evoked otoacoustic emissions in humans: intensity effects. *Hearing research*, *101*(1-2), 113-118.

Housley, G. D., & Ashmore, J. F. (1991). Direct measurement of the action of acetylcholine on isolated outer hair cells of the guinea pig cochlea. Proceedings of the Royal Society of London. Series B: Biological Sciences, 244(1310), 161-167.

Hubel, D. H., Henson, C. O., Rupert, A., & Galambos, R. (1959). " Attention" Units in the Auditory Cortex. Science, 129(3358), 1279-1280.

Hughson, W., & Westlake, H. D. (1944). Manual for program outline for rehabilitation of aural casualties both military and civilian. *Transactions of the American Academy of Ophthalmology and Otolaryngology, 48*(Suppl.), 1–15.

International Organization for Standardization (1997). 389-2:1997, Acoustics – Reference Zero for the Calibration of Audiometric Equipment—Part 2: Reference

Equivalent Threshold Sound Pressure Levels for Pure Tones and Insert Earphones (International Organization for Standardization, Geneva, Switzerland).

International Organization for Standardization (2006). 389-5:2006, Acoustics – Reference Zero for the Calibration of Audiometric Equipment – Part 5: Reference Equivalent Threshold Sound Pressure Levels for Pure Tones in the Frequency Range 8 kHz to 16 kHz (International Organization for Standardization, Geneva, Switzerland).

Jacobson, J., & Jacobson, C. (2004). Evaluation of hearing loss in infants and young children. Pediatric annals, 33(12), 811-821.

Jerger, J., & Hall, J. (1980). Effects of age and sex on auditory brainstem response. *Archives of Otolaryngology*, *106*(7), 387-391.

Jewett, D. L., & WILLISTON, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans. *Brain*, *94*(4), 681-696.

Jewett, D. L., & WILLISTON, J. S. (1971). Auditory-evoked far fields averaged from the scalp of humans.Brain, 94(4), 681-696.

Johnson, J. A., & Zatorre, R. J. (2006). Neural substrates for dividing and focusing attention between simultaneous auditory and visual events. Neuroimage, 31(4), 1673-1681.

Jones EG (2002) Microcolumns in the cerebral cortex. Proc Natl Acad Sci USA 97:5019–5021.

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

Kalaiah, M. K., Theruvan, N. B., Kumar, K., & Bhat, J. S. (2017a). Role of Active Listening and Listening Effort on Contralateral Suppression of Transient Evoked Otoacousic Emissions. *Journal of Audiology & Otology*, *21*(1), 1–8.

Kalaiah, M. K., Nanchirakal, J. F., Kharmawphlang, L., & Noronah, S. C. (2017b). Contralateral suppression of transient evoked otoacoustic emissions for various noise signals. *Hearing, Balance and Communication*, 1-7.

Karns, C. M., & Knight, R. T. (2009). Intermodal auditory, visual, and tactile attention modulates early stages of neural processing. Journal of cognitive neuroscience, 21(4), 669-683.

Kawase, T. E. & Liberman, M. C. (1993). Antimasking effects of the olivocochlear reflex. I. Enhancement of compound action potentials to masked tones. Journal of neurophysiology, 70(6), 2519-2532.

Kawase, T., Delgutte, B., & Liberman, M. C. (1993). Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. Journal of Neurophysiology, 70(6), 2533-2549.

Kayser, C., Petkov, C. I., Lippert, M., & Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: an auditory saliency map. *Current Biology*, *15*(21), 1943-1947.

Keidser, G., Ching, T., Dillon, H., Agung, K., Brew, C., Brewer, S., ... & Storey, L. (2002). The National Acoustic Laboratories (NAL) CDs of speech and noise for hearing aid evaluation: Normative data and potential applications. Australian and New Zealand Journal of Audiology, The, 24(1), 16.

Kemp, D. T. (1978). Stimulated acoustic emissions from within the human auditory system. The Journal of the Acoustical Society of America, 64(5), 1386-1391.

Kemp, D. T., & Brown, A. M. (1983). An integrated view of cochlear mechanical nonlinearities observable from the ear canal. In Mechanics of hearing (pp. 75-82). Springer Netherlands.

Kemp, D. T. (2002). Otoacoustic emissions, their origin in cochlear function, and use. *British medical bulletin*,*63*(1), 223-241.

Khalfa, S., Bougeard, R., Morand, N., Veuillet, E., Isnard, J., Guenot, M., & Collet, L. (2001). Evidence of peripheral auditory activity modulation by the auditory cortex in humans. Neuroscience, 104(2), 347-358.

Kirk, R. E. (2007). Experimental design. The Blackwell Encyclopedia of Sociology.

Knudsen, E. I. (2007). Fundamental components of attention. Annu. Rev. Neurosci., 30, 57-78.

Kössl, M., & Russell, I. J. (1995). Basilar membrane resonance in the cochlea of the mustached bat. Proceedings of the National Academy of Sciences, 92(1), 276-279.

Kumar, U. A., & Vanaja, C. (2004). Functioning of olivocochlear bundle and speech perception in noise. Ear and hearing, 25(2), 142-146.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207(4427), 203-205.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, *62*, 621-647.

Laurienti, P. J., Burdette, J. H., Wallace, M. T., Yen, Y. F., Field, A. S., & Stein, B. E. (2002). Deactivation of sensory-specific cortex by cross-modal stimuli. *Journal of cognitive neuroscience*, *14*(3), 420-429.

Lavie, N. (2005). Distracted and confused? Selective attention under load. Trends in cognitive sciences, 9(2), 75-82.

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. Phil. Trans. R. Soc. B, 369(1641), 20130205.

Lee, D. J., de Venecia, R. K., Guinan, J. J., & Brown, M. C. (2006). Central auditory pathways mediating the rat middle ear muscle reflexes. *The Anatomical Record*, *288*(4), 358-369.

Lehiste, I., & Peterson, G. E. (1959). Linguistic considerations in the study of speech intelligibility. The Journal of the Acoustical Society of America, 31(3), 280-286

Leinonen, A., Grönholm-Nyman, P., Järvenpää, M., Söderholm, C., Lappi, O., Laine, M., & Krause, C. M. (2009). Neurocognitive processing of auditorily and visually presented inflected words and pseudowords: evidence from a morphologically rich language. *Brain research*, *1275*, 54-66.

León, A., Elgueda, D., Silva, M. A., Hamamé, C. M., & Delano, P. H. (2012). Auditory cortex basal activity modulates cochlear responses in chinchillas. PloS one, 7(4), e36203.

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

Liberman, M. C., & Brown, M. C. (1986). Physiology and anatomy of single olivocochlear neurons in the cat. Hearing research, 24(1), 17-36

Liberman, M. C. (1988). Physiology of cochlear efferent and afferent neurons: direct comparisons in the same animal. Hearing research, 34(2), 179-191.

Liberman, M. C., Dodds, L. W., & Pierce, S. (1990). Afferent and efferent innervation of the cat cochlea: quantitative analysis with light and electron microscopy. Journal of comparative neurology, 301(3), 443-460.

Liberman MC, Guinan JJ Jr. 1998. Feedback control of the auditory periphery: anti-masking effects of middle ear muscles vs. olivocochlear efferents. *J Commun Disord* **31**: 471–483.

Lilaonitkul, W., & Guinan Jr, J. J. (2009). Human medial olivocochlear reflex: effects as functions of contralateral, ipsilateral, and bilateral elicitor bandwidths. Journal of the Association for Research in Otolaryngology, 10(3), 459-470.

Linden, J. F., & Schreiner, C. E. (2003). Columnar transformations in auditory cortex? A comparison to visual and somatosensory cortices. Cerebral Cortex, 13(1), 83-89.

Lopez-Poveda, E. A., Eustaquio-Martín, A., Stohl, J. S., Wolford, R. D., Schatzer, R., & Wilson, B. S. (2016). A binaural cochlear implant sound coding strategy inspired by the contralateral medial olivocochlear reflex. *Ear and hearing*, *37*(3), e138.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. Ear and hearing, 19(1), 1.

