EXAMINING DISTRIBUTED CHANGE-DETECTION PROCESSES THROUGH CONCURRENT MEASUREMENT OF SUBCORTICAL AND CORTICAL COMPONENTS OF THE AUDITORY-EVOKED POTENTIAL

EXAMINING DISTRIBUTED CHANGE-DETECTION PROCESSES THROUGH CONCURRENT MEASUREMENT OF SUBCORTICAL AND CORTICAL COMPONENTS OF THE AUDITORY-EVOKED POTENTIAL

BY

CHRISTOPHER SLUGOCKI, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF PSYCHOLOGY, NEUROSCIENCE & BEHAVIOUR

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Doctor of Philosophy (2017) McMaster University

(Psychology, Neuroscience & Behaviour) Hamilton, Ontario, Canada

TITLE:	EXAMINING DISTRIBUTED CHANGE-DETECTION PROCESSES THROUGH CONCURRENT MEASUREMENT OF SUBCORTICAL AND CORTICAL COMPONENTS OF THE AUDITORY-EVOKED POTENTIAL
AUTHOR:	Christopher Slugocki B.Sc., (Psychology) McMaster University, Hamilton, ON, Canada
SUPERVISOR:	Dr. Laurel Trainor
NUMBER OF PAGES:	xvii, 231

To my loving wife Sarah, my caring brother Michael, and my supportive parents Zbigniew and Henryka Slugocki.

Abstract

Study of the mammalian auditory system suggests that processes once thought exclusive to cortical structures also operate subcortically. Recently, this observation has extended to the detection of acoustic change. This thesis uses methods designed for the concurrent capture of auditory-evoked potential (AEP) components attributed to different subcortical and cortical sources. Using such an approach, Chapter 2 shows that 2-month-old infants respond to infrequent changes in sound source location with neural activity implicating both subcortically- and cortically-driven mechanisms of change-detection. Chapter 3 describes the development of a new stimulation protocol and presents normative data from adult listeners showing that the morphologies of several well-known subcortical and cortical AEP components are related. Finally, Chapter 4 uses the new methods developed in Chapter 3 to demonstrate that stimulus regularity not only affects neural activity at both subcortical and cortical structures, but that the activity localized to these structures is linked. Together, the studies presented in this thesis emphasize the potential for existing technologies to study the interaction of subcortical and cortical processing in human listeners. Moreover, the results of Chapters 2 through 4 lend support to models wherein change-detection is considered a distributed, and perhaps fundamental, attribute of the auditory hierarchy.

Acknowledgements

First and foremost, I wish to express the deepest of gratitude to my advisor Dr. Laurel J. Trainor for her unwavering support throughout the many years of my Ph.D research. My successes, in no small part, speak to the strength of Laurel's patience, motivation, and academic mentorship. I consider it one of the great privileges of my life to count myself among her students.

I also wish to convey my sincerest thanks to Dr. Daniel J. Bosnyak and Dr. Paul Faure for serving as members of my supervisory committee. Their insight, encouragement, and willingness to ask difficult questions have been invaluable to strengthening the quality of my investigations and sharpening my approach to all future scientific inquiry.

I must also acknowledge the tireless efforts of Elaine Whiskin and Dave Thompson. Elaine, I am certain your hard work in the ERP Lab will continue to provide the backbone for many future theses, as it has for mine. Dave, without your technical guidance the SEAP method would have remained but an idea.

Thank you to the many members of the Auditory Development Lab, as well as the graduate community and faculty at McMaster's Department of Psychology, Neuroscience & Behaviour. It has been an honour to share in this academic journey alongside such wonderful people.

A special thanks those fine humans who tolerate my idiosyncrasies and whom I have the good fortune to call friends.

Last, but certainly not least, I would like to thank my wonderful family, my wife Sarah, my brother Michael, and my parents Zbigniew and Henryka Slugocki for supporting me through the many difficult and uncertain periods both in my graduate career and life in general. To you, I owe all things.

List of Abbreviations

A1 . . . . . primary auditory cortex

A2 . . . . . secondary auditory cortex

ABR . . . . . auditory brainstem response

AC . . . . . auditory cortex

AEP . . . . . auditory-evoked potential

AN . . . . . auditory nerve

ANF . . . . . auditory nerve fiber

ASSR . . . . . auditory steady-state response

AUC . . . . . area under the curve

CN . . . . . cochlear nucleus

dB . . . . . decibels

DFT . . . . . discrete Fourier transform

EEG . . . . . electroencephalography

ERP . . . . . event-related potential

FFR . . . . . frequency-following response

fMRI . . . . . functional magnetic resonance imaging

IC . . . . . inferior colliculus

ICc . . . . . central nucleus of the inferior colliculus

ICd . . . . . dorsal nucleus of the inferior colliculus

IIR . . . . . infinite impulse response

ISI . . . . . inter-stimulus interval

LL . . . . . lateral lemniscus

LLR . . . . . long latency response

MAA . . . . . minimum auditory angle

MGB . . . . . medial geniculate body

MGBv . . . . . ventral nucleus of the medial geniculate body

MLR . . . . . middle latency response

MMN . . . . . mismatch negativity

MMR . . . . . mismatch response

SEAP . . . . . simultaneously-evoked auditory potential

SLM . . . . . sound level meter

SPL . . . . . sound pressure level

SPL-C . . . . . C-weighted sound pressure level

SOC . . . . . superior olivary complex

SSA . . . . . stimulus-specific adaptation

TDT . . . . . Tucker-Davis Technologies

TFR . . . . . time-frequency representation

Preface

This thesis comprises three manuscripts which detail work conducted in the Auditory Development Laboratory in the Department of Psychology, Neuroscience & Behaviour at McMaster University under the supervision of Laurel J. Trainor. Each empirical chapter represents a single manuscript for which I am the primary author. The first and second empirical chapters (i.e. Chapter 2 and Chapter 3) have been published in peer-reviewed journals. Chapter 4 represents a manuscript of empirical work that has been prepared for submission to a peer-reviewed journal.

The first empirical chapter is a reprint of Slugocki, C., and Trainor, L.J. (2014). Cortical indices of sound localization mature monotonically in early infancy. *European Journal of Neuroscience, 40*, 3608–3619. I was the lead investigator on this project, responsible for experimental design, programming, data collection and analysis, as well as manuscript composition. Laurel J. Trainor was the supervisory author on this project and was involved throughout in concept formation and manuscript composition.

The second empirical chapter is a reprint of Slugocki, C., Bosnyak, D.J., and Trainor, L.J. (2017). Simultaneously-evoked auditory potentials (SEAP): A new method for concurrent measurement of cortical and subcortical auditory-evoked activity. *Hearing Research, 345*, 30–42. I was the lead investigator on this project,

responsible for all major areas of concept formation, experimental design, programming, data collection and analysis, as well as manuscript composition. Daniel J. Bosnyak was involved in the early stages of concept formation as well as approaches to data collection and analysis, and Laurel J. Trainor was the supervisory author on this project and was involved throughout in concept formation and manuscript composition.

The third empirical chapter is a manuscript, Slugocki, C., and Trainor, L.J. (*in prep*) for submission to the *Journal of the Acoustical Society of America*. I was the lead investigator on this project, responsible for all major areas of concept formation, experimental design, programming, data collection and analysis, as well as manuscript composition. Laurel J. Trainor was the supervisory author on this project and was involved throughout in concept formation and manuscript composition.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 General

The mammalian auditory system is adept at detecting change in the acoustic environment. Registering unexpected sounds can be vital to an organism's survival as such events might indicate the presence of a threat or an important communicative signal from a conspecific. In humans the mechanisms implicated in auditory change-detection involve large populations of cortical neurons and their activity can be reliably measured with scalp electrodes using electroencephalography (EEG). Change-detection neural activity can be measured through the use of a simple oddball presentation paradigm wherein a relatively infrequent deviant sound is presented pseudo-randomly in an otherwise repetitive sequence of frequently occurring standard sounds. Subtracting the averaged response to the deviant sounds from that of the standards reveals a negative peak over frontocentral electrode sites which reaches a maximum between 100-200 ms after onset of the deviant sound. This negative peak

is referred to as the mismatch negativity (MMN) and is one so-called component of the human auditory-evoked potential (AEP; Picton *et al.*, 2000).

The MMN has been the subject of much research in the field of auditory neuroscience over the last several decades (for a review see: Paavilainen, 2013). Classic conjecture regarding the neural origin of the MMN signal suggests a comparative process which judges incoming auditory information against a temporary representation of preceding sounds held in sensory memory (Näätänen *et al.*, 1989; Näätänen and Winkler, 1999). In contrast, the neural adaptation hypothesis suggests that MMN need not reflect a higher level comparative process, but simply diminished responsiveness of neurons to features of a repetitive standard sound (May and Tiitinen, 2010). More recent conceptualizations see MMN as the error signal in a predictive coding mechanism which is triggered by deviations from ongoing regularity in the acoustic environment (Lieder *et al.*, 2013; Michie *et al.*, 2016; Winkler *et al.*, 2009). This error signal might then initiate processes which update model predictions or perhaps engage attentional mechanisms and re-tune exogenous feature encoding at antecedent processing stages (Chandrasekaran *et al.*, 2014; Winkler *et al.*, 2009). Evidence that neural adaptation and perhaps even change-detection processes are measurable in AEP components localized to pre-cortical sources (Escera and Malmierca, 2014) suggests that these proposed processes require further refinement. Should predictive coding strategies indeed operate in human listeners, then their existence might be found in the relationship between cortically-generated error signals, such as the MMN, and signals reflecting the representation of stimulus periodicity and/or regularity at subcortical levels of auditory processing.

In this thesis I present three EEG studies which share a common focus on the processing of acoustic change as indexed by the MMN response. Unique to these studies is the concurrent measurement of MMN and one or more AEPs which primarily reflect processing at subcortical structures. The first study measures mismatch responses in early infancy to ask whether the development of sound localization abilities involves the abrupt transition from a reflexive subcortically-mediated localization mechanism to a volitional cortically-mediated mechanism. The second study details the development of new stimulation method for concurrent recording of cortical change-detection responses along with AEPs from multiple subcortical and cortical stages of the auditory neuraxis, and explores the interaction of these AEPs in normal-hearing adults. The third study applies this newly developed method to the investigation of distributed change-detection along the auditory hierarchy as it might involve local stimulus-specific changes to neural responsiveness at pre-cortical nuclei and/or the action of corticofugal feedback mechanisms. The thesis concludes with an examination of how these three studies reflect upon the use of concurrent measurement approaches in furthering our understanding regarding the complex interaction of change-detection processes along the human auditory pathway and a reflection on the limitations and challenges inherent therein.