Luck, S. J. (2005). Ten simple rules for designing ERP experiments. Event-related potentials: A methods handbook, 262083337.

Luck, S. J. (2005). An introduction to the event-related potential technique MIT press. Cambridge, Ma, 45-64.

Lukas, J. H. (1980). Human auditory attention: the olivocochlear bundle may function as a peripheral filter. Psychophysiology, 17(5), 444-452.

Lukas, J. H. (1981). The role of efferent inhibition in human auditory attention: an examination of the auditory brainstem potentials. International Journal of Neuroscience, 12(2), 137-145.

McFadden D., and Shubel E. (2003). “The relationships between otoacoustic emissions and relative lengths of fingers and toes in humans,” Horm. Behav. 43, 421–429.

McFadden, D., Martin, G. K., Stagner, B. B., & Maloney, M. M. (2009). Sex differences in distortion-product and transient-evoked otoacoustic emissions compared. *The Journal of the Acoustical Society of America*, *125*(1), 239-246.

Macken, W. J., Tremblay, S., Houghton, R. J., Nicholls, A. P., & Jones, D. M. (2003). Does auditory streaming require attention? Evidence from attentional selectivity in short-term memory. Journal of Experimental Psychology: Human Perception and Performance, 29(1), 43.

Maddox, R. K., & Lee, A. K. (2017). Auditory brainstem responses to continuous natural speech in human listeners. *bioRxiv*, 192070.

Maison, S., Micheyl, C., & Collet, L. (1999). Sinusoidal amplitude modulation alters contralateral noise suppression of evoked otoacoustic emissions in humans. Neuroscience, 91(1), 133-138.

Maison, S., Micheyl, C., & Collet, L. (2001). Influence of focused auditory attention on cochlear activity in humans. Psychophysiology, 38(01), 35-40.

Maison, S. F., Adams, J. C., & Liberman, M. C. (2003). Olivocochlear innervation in the mouse: immunocytochemical maps, crossed versus uncrossed contributions, and transmitter colocalization. Journal of Comparative Neurology, 455(3), 406-416.

Maison, S., Liberman, L. D., & Liberman, M. C. (2016). Type II cochlear ganglion neurons do not drive the olivocochlear reflex: re-examination of the cochlear phenotype in peripherin knock-out mice. eNeuro, 3(4), ENEURO-0207.

Malmierca, M. S., & Ryugo, D. K. (2011). Descending connections of auditory cortex to the midbrain and brain stem. In *The auditory cortex* (pp. 189-208). Springer, Boston, MA.

Marslen-Wilson, W., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. Cognition, 8(1), 1-71.

McCarthy, G., & Nobre, A. C. (1993). Modulation of semantic processing by spatial selective attention. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, *88*(3), 210-219.

McCusker, L. X., Hillinger, M. L., & Bias, R. G. (1981). Phonological recoding and reading. Psychological Bulletin, 89(2), 217.

Meric, C., & Collet, L. (1992). Visual attention and evoked otoacoustic emissions: a slight but real effect. International journal of psychophysiology, 12(3), 233-235.

Meric, C., & Collet, L. (1994). Differential effects of visual attention on spontaneous and evoked otoacoustic emissions. International journal of psychophysiology, 17(3), 281-289.

Mertes, I. B., & Goodman, S. S. (2016). Within-and across-subject variability of repeated measurements of medial olivocochlear-induced changes in transient-evoked otoacoustic emissions. Ear and hearing, 37(2), e72-e84.

Merzenich, M. M., Knight, P. L., & Roth, G. L. (1975). Representation of cochlea within primary auditory cortex in the cat. Journal of neurophysiology, 38(2), 231-249.

Messing, D. P., Delhorne, L., Bruckert, E., Braida, L. D., & Ghitza, O. (2009). A non-linear efferent-inspired model of the auditory system; matching human confusions in stationary noise. Speech communication, 51(8), 668-683.

Mishra, S. K., & Lutman, M. E. (2013). Repeatability of click-evoked otoacoustic emission-based medial olivocochlear efferent assay. *Ear and hearing*, *34*(6), 789-798.

Mishra, S. K., & Lutman, M. E. (2014). Top-Down Influences of the Medial Olivocochlear Efferent System in Speech Perception in Noise. PloS one, 9(1), 57-56.

Møller, A. R., & Jannetta, P. J. (1985). Neural generators of the auditory brainstem response. The auditory brainstem response, 13-31.

Molloy, K., Griffiths, T. D., Chait, M., & Lavie, N. (2015). Inattentional deafness: visual load leads to time-specific suppression of auditory evoked responses. Journal of Neuroscience, 35(49), 16046-16054.

Moore, J. K. (1987). The human auditory brain stem as a generator of auditory evoked potentials. *Hearing research*, *29*(1), 33-43.

Motowildo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human performance*, *10*(2), 71-83.

Moulin, A., & Carrier, S. (1998). Time course of the medial olivocochlear efferent effect on otoacoustic emissions in humans. *Neuroreport*, *9*(16), 3741-3744.

Muchnik C., Ari-Even Roth D., Othman-Jebara R., Putter-Katz H., Shabtai E. L., Hildesheimer M. (2004). Reduced medial olivocochlear bundle system function in children with auditory processing disorders. Audiol. Neurootol. 9, 107–114

Mukari, S. Z. M. S., & Mamat, W. H. W. (2008). Medial olivocochlear functioning and speech perception in noise in older adults. *Audiology and Neurotology*, *13*(5), 328-334.

Mulders, W. H. A. M., & Robertson, D. (2000a). Effects on cochlear responses of activation of descending pathways from the inferior colliculus. Hearing research, 149(1), 11-23.

Mulders, W. H. A. M., & Robertson, D. (2000b). Evidence for direct cortical innervation of medial olivocochlear neurones in rats. Hearing research, 144(1), 65-72.

Murata K, Ito S, Horikawa J, Minami S. 1986. The acoustic middle ear muscle reflex in albino rats. *Hear Res* **23**: 169–183.

Murugasu, E., & Russell, I. J. (1996). The effect of efferent stimulation on basilar membrane displacement in the basal turn of the guinea pig cochlea. Journal of Neuroscience, 16(1), 325-332.

Näätänen, R. and Picton, T. (1987), The N1 Wave of the Human Electric and Magnetic Response to Sound: A Review and an Analysis of the Component Structure. Psychophysiology, 24: 375–425.

Nelson, P. B., Jin, S. H., Carney, A. E., & Nelson, D. A. (2003). Understanding speech in modulated interference: Cochlear implant users and normal-hearing listeners. *The Journal of the Acoustical Society of America*, *113*(2), 961-968.

Nobre, K. (2014). The Oxford handbook of attention. Oxford University Press

Norma, I. S. O. 8253-1: 2010 Acoustics-Audiometrics test methods.

Norman, M., & Thornton, A. R. D. (1993). Frequency analysis of the contralateral suppression of evoked otoacoustic emissions by narrow-band noise. British journal of audiology, 27(4), 281-289.

Nowotny, M., & Gummer, A. W. (2006). Nanomechanics of the subtectorial space caused by electromechanics of cochlear outer hair cells. Proceedings of the National Academy of Sciences of the United States of America, 103(7), 2120-2125.

Nowotny, M., & Gummer, A. W. (2011). Vibration responses of the organ of Corti and the tectorial membrane to electrical stimulation. The Journal of the Acoustical Society of America, 130(6), 3852-3872.

Oatman, L. C. (1971). Role of visual attention on auditory evoked potentials in unanesthetized cats. Experimental neurology, 32(3), 341-356.

Oatman, L. C. (1976). Effects of visual attention on the intensity of auditory evoked potentials. Experimental neurology, 51(1), 41-53.

O'Beirne, G. A., & Patuzzi, R. B. (1999). Basic properties of the sound-evoked post-auricular muscle response (PAMR). Hearing research, 138(1-2), 115-132.

Okita, T., & Jibu, T. (1998). Selective attention and N400 attenuation with spoken word repetition. *Psychophysiology*, *35*(3), 260-271.