1.2 Background

1.2.1 Acoustic transduction

In mammals, the encoding of acoustic information begins with transduction of sound pressure fluctuations via the inner hair cells of the organ of Corti. Input vibrations

to the cochlear partition reflect the temporal and spectral properties of the stimulating acoustic waveform after having been filtered by the head, pinnae, ear canal, and middle ear (Moore, 1997). The anatomy and physiology of the cochlea permit the spectral analysis of these input vibrations with a high degree of resolution. Through its physical properties, the basilar membrane converts vibratory patterns into travelling waves which peak more basally for stimuli of higher frequencies and more apically for stimuli of lower frequencies (Von Békésy, 1960). In turn, nonlinear mechanoelectric transduction mechanisms of the organ of Corti convert the patterned displacement of the basilar membrane into impulses in the auditory nerve (AN) fibers that innervate the inner hair cells (Moore, 1997).

Fibers constituting the AN bundle reflect an orderly innervation of hair cells along the length of the cochlear partition. By this arrangement, the frequency resolution of the basilar membrane is imparted spatially upon the AN in the frequency-selective thresholds of individual nerve fibers—a property known as tonotopy (Moore, 1997). For lower frequency sounds, the activation of AN fibers further reflects a preference for displacement of the stereocilia in the effective direction, corresponding to a compression in the stimulating waveform. In this way, the discharge of AN fibers is said to be phase-locked, albeit probabilistically, to the individual cycles of a stimulating waveform, at least up to 5 kHz (Rose *et al.*, 1971). Owing to the nonlinearity of inner hair cells, the stimulating waveform is also half-wave rectified, which in turn allows for the demodulation and encoding of envelope information (Regan and Regan, 1988). Hence, the spectral and temporal qualities of an acoustic stimulus are conveyed from the AN to the central auditory system via two mechanisms: the "place"

or spatial location of active AN fibers corresponding to particular sites of displacement on the basilar membrane, and the patterned "rate" of discharges in AN fiber ensembles. These spatial and temporal codes form the input to all later stages of the auditory system (Pickles, 2015) and are thus the basis for mechanisms of auditory discrimination and the encoding of acoustic regularity.

1.2.2 Components of the human auditory-evoked potential

Transduced acoustic energy initiates volleys of synchronous neural discharge at the many ascendant stages of the central auditory nervous system. Under certain conditions, the post-synaptic potentials generated by populations of depolarizing neurons will summate and create propagating fluctuations in voltage that can be measured at the listener's scalp using EEG electrodes. Typically, the acoustic events which elicit such activity involve relatively large changes in spectral energy occurring over a short period of time (i.e., sound onsets/offsets or gross changes in acoustic features) as such events are able to rapidly synchronize large populations of individual neurons. However, given suitable recording conditions as well as sufficiently accurate stimulating/recording equipment, one can also measure neural activity which follows periodicities in the evoking stimulus (Skoe and Kraus, 2010). In either case, visualizing the AEP commonly involves averaging operations which effectively attenuate spontaneous neural activity and thereby increase the signal-to-noise ratio for that EEG activity which is specifically related to the acoustic event. Hence, the resultant response does not capture the neural activity evoked by a single acoustic event, but rather that which is reliably evoked over anywhere from 50 to upwards of 3000 stimulus presentations (Luck, 2005).

Several so-called components have been identified in the human AEP. The exact combination of components that can be recorded in any particular experimental paradigm depends largely on the properties of the acoustic stimulus, the presentation parameters, and the task of the listener. Components can be broadly categorized as *obligatory/exogenous* if elicited on every instance of an acoustic event, or *endogenous* if elicited by certain contingencies between events or the internal state of the listener (Winkler *et al.*, 2015). Components can also be broadly categorized as *transient*, if defined by a single voltage deflection at some latency following stimulus onset, or as *sustained* if defined by ongoing responsiveness at some latency throughout stimulus duration. Of particular relevance to the current thesis, is that different transient and sustained components have been attributed to functionally discrete stages of auditory processing and, more notably, localized to different subcortical and cortical generator sites along the auditory pathway (Eggermont, 2007). A selection of these components and their purported generators is summarized below (Fig. 1.1).

Sustained components tend to reflect the activity of predominantly subcortical generators though they can also reflect the activity of thalamocortical circuits and primary auditory cortex (A1). The auditory steady-state responses (ASSRs) are one such class of sustained components defined by oscillatory activity which phase-locks to periodicities in the amplitude envelope of the evoking stimulus. Given that envelope extraction occurs peripherally (Regan and Regan, 1988), envelope-following activity can be preserved in either subcortical or cortical sources insofar as neurons at these sources are capable of sustained phase-locking at the envelope frequency (Lins *et al.*, 1995). However, the upper limit of neural phase-locking diminishes as one ascends the auditory pathway. For example, whereas neurons in the AN can exhibit