Oonishi, S., & Katsuki, Y. (1965). Functional organization and integrative mechanism on the auditory cortex of the cat. The Japanese Journal of Physiology, 15(4), 342-365.

Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Computational intelligence and neuroscience, 2011, 1.

Ota, Y., Oliver, D. L., & Dolan, D. F. (2004). Frequency-specific effects on cochlear responses during activation of the inferior colliculus in the Guinea pig. Journal of neurophysiology, 91(5), 2185-2193.

Otte, J., Schuknecht, H. F., & Kerr, A. G. (1978). Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *The Laryngoscope*, *88*(8), 1231-1246.

Papanicolaou, A. C., Raz, N., Loring, D. W., & Eisenberg, H. M. (1986). Brain stem evoked response suppression during speech production. Brain and Language, 27(1), 50-55.

Patuzzi, R. (1996). Cochlear micromechanics and macromechanics. In *The cochlea* (pp. 186-257). Springer, New York, NY.

Perrot, X., Ryvlin, P., Isnard, J., Guénot, M., Catenoix, H., Fischer, C., & Collet, L. (2006). Evidence for corticofugal modulation of peripheral auditory activity in humans. Cerebral Cortex, 16(7), 941-948.

Perrot, X., & Collet, L. (2014). Function and plasticity of the medial olivocochlear system in musicians: a review. Hearing research, 308, 27-40.

Picton, T. W., Hillyard, S. A., Galambos, R., & Schiff, M. (1971). Human auditory attention: a central or peripheral process? *Science*, *173*(3994), 351-353.

Picton, T. W., Stapells, D. R., & Campbell, K. B. (1981). Auditory evoked potentials from the human cochlea and brainstem. *The Journal of Otolaryngology. Supplement*, *9*, 1-41.

Plomp, R. (2001). The intelligent ear: On the nature of sound perception. Psychology Press.

Popelar, J., Mazelova, J., & Syka, J. (2002). Effects of electrical stimulation of the inferior colliculus on 2f 1− f 2 distortion product otoacoustic emissions in anesthetized guinea pigs. Hearing research, 170(1), 116-126.

Potashner, S. J., Dymczyk, L., & Deangelis, M. M. (1988). D-aspartate uptake and release in the guinea pig spinal cord after partial ablation of the cerebral cortex. *Journal of neurochemistry*, *50*(1), 103-111.

Pratt, H., & Sohmer, H. (1976). Intensity and rate functions of cochlear and brainstem evoked responses to click stimuli in man. European Archives of Oto-Rhino-Laryngology, 212(2), 85-92.

Puel, J. L., Bonfils, P., & Pujol, R. (1988). Selective attention modifies the active micromechanical properties of the cochlea. Brain research, 447(2), 380-383.

Qin, M. K., & Oxenham, A. J. (2003). Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. *The Journal of the Acoustical Society of America*, *114*(1), 446-454.

Quinn, G. P., & Keough, M. J. (2002). Experimental design and data analysis for biologists. Cambridge University Press.

Rasmussen, G. L. (1946). The olivary peduncle and other fibre projections of the superior olivary complex. Journal of Comparative Neurology, 84(2), 141-219.

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC nonword database. The Quarterly Journal of Experimental Psychology: Section A, 55(4), 1339-1362.

Reichenbach, C. S., Braiman, C., Schiff, N. D., Hudspeth, A. J., & Reichenbach, T. (2016). The Auditory-Brainstem Response to Continuous, Non-repetitive Speech Is

Modulated by the Speech Envelope and Reflects Speech Processing. *Frontiers in computational neuroscience*, *10*.

Riecke, L., Formisano, E., Sorger, B., Başkent, D., & Gaudrain, E. (2017). Neural Entrainment to Speech Modulates Speech Intelligibility. *Current Biology*.

Roberts, W., J. Howard, et al. (1988). "Hair cells: transduction, tuning, and transmission in the inner ear." Annual review of cell biology 4(1): 63-92.

Robertson, D. (1984). Horseradish peroxidase injection of physiologically characterized afferent and efferent neurones in the guinea pig spiral ganglion. Hearing research, 15(2), 113-121.

Robertson, D., & Gummer, M. (1985). Physiological and morphological characterization of efferent neurones in the guinea pig cochlea. Hearing research, 20(1), 63-77.

Robertson, D., & Winter, I. M. (1988). Cochlear nucleus inputs to olivocochlear neurones revealed by combined anterograde and retrograde labelling in the guinea pig. Brain research, 462(1), 47-55.

Robertson D, Sellick PM, Patuzzi R (1999) The continuing search for outer hair cell afferents in the guinea pig spiral ganglion. Hear Res **136**:151–158.

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

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *336*(1278), 367-373.

Ryugo, D. K., Fay, R. R., & Popper, A. N. (Eds.). (2010). Auditory and vestibular efferents (Vol. 38). Springer Science & Business Media.

Ryugo, D.K., and T.N Parks (2003). Innervation of the cochlear nucleus in birds and mammals. Brain Res. Bull. 60:435-456.

Sanches S. G., Carvallo R. M. (2006). Contralateral suppression of transient evoked otoacoustic emissions in children with auditory processing disorder. Audiol. Neurootol. 11, 366–372.

Sarter, M., Givens, B., & Bruno, J. P. (2001). The cognitive neuroscience of sustained attention: where top-down meets bottom-up. *Brain research reviews*, *35*(2), 146-160.

Saupe, K., Schröger, E., Andersen, S. K., & Müller, M. M. (2009). Neural mechanisms of intermodal sustained selective attention with concurrently presented auditory and visual stimuli. *Frontiers in human neuroscience*, *3*.

Schaette, R., & McAlpine, D. (2011). Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. The Journal of Neuroscience, 31(38), 13452-13457.

Scharf, B., Magnan, J., & Chays, A. (1997). On the role of the olivocochlear bundle in hearing: 16 case studies. Hearing research, 103(1), 101-122.

Scharf, B., Magnan, J., Collet, L., Ulmer, E., & Chays, A. (1994). On the role of the olivocochlear bundle in hearing: a case study. Hearing research, 75(1), 11-26.

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.

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.

Schimmel, H. (1967). The (±) reference: accuracy of estimated mean components in average response studies. Science, 157(3784), 92-94.

Schofield, B. R. (2011). Central descending auditory pathways. In Auditory and vestibular efferents (pp. 261-290). Springer New York.

Schofield, B. R., & Coomes, D. L. (2005). Projections from auditory cortex contact cells in the cochlear nucleus that project to the inferior colliculus. Hearing research, 206(1), 3-11.

Shackleton, T. M., and Carlyon, R. P. (1994). “The role of resolved and unresolved harmonics in pitch perception and frequency-modulation discrimination,”J. Acoust. Soc. Am. 95, 3529–3540.

Shahin, A. J. (2011). Neurophysiological influence of musical training on speech perception. Frontiers in psychology, 2.

Shannon, R. V., Zeng, F. G., & Wygonski, J. (1998). Speech recognition with altered spectral distribution of envelope cues. The Journal of the Acoustical Society of America, 104(4), 2467-2476.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. Science, 270(5234), 303.

Shera, C. A. (2004). Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions. Ear and hearing, 25(2), 86-97.

Shera, C. A., & Guinan Jr, J. J. (1999). Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. The Journal of the Acoustical Society of America, 105(2), 782-798.

Shomstein, S., & Yantis, S. (2004). Control of attention shifts between vision and audition in human cortex. *Journal of neuroscience*, *24*(47), 10702-10706.

Siegel, J. H., & Kim, D. O. (1982). Efferent neural control of cochlear mechanics? Olivocochlear bundle stimulation affects cochlear biomechanical nonlinearity. Hearing research, 6(2), 171-182.

Siegel, J. H., and Hirohata, E. T. (1994). “Sound calibration and distortion product otoacoustic emissions at high frequencies,” Hear Res. 80, 146–152.

Silva, I. (2009). Estimation of postaverage SNR from evoked responses under nonstationary noise. IEEE Transactions on Biomedical Engineering, 56(8), 2123-2130.