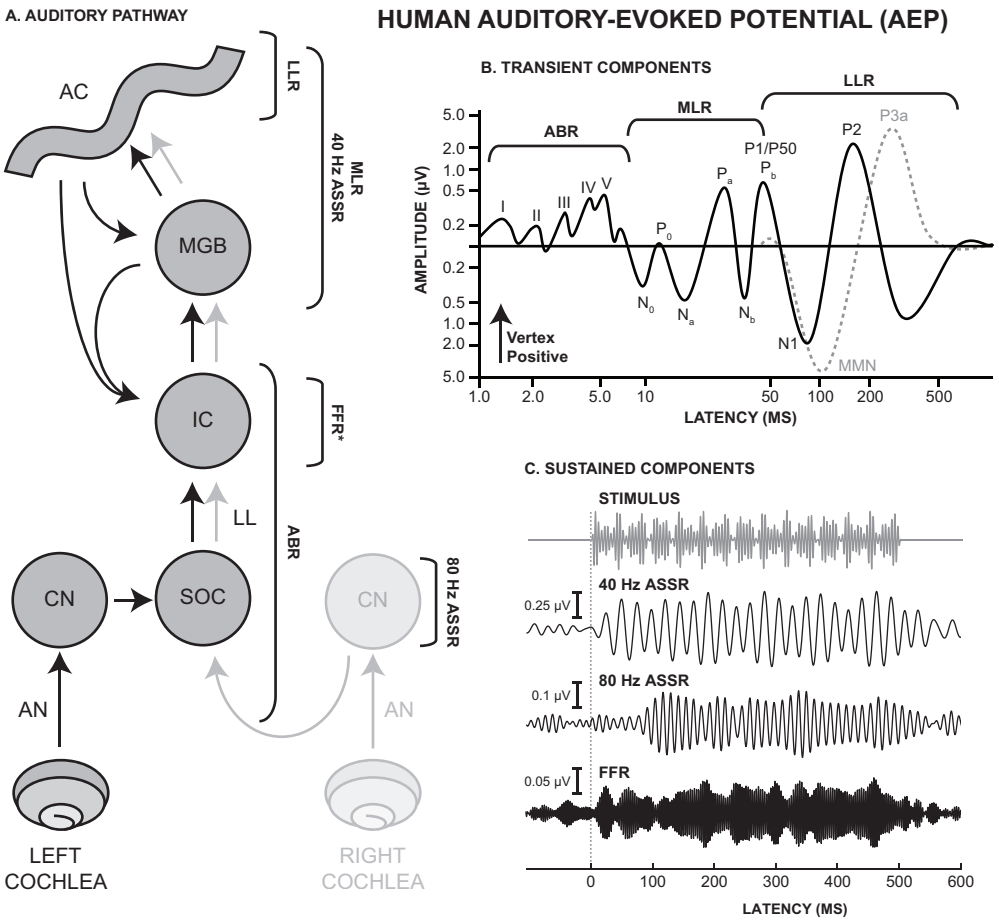

FIGURE 1.1: Summary of notable human auditory-evoked potential (AEP) components and their purported generators in the auditory system. (A) Highly simplified schematic of the mammalian auditory system showing ascending projections from the cochlea, through auditory nerve (AN) to cochlear nucleus (CN), superior olivary complex (SOC), lateral lemniscus (LL), inferior colliculus (IC), medial geniculate body of the thalamus (MGB) and auditory cortex (AC). Projections from CN cross en route to SOC for binaural integration. Part of the corticofugal system is shown in the projections from AC to MGB/IC and from MGB to IC. (B) Transient components of the human AEP at typical latencies following stimulus onset. Labels denote common peaks of the auditory brainstem response (ABR), middle latency response (MLR), and long latency responses (LLRs). The mismatch negativity (MMN) and P3a are shown in light grey as these components are only visualized in the difference between responses elicited by an infrequent compared to a frequent acoustic stimulus. (C) Sustained components of the human AEP as elicited by an amplitude modulated pure tone stimulus (top). Waveforms have been filtered to visualize 40 and 80 Hz auditory steady-state responses (ASSRs) and well as the frequency-following response (FFR*). *FFR might also contain activity from AN and AC depending on frequency content in the evoking stimulus (Coffey *et al.*, 2016, 2017).

phase-locked activity up to about 5 kHz, neurons in auditory cortex (AC) are only considered to reliably phase-lock up to about 60 Hz (for a review see: Joris *et al.*, 2004), though recent multi-unit recordings in guinea pig suggest that the upper limit of phase-locking at the population level in AC might actually be higher (∼100–120 Hz; Abrams *et al.*, 2017). Thus, different classes of ASSRs have been defined by the modulation rates evoking the largest peaks in the response spectrum. Two of the largest human ASSRs are evoked by modulation rates of around 40 and 80 Hz. Dipole solutions for these human ASSRs do indeed reflect the phase-locking gradient of the ascending auditory pathway. Whereas sources of the 40 Hz ASSR fit to both cortical and subcortical dipoles, the largest source of the 80 Hz ASSR fits mainly to brainstem generators with minimal contribution from cortical dipoles (Herdman *et al.*, 2002). The morphology of the 80 Hz ASSR is also not affected by sleep (Aoyagi *et al.*, 1993; Cohen *et al.*, 1991) or general anesthetic (e.g. propofol; Plourde *et al.*, 2008), further suggesting a pre-thalamic source. Current localization efforts based upon intracellular recordings in animal models point to cochlear nucleus (CN) as the primary subcortical source of scalp-recorded 80 Hz ASSR activity as neurons there respond best to amplitude-modulation rates above 80 Hz (Frisina *et al.*, 1990; Suzuki, 2000).

The frequency-following response (FFR) also belongs to the category of sustained AEP components. First identified by Worden and Marsh (1968), the FFR is characterized by oscillatory activity which can phase lock to periodicities of an evoking stimulus at least up to 1000 Hz (Chandrasekaran and Kraus, 2010; Smith *et al.*, 1975, 1978; Sohmer *et al.*, 1977; Worden and Marsh, 1968). Further, by applying the

appropriate recording and analysis methods, one can isolate either spectral or temporal/envelope stimulus features as they might be represented at the neural sources of the FFR signal (Aiken and Picton, 2008). The latency of the FFR signal (e.g., 5–10 ms post-stimulus onset) suggests that these sources predominantly reside in the nuclei of the midbrain, namely the inferior colliculus (IC; for a review see: Skoe and Kraus, 2010), though the FFR may also represent significant contributions from auditory nerve (Bidelman, 2015) and, at least for activity below 100 Hz, the activity of cortical generators (Coffey *et al.*, 2016).