Smith, D. W., et al. (2012). "Cognitive task demands modulate the sensitivity of the human cochlea." Frontiers in psychology 3.

Smith, P. H., Joris, P. X., Banks, M. I., & Yin, T. C. (1993). Responses of cochlear nucleus cells and projections of their axons. In *The Mammalian Cochlear Nuclei* (pp. 349-360). Springer, Boston, MA.

Smith, S. B., & Cone, B. (2015). The medial olivocochlear reflex in children during active listening. *International journal of audiology*, *54*(8), 518-523.

Smith, S., Kei, J., McPherson, B., & Smyth, V. (2001). Effects of speech babble on transient evoked otoacoustic emissions in normal-hearing adults. *Journal of the American Academy of Audiology*, *12*(7), 371-378.

Souza, N. N., Dhar, S., Neely, S. T., and Siegel, J. H. (2014). “Comparison of nine methods to estimate ear-canal stimulus levels,” J. Acoust. Soc. Am. 136, 1768–1787.

Sridhar, T. S., Liberman, M. C., Brown, M. C., & Sewell, W. F. (1995). A novel cholinergic" slow effect" of efferent stimulation on cochlear potentials in the guinea pig. The Journal of neuroscience, 15(5), 3667-3678.

Srinivasan, S., et al. (2012). "Effects of cross-modal selective attention on the sensory periphery: cochlear sensitivity is altered by selective attention." Neuroscience 223: 325-332.

Strauss, D. J., & Francis, A. L. (2017). Toward a taxonomic model of attention in effortful listening. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(4), 809-825.

Stuart, A., & Butler, A. K. (2012). Contralateral suppression of transient otoacoustic emissions and sentence recognition in noise in young adults. Journal of the American Academy of Audiology, 23(9), 686-696.

Suga, N., & Ma, X. (2003). Multiparametric corticofugal modulation and plasticity in the auditory system. Nature Reviews Neuroscience, 4(10), 783-794.

Sussman, E., Ritter, W., & Vaughan, H. G. (1999). An investigation of the auditory streaming effect using event-related brain potentials. Psychophysiology, 36(1), 22-34.

Suzuki, T., Kobayashi, K., & Takagi, N. (1986). Effects of stimulus repetition rate on slow and fast components of auditory brain-stem responses. Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 65(2), 150-156.

Talmadge, C. L., Tubis, A., Long, G. R., & Piskorski, P. (1998). Modeling otoacoustic emission and hearing threshold fine structures. The Journal of the Acoustical Society of America, 104(3), 1517-1543.

Tan, M. N., Robertson, D., & Hammond, G. R. (2008). Separate contributions of enhanced and suppressed sensitivity to the auditory attentional filter. Hearing research, 241(1), 18-25.

Terhardt, E., Stoll, G., & Seewann, M. (1982). Algorithm for extraction of pitch and pitch salience from complex tonal signals. The Journal of the Acoustical Society of America, 71(3), 679-688.

Terreros, G., & Delano, P. H. (2015). Corticofugal modulation of peripheral auditory responses. *Frontiers in systems neuroscience*, *9*, 134.

Thompson, A. M., & Thompson, G. C. (1991). Projections from the posteroventral cochlear nucleus to the superior olivary complex in guinea pig: Light and em observations with the PHA-L method. Journal of Comparative Neurology, 311(4), 495-508.

Thompson, A. M., & Thompson, G. C. (1993). Relationship of descending inferior colliculus projections to olivocochlear neurons. Journal of Comparative Neurology, 335(3), 402-412.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology*, *67*(6), 1176-1190.

Vaughan Jr, H. G., & Ritter, W. (1970). The sources of auditory evoked responses recorded from the human scalp. *Electroencephalography and clinical neurophysiology*, *28*(4), 360-367.

Vetter, D. E., Saldaña, E., & Mugnaini, E. (1993). Input from the inferior colliculus to medial olivocochlear neurons in the rat: a double label study with PHA-L and cholera toxin. Hearing research, 70(2), 173-186.

Veuillet, E., Collet, L., & Bazin, F. (1999). Objective evidence of peripheral auditory disorders in learning-impaired children. Journal of Audiological Medicine, 8(1), 18-29.

Veuillet, E., Magnan, A., Ecalle, J., Thai-Van, H., & Collet, L. (2007). Auditory processing disorder in children with reading disabilities: effect of audiovisual training. Brain, 130(11), 2915-2928.

Voldrich, L. (1978). Mechanical properties of basilar membrane. Acta oto-laryngologica, 86(5-6), 331-335.

Wagner, W., Heppelmann, G., Müller, J., Janssen, T., & Zenner, H. P. (2007). Olivocochlear reflex effect on human distortion product otoacoustic emissions is largest at frequencies with distinct fine structure dips. Hearing research, 223(1), 83-92.

Wagner, W., Frey, K., Heppelmann, G., Plontke, S. K., & Zenner, H. P. (2008). Speech-in-noise intelligibility does not correlate with efferent olivocochlear reflex in humans with normal hearing. *Acta oto-laryngologica*, *128*(1), 53-60.

Walsh, K. P., Pasanen, E. G., & McFadden, D. (2014a). Selective attention reduces physiological noise in the external ear canals of humans. I: Auditory attention. *Hearing research*, *312*, 143-159.

Walsh, K. P., Pasanen, E. G., & McFadden, D. (2014b). Selective attention reduces physiological noise in the external ear canals of humans. II: Visual attention. *Hearing research*, *312*, 160-167.

Waltner, J. G., & Raymond, S. (1950). On the chemical composition of the human perilymph and endolymph. The Laryngoscope, 60(9), 912-918.

Warr WB, Guinan JJJ, White JS (1986) Organization of the efferent fibres: the lateral and medial olivocochlear systems. In: Altschuler RA, Hoffman DW, Bobbin RP (ed) Neurobiology of hearing: the cochlea. Raven Press, New York, pp 333–348

Warr, W. B., & Guinan Jr, J. J. (1979). Efferent innervation of the organ of Corti: two separate systems. Brain research, 173(1), 152-155.

Warr, W. B., Boche, J. B., & Neely, S. T. (1997). Efferent innervation of the inner hair cell region: origins and terminations of two lateral olivocochlear systems. *Hearing research*, *108*(1-2), 89-111.

Warr, W. B. (1992). Organization of olivocochlear efferent systems in mammals. In *The mammalian auditory pathway: Neuroanatomy* (pp. 410-448). Springer, New York, NY.

Warren III, E. H., & Liberman, M. C. (1989). Effects of contralateral sound on auditory-nerve responses. I. Contributions of cochlear efferents. *Hearing research*, *37*(2), 89-104.

Wersinger, E., & Fuchs, P. A. (2011). Modulation of hair cell efferents. Hearing research, 279(1), 1-12.

Weisz, C., Glowatzki, E., & Fuchs, P. (2009). The postsynaptic function of type II cochlear afferents. Nature, 461(7267), 1126.

Widmann, A., Schröger, E., & Maess, B. (2015). Digital filter design for electrophysiological data–a practical approach. Journal of neuroscience methods, 250, 34-46.

Wiederhold ML (1970) Variations in the effect of electrical stimulation of the crossed olivocochlear bundle on cat single auditory nerve fibre responses to tone bursts. J Acoust Sot Am 48:966-977.

Wiederhold, M. L., & Peake, W. T. (1966). Efferent Inhibition of Auditory-Nerve Responses: Dependence on Acoustic-Stimulus Parameters. *The Journal of the Acoustical Society of America*, *40*(6), 1427-1430..

Wilson, U. S., Sadler, K. M., Hancock, K. E., Guinan Jr, J. J., & Lichtenhan, J. T. (2017). Efferent inhibition strength is a physiological correlate of hyperacusis in children with autism spectrum disorder. Journal of neurophysiology, 118(2), 1164-1172.

Winer, J. A. (1992). The functional architecture of the medial geniculate body and the primary auditory cortex. In The mammalian auditory pathway: Neuroanatomy (pp. 222-409). Springer New York.