The earliest detectable transient components of the human AEP are collectively referred to as the auditory brainstem response (ABR). The ABR is obligatory and is commonly comprised of a series of five peaks or waves (I–V in Fig. 1.1) occurring within 10 ms after stimulus onset. Roman numerals are used to label the ABR peaks in order of elicitation, following the convention set forth by Jewett and Williston (1971). Though the complete array of possible generators for each wave of the human ABR is complex (Møller and Jannetta, 1985), simplified interpretations of the ABR tend to ascribe each peak to a single anatomic source in the ascending auditory brainstem. For example, wave I is thought to reflect the first volley of action potentials arising in the distal end of the AN, whereas wave V is associated predominantly with activity in the IC. Wave V is then followed by the middle latency response (MLR) which occurs 10-50 ms post-stimulus and reflects activity of thalamocortical origin (Kraus, 1982), and that of A1 (for a review see: Alain *et al.*, 2013). The peaks of the MLR are denoted by their polarity at the vertex (P/N for positive/negative, respectively) and a number or letter indicating their place in the emerging sequence: N_0, P_0, Na, Pa, Nb, and Pb.

The transient waveforms which follow are called long-latency responses (LLRs). Those LLRs reflecting endogenous processes typically adopt names in reference to their purported functional significance, such as the ORN (object related negativity) or, as central to this thesis, the MMN (mismatch negativity; Luck, 2005). For obligatory LLRs, the naming conventions are similar to those of the MLR wherein component polarity at frontocentral sites is denoted by P/N. The numbering of LLR components can further reflect the order of component detection (i.e., N1 is the first negative-going LLR) or the component's typical peak latency (i.e., P50 is a positive-going component which occurs 50 ms after stimulus onset). The P50 is found at the border between MLR and LLRs and so is also referred to as the Pb or P1 depending on the focus of the recording. When considered part of the LLRs, P1 represents the first wave of the obligatory P1-N1-P2 complex and is thought to be generated bilaterally for pure tones and other pitch-evoking sounds (Butler and Trainor, 2012; Godey *et al.*, 2001). Little is known about the auditory P2 outside of consideration as part of the P1-N1-P2 or N1-P2 complex. Generators of the P2 have been localized to areas of secondary auditory cortex (A2) and the component has been shown sensitive to the effect of training with speech and music (Bosnyak *et al.*, 2004; Tremblay *et al.*, 2001). The auditory N1 (i.e., N100) is the most widely studied AEP that was first discovered due to its prominent negativity at the vertex, peaking at around 100 ms after the acoustic event (Davis *et al.*, 1939). The N1 is composed of subcomponents (Woods, 1995) and so its generator structure is complex. Notably, the subcomponent localizing to generators in A2 (supratemporal N1) happens to be most strongly associated with auditory processing (Godey *et al.*, 2001). Activity reflected in N1 has been associated with the indexing of exogenous stimulus features

(Alain *et al.*, 2007; Bidelman *et al.*, 2013). Similar to the preceding P1 component, N1 exhibits sensitivity to stimulus features as well as presentation rate, but N1 recovery to repeated stimulation is much slower (up to 10 s) compared to that of the P1 (∼200 ms). These observations have led some to speculate that neural adaptation at N1 generators forms the basis for auditory change-detection mechanisms (May and Tiitinen, 2010) though, as will be reviewed below, this hypothesis has largely been refuted (Näätänen *et al.*, 2011). Some have also proposed N1 to reflect feedback to AC from higher attentional centers (Shahin *et al.*, 2004; Trainor, 2007, 2012a). This hypothesis is largely based on observations that the N1 component does not become adult-like in morphology until the late teenage years (Ponton *et al.*, 2000).

1.2.3 Acoustic change-detection and the MMN

For years the mechanisms involved in detecting acoustic change were thought to rely on processes occurring at the level of the AC or beyond. Such an assumption is warranted when considered in a hierarchical framework of auditory processing. In hierarchical models, subcortical nuclei merely serve as requisite relays whose purpose is to extract and preserve the most accurate spectrotemporal representation of an impinging acoustic stimulus. Auditory cortex is the ultimate terminus of this ascending auditory information, wherein higher-level processes are then responsible for segregating the auditory sound stream into meaningful percepts such as voices or musical instruments (Näätänen and Winkler, 1999; Nelken, 2004; Zatorre *et al.*, 2004). This process is thought to be contingent on the consolidation and, thus, the transformation of spectrotemporal feature representations into discrete auditory objects which can

then be acted upon by higher cortical centers (for reviews see: Griffiths and Warren, 2004; Näätänen and Winkler, 1999). Though the precise mechanisms involved in auditory object formation are still poorly understood, they likely involve the integration of spectral, temporal, and spatial information via the complex receptive fields of cortical neurons (Linden, 2003; Nelken, 2004; Schnupp *et al.*, 2001; Shamma, 2001).