Winslow, R. L., & Sachs, M. B. (1987). Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve response to tones in noise. J Neurophysiol, 57(4), 1002-1021.

Wirth, M., Horn, H., König, T., Stein, M., Federspiel, A., Meier, B., ... & Strik, W. (2006). Sex differences in semantic processing: event-related brain potentials distinguish between lower and higher order semantic analysis during word reading. *Cerebral Cortex*, *17*(9), 1987-1997.

Wittekindt, A., Kaiser, J., & Abel, C. (2014). Attentional Modulation of the Inner Ear: A Combined Otoacoustic Emission and EEG Study. The Journal of Neuroscience, 34(30), 9995-10002.

Woldorff, M. G., Hansen, J. C. & Hillyard, S. A. (1987) in Current Trends in Event-Related Potential Research, eds. Johnson, R., Rohrbaugh, J. W. & Parasuraman, R. (Elsevier, London), pp. 146-154.

Woldorff, M. G., Gallen, C. C., Hampson, S. A., Hillyard, S. A., Pantev, C., Sobel, D., & Bloom, F. E. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. *Proceedings of the National Academy of Sciences*, *90*(18), 8722-8726.

Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Attention, Perception, & Psychophysics*, *72*(8), 2031-2046.

Woodruff, P. W., Benson, R. R., Bandettini, P. A., Kwong, K. K., Howard, R. J., Talavage, T., ... & Rosen, B. R. (1996). Modulation of auditory and visual cortex by selective attention is modality-dependent. *Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience*.

Wright, A. (1984). Dimensions of the cochlear stereocilia in man and the guinea pig. Hearing research, 13(1), 89-98.

Wu CT, Weissman DH, Roberts KC, Woldorff MG: The neural circuitry underlying the executive control of auditory spatial attention. Brain Res 2007, 1134:187-198.

Xiao, Z., & Suga, N. (2002). Modulation of cochlear hair cells by the auditory cortex in the mustached bat. Nature neuroscience, 5(1), 57-63.

Yamagishi, S., Otsuka, S., Furukawa, S., & Kashino, M. (2016). Subcortical correlates of auditory perceptual organization in humans. Hearing research, 339, 104-111.

Yao, W., & Godfrey, D. A. (1999). Immunolocalization of α4 and α7 subunits of nicotinic receptor in rat cochlear nucleus. Hearing research, 128(1), 97-102.

Ye, Y., Machado, D. G., & Kim, D. O. (2000). Projection of the marginal shell of the anteroventral cochlear nucleus to olivocochlear neurons in the cat. *Journal of Comparative Neurology*, *420*(1), 127-138.

Zatorre, R. J., Mondor, T. A., & Evans, A. C. (1999). Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage*, *10*(5), 544-554.

Zeng, F. G., & Shannon, R. V. (1994). Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. *Science-AAAS-Weekly Paper Edition-including Guide to Scientific Information*, *264*(5158), 564-565.

Zeng, F. G., Martino, K. M., Linthicum, F. H., & Soli, S. D. (2000). Auditory perception in vestibular neurectomy subjects. Hearing research, 142(1), 102-112.

Zhang, W., & Dolan, D. F. (2006). Inferior colliculus stimulation causes similar efferent effects on ipsilateral and contralateral cochlear potentials in the guinea pig. Brain research, 1081(1), 138-149.

Zheng, J., Shen, W., He, D. Z., Long, K. B., Madison, L. D., & Dallos, P. (2000). Prestin is the motor protein of cochlear outer hair cells. *Nature*, *405*(6783), 149.

Zweig, G., & Shera, C. A. (1995). The origin of periodicity in the spectrum of evoked otoacoustic emissions. The Journal of the Acoustical Society of America, 98(4), 2018-2047.

Appendix

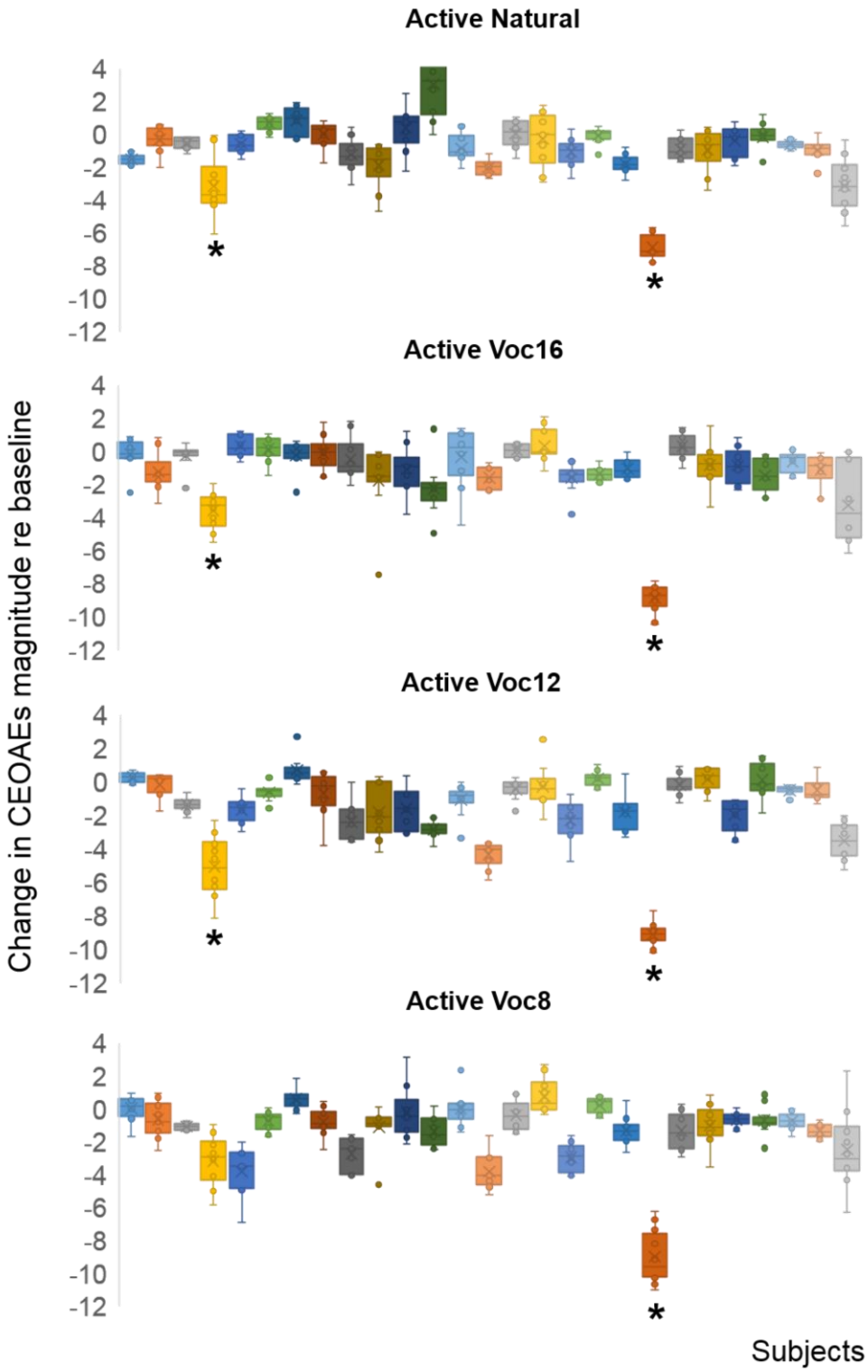

Figure 1. Individual CEOAE magnitudes relative to the baseline during the Active listening conditions in Experiment 1. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. The stars represent outliers that were not considered for statistical analysis.