As with other sensory cortices, AC can be anatomically and physiologically subdivided into a central primary auditory field (i.e., A1) surrounded by secondary belt areas (i.e., A2; Hackett, 2015). In humans, AC is thought to comprise parts of posterior superior temporal cortex (including Heschl's gyrus), the planum temporale, and at least part of the posterior superior temporal gyrus. However, there is still no consensus regarding the number or arrangement of existing AC areas as the targets of the major input projections from thalamus are unknown. Intracortical connections in AC also suggest hierarchical organization wherein A1 sends forward projections to, and receives feedback projections from, areas in A2 (Hackett, 2011). It is in A2 where presumably more complex calculations are carried out, such as those involved in the formation and comparison of sensory memory traces (Näätänen and Winkler, 1999).

Indeed, current thought regarding A2 function can at least in part be traced to its fit as a source for MMN. Early accounts of MMN proposed that generators of the component reflect the neural basis of auditory sensory memory. In this view, spectrotemporal features of incoming sounds are stored as neural representations over several seconds so that higher-level mechanisms might extract behaviourally-relevant information. The MMN itself then reflects an automatic change-detection mechanism which responds to any discrepancy between fresh sensory input and the

regular aspects of auditory stimulation stored in sensory memory (Näätänen *et al.*, 1989). Though this mechanism of change-detection is considered to reflect a pre-attentive process, it could also trigger activity in frontal cortex resulting in attentional switch and the conscious perception of acoustic change (Näätänen *et al.*, 2007). The neural processes associated with the involuntary capture of attention are manifest in the P3a potential: a late positive component which peaks maximally over centro-parietal electrodes at 250–300 ms post-stimulus and reflects the activity of generators in A2 and frontal lobe (Polich, 2007).

Alternative explanations for the origin of MMN are based in observations of neural adaptation (e.g., forward masking), wherein the responsiveness of neurons is diminished with repeated stimulation (Brosch and Schreiner, 1997). These adaptation theories propose that the neural population encoding features of a repetitive "standard" sound becomes adapted over the course of the regular acoustic stimulation. Conversely, the neural population responding to the features of a rare "deviant" sound is not adapted and thus capable of responding more strongly compared to the "standard" counterpart. As the MMN response is constructed from the subtraction of two averaged AEPs, the differential adaptation between populations responding to standard and deviant sounds could be argued to produce the enhanced negativity which defines the MMN (May and Tiitinen, 2010).

As theories for the mechanistic origins of MMN, sensory memory and neural adaptation interpretations are alleged to be mutually exclusive (Näätänen and Alho, 1995). That is, either MMN generation involves the activity of specialized "comparator" neurons which judge incoming auditory information against a sensory memory trace or else it simply reflects a difference in adaptation to frequent and infrequent sounds.

The relative merits of each model are eloquently contrasted by Näätänen *et al.* (2005), who advocate for a memory-based interpretation, and May and Tiitinen (2010) who advocate for an interpretation based on neural adaptation. In brief, much of the debate is focused on whether the MMN can be separated from the adaptation observed in the N1 component. The discovery of stimulus-specific adaption (SSA)—a diminution in the responsiveness of single neurons to the specific features of a repetitive sound—as a property of neurons in A1 further led to speculation that a single unit correlate of scalp-recorded MMN had been found (e.g., Ulanovsky *et al.*, 2003). However, neural adaption theories must contend with decades of research showing that the processes indexed by MMN are not restricted to discriminations of basic physical sound features (i.e., frequency, amplitude, duration, or spatial location). Rather, MMN is reliably evoked by deviations from invariances which can be based in either the relationship between combinations of physical sound features and/or the rules determining stimulus occurrence in a particular auditory stream (for a review see: Paavilainen, 2013). To date, there is still no conclusive evidence to suggest SSA is sensitive to such complex acoustic regularities (Grimm *et al.*, 2016). Further, neurons exhibiting SSA are primarily "onset" types whose robust response is confined to the first 40 ms of the stimulus (Pérez-González *et al.*, 2005; Malmierca *et al.*, 2009; Pérez-González and Malmierca, 2012; Lumani and Zhang, 2010; Duque *et al.*, 2012). Thus, renewed responsiveness to novel sounds manifests much earlier than would be required to define the MMN waveform, which peaks at 100–200 ms. The short time course of SSA, of around 2000 ms (Ulanovsky *et al.*, 2004), also cannot accommodate studies showing that the traces underlying MMN generation can last considerably longer than a few seconds (for a review see: Winkler and Cowan, 2005). Nor can

SSA easily explain how MMN is evoked by the infrequent omission of a stimulus or of the second of two paired stimuli separated by a short inter-stimulus interval (Yabe *et al.*, 1998, 1997) or by infrequent decreases in stimulus amplitude (Näätänen *et al.*, 1989). Of course, it may be that both types of mechanisms operate in the human brain simultaneously, though the relative contribution of each remains unknown.

1.2.4 Subcortical mechanisms of acoustic change-detection

Though SSA might not be the single unit correlate of the MMN signal, its role in acoustic change-detection should not be readily discounted. Indeed, evidence from rodents suggests that SSA is a pervasive property of the non-lemniscal (i.e., non-tonotopic) auditory pathway from IC through the medial geniculate body (MGB) of the thalamus and up to AC (for reviews see: Ayala *et al.*, 2016; Grimm *et al.*, 2016). The discovery of SSA in the midbrain and in particular all three subdivisions of the IC (Duque *et al.*, 2012; Malmierca *et al.*, 2009; Pérez-González *et al.*, 2005), has led to the suggestion that change-detection may be a fundamental attribute of auditory processing (Grimm *et al.*, 2016).