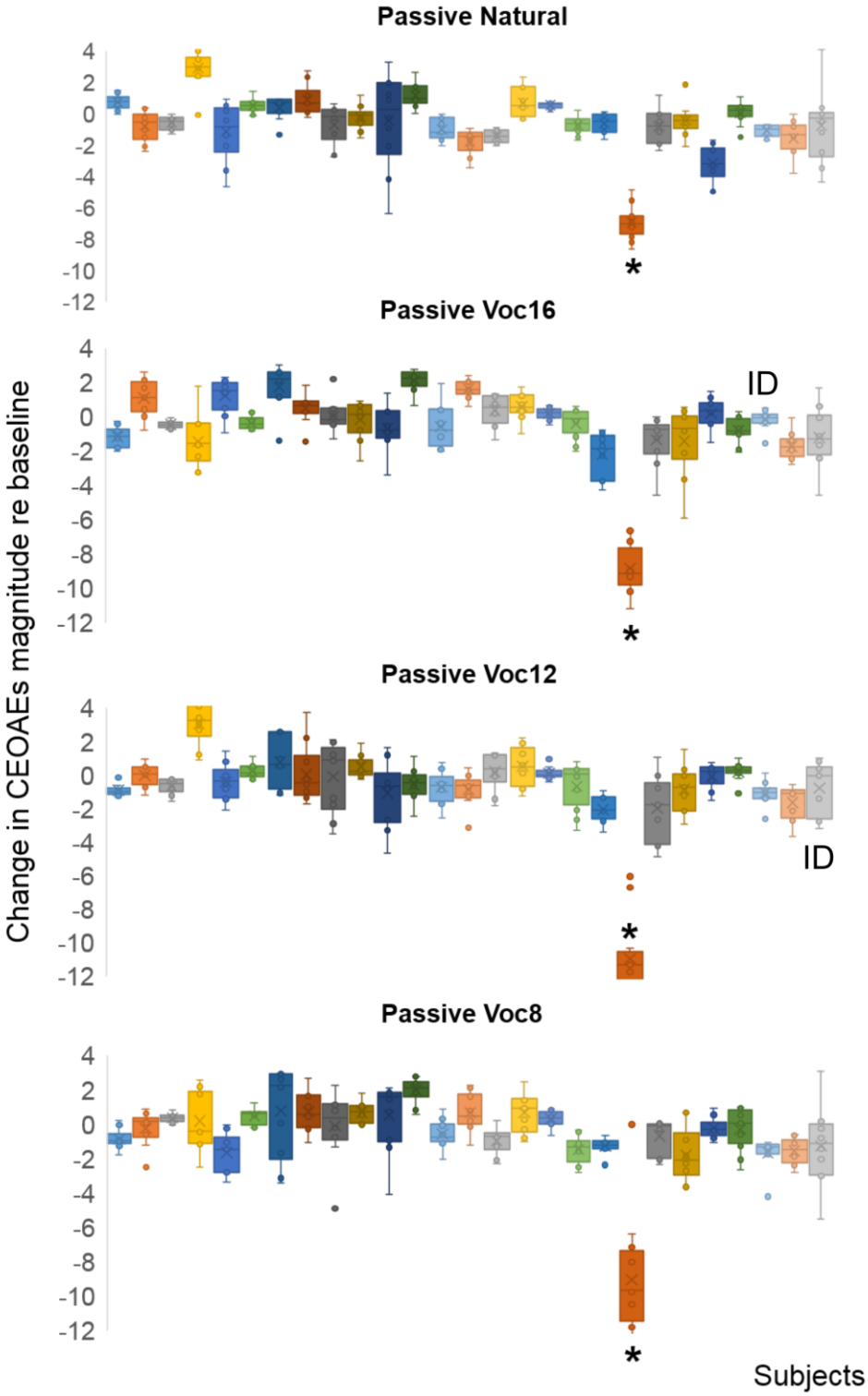

Figure 2. Individual CEOAE magnitudes relative to the baseline during the Passive listening conditions in Experiment 1. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. The stars represent outliers and IDs correspond to invalid data: both of which were not considered for statistical analysis.

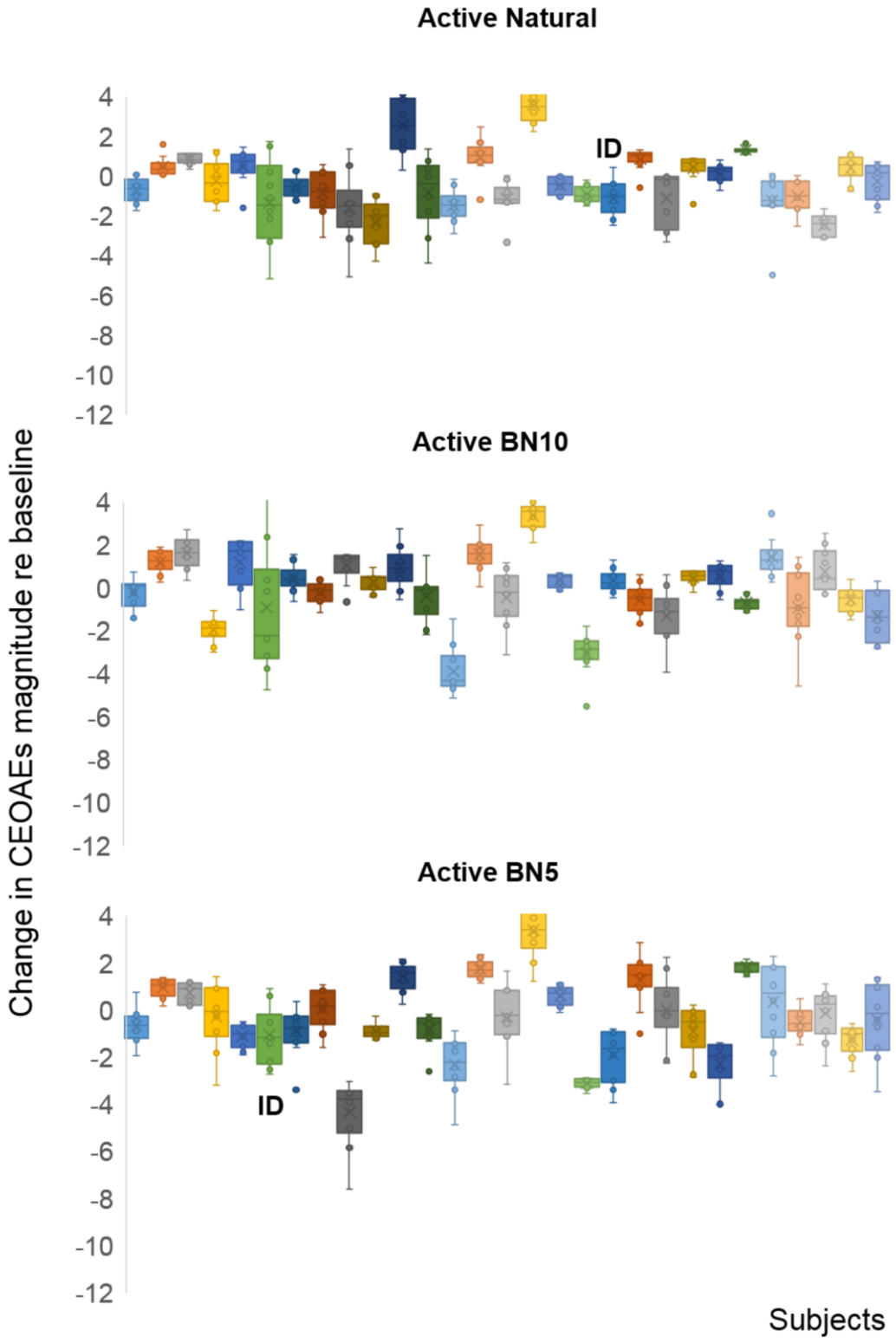

Figure 3. Individual CEOAE magnitudes relative to the baseline during the Active listening conditions in Experiment 2. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. IDs correspond to invalid data that were not considered for statistical analysis.