In humans, evidence that change-detection processes operate at stages preceding the generators of MMN has been reported in the modulation of middle-latency responses (MLRs). A series of studies has shown that certain MLR peaks are enhanced in response to frequency deviants (Nb; Grimm *et al.*, 2011; Slabu *et al.*, 2010), location deviants (Na; Grimm *et al.*, 2012; Sonnadara *et al.*, 2006), intensity deviants (Pa; Althen *et al.*, 2011), and temporal deviants (Pa/Nb; Leung *et al.*, 2013). And for frequency deviants, modulation of the MLR components appears to reflect more than differential adaptation to the deviant stimulus. For example, enhancement of

deviant-evoked MLR peaks still occurs in experiments employing equiprobable control blocks. In such experiments, the responses evoked by infrequent deviants in an auditory oddball block are compared against those evoked by the same physical stimulus presented at the same probability as the oddball deviant, but interspersed among different randomly occurring equiprobable stimuli. Hence, the mismatch component in the resultant difference wave cannot be wholly attributed to differences in stimulus probability or to differential rates of associated neural adaptation (Grimm *et al.*, 2011; Recasens *et al.*, 2014b; Slabu *et al.*, 2010). Though these studies suggest that the MLR is sensitive to early mechanisms of deviance detection beyond SSA, such sensitivity remains limited to the encoding of simple acoustic regularities (i.e., comparisons against the immediately preceding sound) and not to the complex abstract regularities encoded by generators of the MMN (Grimm *et al.*, 2016).

Evidence that change-detection processes operate pre-cortically has also been reported in recent studies of the FFR component. Slabu *et al.* (2012) demonstrated that the second formant of a speech syllable, as represented in the FFR, could be modulated based on whether the speech stimulus was presented in a deviant or standard context. Similarly, Shiga *et al.* (2015) found that amplitude-modulated frequency deviants not only evoked MMN but also concurrently modulated FFR and MLR components. Functional magnetic resonance imaging (fMRI) experiments further suggest that activity related to novelty-detection can be localized to IC as well as the MGB (Cacciaglia *et al.*, 2015; Gao *et al.*, 2014). Given these developments, it is not unreasonable to speculate that change-detection processes might function along the processing hierarchy, with deviance from simple feature repetition being detected early on, and deviance from abstract regularities detected later (Grimm *et al.*, 2016).

1.2.5 Change-detection and theories of predictive coding

The evidence reviewed above has led to the hypotheses that 1) any mechanisms representing the detection of regularities might in and of themselves constitute generative models for predicting upcoming sensory events, and 2) MMN is elicited when the current sensory information does not match such predictions (Winkler *et al.*, 2009). This recent reinterpretation provides a link between MMN and predictive coding theories which posit that the objective of any sensory system is to minimize disparity between internal generative models of the external environment and coding of exogenous inputs to the system (Friston, 2005). In these theories, internal models are structured hierarchically with each increasing level representing a greater level of feature abstraction. Predictions of higher levels are tested against the inflow of information from antecedent processing stages. Any disagreement between prediction and sensory input leads to the generation of an error signal (e.g., MMN) which then guides model adjustment in a way that minimizes prediction error at all levels of the system (Winkler and Czigler, 2012; Winkler *et al.*, 2009).

1.3 Motivation for current thesis

In a predictive coding model of the auditory system, SSA in neurons of IC, MGB, and A1 might represent an early generative stage of the prediction model. Release from SSA by an unexpected acoustic event would then constitute a type of error signal to be passed up the processing hierarchy to potentially trigger processes related to MMN and model refinement. Determining whether such a system is indeed at work in human listeners requires careful consideration of how information related to the

encoding of acoustic regularities is handled at multiple stages of the processing hierarchy. Fortunately, as reviewed at the outset of this introduction and as summarized in Fig. 1.1, there are many components of the human AEP which are considered to localize to either subcortical and/or cortical sources. Further, at least some of these components exhibit sensitivity to acoustic change; however, to date relatively few studies have recorded subcortical and cortical components in the same listeners and fewer still have made efforts to limit between-subject and/or between-session variability by recording subcortical and cortical components simultaneously (Bidelman, 2015,b; Bidelman and Alain, 2015; Bidelman *et al.*, 2013, 2014a,b; Krishnan *et al.*, 2012; Musacchia *et al.*, 2008; Shiga *et al.*, 2015; Sohmer and Feinmesser, 1970; Tietze, 1979; Woods *et al.*, 1993). Of those that have, only one has included the MMN among its measures (Shiga *et al.*, 2015). Therefore, the relationship between cortical and subcortical components as well as their possible interaction in the detection of acoustic change remains an area for active investigation.

This thesis employs two approaches in its examination of distributed acoustic-change detection mechanisms along the auditory neuraxis. The first approach exploits the immaturity of cortical change-detection process in early infancy to measure the relative contribution of both processing levels to spatial deviants as such processing develops over the first 12 months of life. The second approach measures distributed auditory processing in the brains of normal-hearing adult listeners using a new stimulation method that captures the MMN, along with activity from subcortical generators, as well as A1 and A2. In this way, this thesis addresses how deviations from acoustic regularity might be represented in the differential activity of generators at several pre-cortical levels of the central auditory pathway. This thesis

also addresses how such information is relayed and related to higher level processing and how higher level processing might feedback to modulate auditory encoding at lower levels of the auditory hierarchy. The rationale for each approach and general expectations of relevant experiments are outlined in the remainder of this chapter.