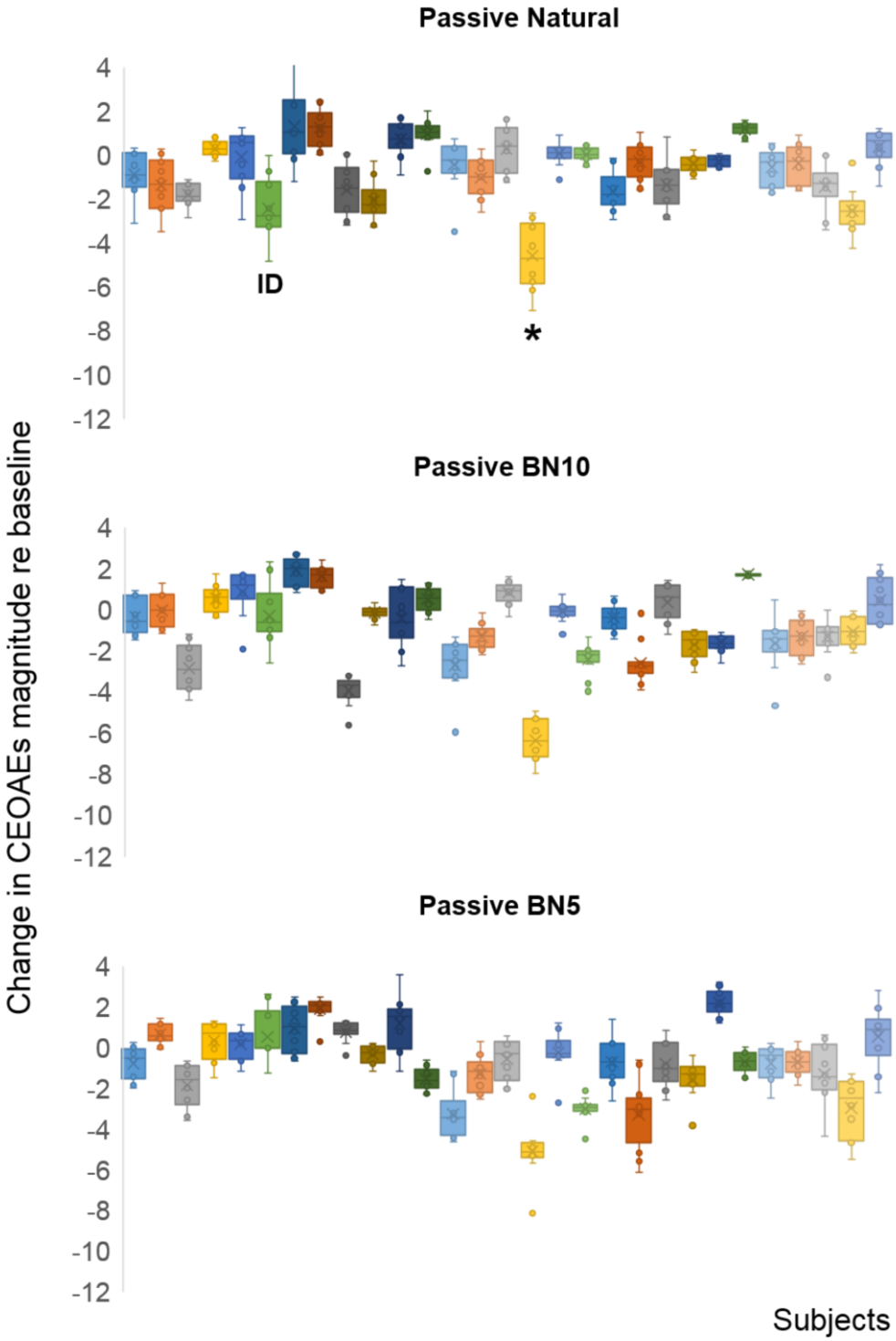

Figure 4. Individual CEOAE magnitudes relative to the baseline during the Passive listening conditions in Experiment 2. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. The star represents an outlier and ID corresponds to invalid data: both of which were not considered for statistical analysis.

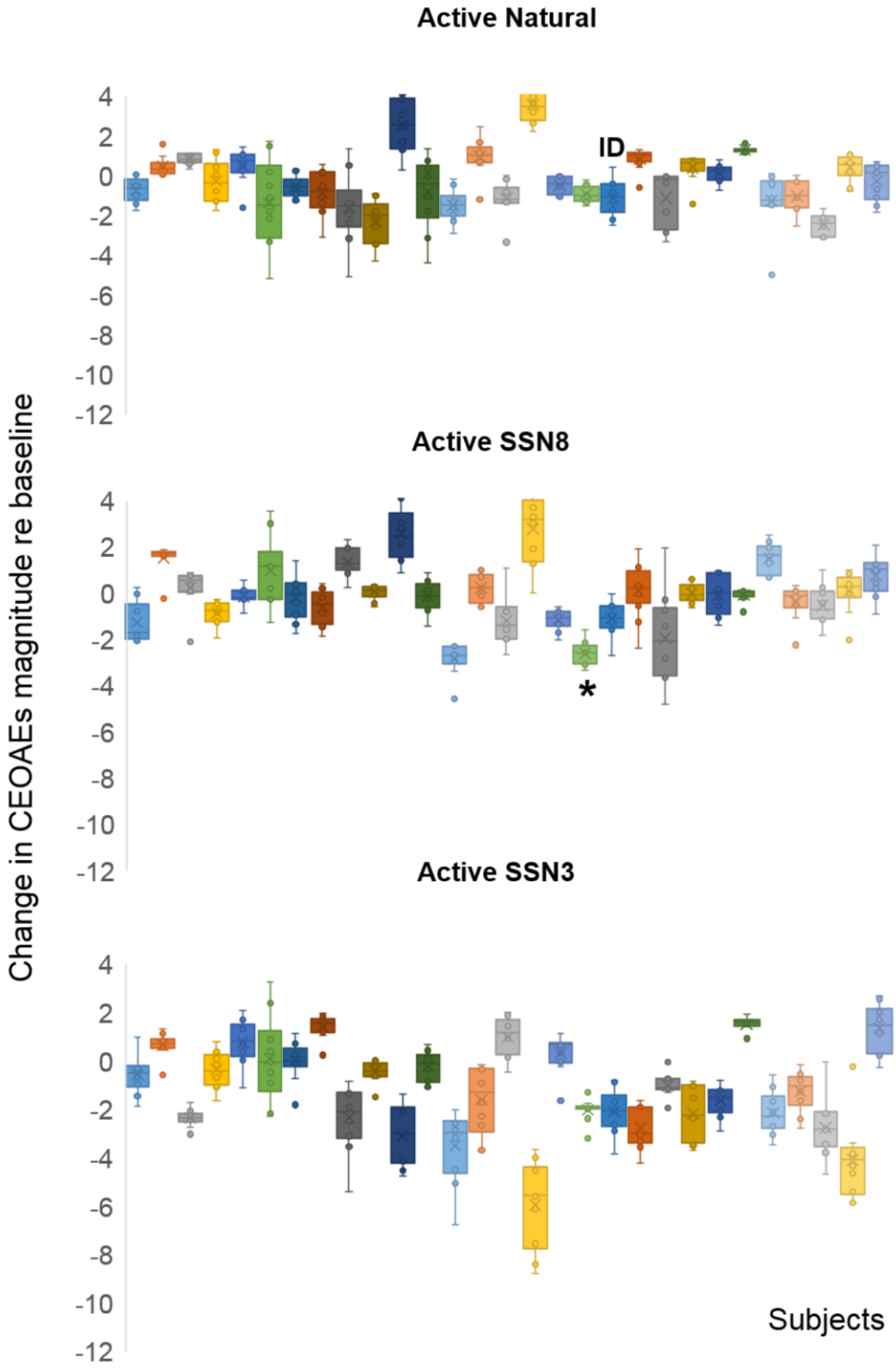

Figure 5. Individual CEOAE magnitudes relative to the baseline during the Active listening conditions in Experiment 3. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. The star represents an outlier and ID corresponds to invalid data: both of which were not considered for statistical analysis.

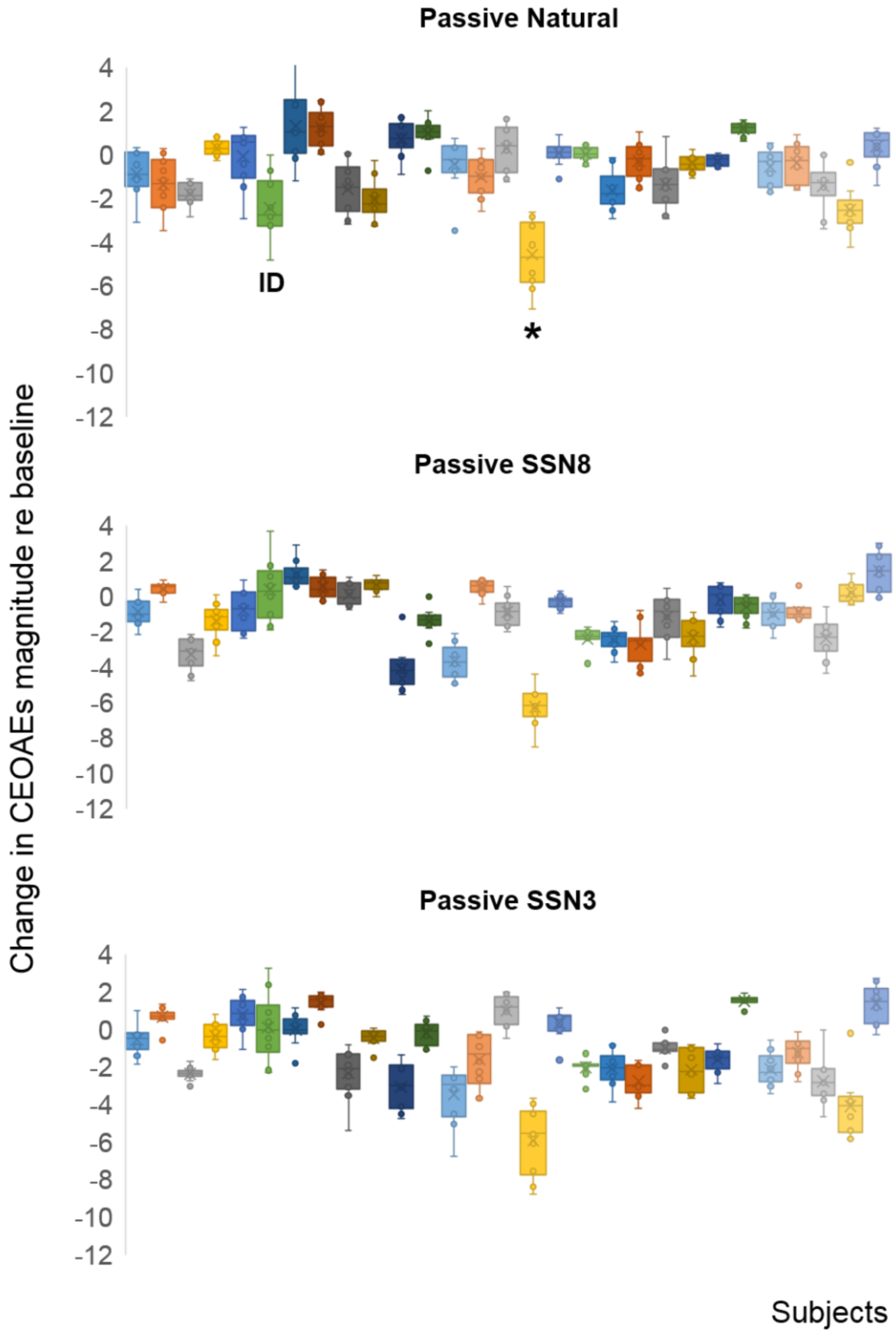

Figure 6. Individual CEOAE magnitudes relative to the baseline during the Passive listening conditions in Experiment 3. Boxes and whiskers show the distribution of the data in quartiles. Whiskers indicate the variability outside the upper and lower quartiles. The star represents an outlier and ID corresponds to invalid data: both of which were not considered for statistical analysis.

Table 1. Experiment 1 (Noise Vocoded). Description of participants with invalid ABR components.

Experimental Condition	Subject ID with invalid Wave V	Subject ID with invalid Wave III
Active Natural	5	5
Active Voc16	26	26
Active Voc12	27	27
Active Voc8		27
Passive Natural		12
Passive Voc16	25	25

Table 2. Experiment 2 (Babble Noise). Description of participants with invalid ABR components

Experimental Condition	Subject ID with invalid Wave V	Subject ID with invalid Wave III
Active Natural	16	16
Active BN10	16	16
Active BN5	16	16
Passive Natural		16
Passive BN10		16
Passive BN5		16

Table 3. Experiment 3 (Speech-Shaped Noise). Description of participants with invalid ABR components

Experimental Condition	Subject ID with invalid Wave V	Subject ID with invalid Wave III
Active Natural	16	16
Active SSN8	16	16
Active SSN3	16	16
Passive Natural		16
Passive SSN8	16	16
Passive SSN3		12

Ethics approval letter

Office of the Deputy Vice-Chancellor
(Research)

Research Office
Research Hub, Building C5C East
Macquarie University
NSW 2109 Australia
T: +61 (2) 9850 4459
http://www.research.mq.edu.au/
ABN 90 952 801 237

30 April 2015

A/Prof Catherine McMahon
Department of Linguistics
Faculty of Human Sciences
Macquarie University
NSW 2109

Dear A/Prof McMahon

Reference No: 5201500235

Title: *Effects of auditory attention on otoacoustic emission amplitudes*

Thank you for submitting the above application for ethical and scientific review. Your application was considered by the Macquarie University Human Research Ethics Committee (HREC (Medical Sciences)) at its meeting on 26 March 2015 at which further information was requested to be reviewed by the Ethics Secretariat.

The requested information was received with correspondence on 16 April 2015.

I am pleased to advise that ethical and scientific approval has been granted for this project to be conducted at:

- Macquarie University

This research meets the requirements set out in the *National Statement on Ethical Conduct in Human Research* (2007 – Updated March 2014) (the *National Statement*).

This letter constitutes ethical and scientific approval only.

Standard Conditions of Approval:

1. Continuing compliance with the requirements of the *National Statement*, which is available at the following website:

http://www.nhmrc.gov.au/book/national-statement-ethical-conduct-human-research

2. This approval is valid for five (5) years, subject to the submission of annual reports. Please submit your reports on the anniversary of the approval for this protocol.

3. All adverse events, including events which might affect the continued ethical and scientific acceptability of the project, must be reported to the HREC within 72 hours.

4. Proposed changes to the protocol must be submitted to the Committee for approval before implementation.

It is the responsibility of the Chief investigator to retain a copy of all documentation related to this project and to forward a copy of this approval letter to all personnel listed on the project.

Should you have any queries regarding your project, please contact the Ethics Secretariat on 9850 4194 or by email ethics.secretariat@mq.edu.au

The HREC (Medical Sciences) Terms of Reference and Standard Operating Procedures are available from the Research Office website at:

http://www.research.mq.edu.au/for/researchers/how_to_obtain_ethics_approval/human_research_ethics

The HREC (Medical Sciences) wishes you every success in your research.

Yours sincerely

Professor Tony Eyers
Chair, Macquarie University Human Research Ethics Committee (Medical Sciences)

This HREC is constituted and operates in accordance with the National Health and Medical Research Council's (NHMRC) *National Statement on Ethical Conduct in Human Research* (2007) and the *CPMP/ICH Note for Guidance on Good Clinical Practice*.

cc. Ms Heivet Hernandez Perez

Details of this approval are as

follows: Approval Date: 28 April

2015

The following documentation has been reviewed and approved by the HREC (Medical Sciences):

Documents reviewed	Version no.	Date
Macquarie University Ethics Application Form	2.3	July 2013
Correspondence from A/Prof McMahon responding to the issues raised by the HREC (Medical Sciences)		Received 16/4/2015
Advertisement		
MQ Participant Information and Consent Form (PICF)	2	16/4/2015
Participant Questionnaire – Edinburgh Handedness Inventory		