1.3.1 Detecting spatial deviants in early infancy

Behavioural research suggests that the postnatal trajectory for human spatial perception is best approximated by a U-shaped function (for a review see: Muir *et al.*, 1989). The neonatal orienting response to an off-midline sound source, as observed within the first postnatal month (Crassini and Broerse, 1980; Wertheimer, 1961), disappears between 1 and 3 months of age only to re-emerge between 3 and 4 months with decreased latency, greater sensitivity, and increasing absolute precision (Clifton *et al.*, 1981; Muir *et al.*, 1979, 1989; Muir and Hains, 2004). The dominant explanation for this dip in orienting behaviour posits that neonatal localization is driven by a subcortically-mediated reflex which becomes suppressed at around 1 month of age. Suppression is thought to have a facilitative role in the transition to a more volitional and cortically-mediated sound localization mechanism (Muir and Clifton, 1985; Muir and Hains, 2004).

The frontocentral negative deflection that defines the adult MMN is predominantly positive in neonates and infants (Eggermont and Moore, 2012; Trainor, 2012a). To avoid confusion, this positivity is often referred to as the mismatch response (MMR). He *et al.* (2007) found that both the MMR and a faster adult-like negative response were observed simultaneously in the AEPs of 2-, 3-, and 4-month-olds when elicited

by infrequent piano tones. Further, the slow positive MMR is present in 2- and 3-month-olds but is insignificant in the 4-month-old group, whereas the fast adult-like negativity is present in 2-month-olds and exhibits a reduced latency and a stronger amplitude with increasing age. Hence, the authors speculated that fast-negative and slow-positive MMRs likely reflect different neural mechanisms. According to Eggermont (2006), a positive MMR suggests a depolarization site in layer IV of AC. Layer IV receives activity from the ventral medial geniculate nucleus of the thalamus via thalamocortical projections. These projections do not undergo substantial myelination until 3 months of age. As such, excitatory input from poorly myelinated thalamocortical fibers could produce poorly synchronized depolarization in layer IV and the frontocentral slow positive MMR. With increasing maturity, the emerging negative MMR/MMN reflects the new contribution of generators in superficial layers I and/or II as well as excitatory modulation of neurons in the layers II and III via intracortical projections. In this way, the slow-positive MMR can be considered to reflect subcortically-driven change-detection processes whereas the fast-negative MMR reflects the kinds of cortical mechanisms involved in generating the mature MMN (Eggermont and Moore, 2012).

The sound-localizing auditory stream begins with the projection of the anterior AN branch to the anteroventral cochlear nucleus (AVCN). Cells in the AVCN then project bilaterally to the medial superior olivary nuclei (MSO), which contribute to sound localization processes by comparing the binaural timing of acoustic information (Beckius *et al.*, 1999; Smith *et al.*, 1993). The response properties of neurons in these early brainstem structures show little sensitivity to repetitive stimulation (Ayala and Malmierca, 2013), and so their direct involvement in a predictive coding system is

unlikely. However, ascending auditory tracts of the ventral stream converge upon targets in the central nucleus of the inferior colliculus (ICc) where auditory information regarding frequency, temporal envelope, and sound source location begins to integrate (Pickles, 2015). Moreover, neurons in the ICc receive information about all three major cues to sound source location (i.e., interaural level/timing differences and monoaural spectral cues) and are often sensitive to more than one spatial cue (Chase and Young, 2008). Some animal studies even suggest that the ICc has a critical role in facilitating goal-related behaviour based on the directionality of auditory sources in the external world. For example, lesioning the brachium of the rat IC severely impairs performance on a two-choice sound localization task even at large (e.g., 180°) speaker separation angles. Such impairment is not observed to follow bilateral ablation of MGB, even at smaller (e.g., 60°) speaker angles (Kelly and Judge, 1985). Neurons in IC are also the earliest in the auditory pathway to exhibit SSA (Malmierca *et al.*, 2009). Afferent fibers relaying information about sound source location from the ICc then target the ventral division of the medial geniculate body (MGBv)—the principle auditory division of the principle auditory nucleus of the thalamus (Wenstrup, 2005)—which then projects to the middle layers (i.e., III and IV) of A1 (Smith *et al.*, 2012). Hence, SSA at any combination of neurons in the developing auditory midbrain could reasonably provide the subcortical mechanism required for detecting changes in sound source location and in turn drive the input which generates a slow-positive MMR.

Whether this mechanism is indeed suppressed in early development is the topic of Chapter 2. If infants are able to detect changes in sound source location during the period of behavioural silence then we expect spatial deviants of 90° relative to midline

standards to elicit at least the slow-positive MMR at 2-months-of-age. Alternatively, if neural localization processes, along with orienting behaviour, are suppressed between 1 and 3 months of age, then we expect to observe both mismatch responses from 5-, 8-, and 13-month-olds, but neither from 2-month-olds. Finally, we expect the slow-positive response to diminish with age and the fast-negative response to increase in amplitude and decrease in latency with increasing age. This final expectation does not suggest that subcortical responsiveness to spatial deviants disappears by 1 year-of-age, as presumably the underlying processes continue to operate throughout development. Rather, here I suggest that the mismatch response elicited by spatial change will follow the same developmental trajectory observed of other acoustic deviants (e.g., He *et al.*, 2007). Cortical maturation might reasonably be expected to direct this commonality, wherein the mechanism by which subcortical activity is able to drive slow-positive MMRs begins to increasingly interact with other largely excitatory input arriving in layer I. With increasing age, this excitatory input ultimately results in the layer I current sinks that give rise to the adult-like negative polarity of MMNs and a cancelling of the positivity generated by the current sinks in layer IV (Eggermont and Moore, 2012).

1.3.2 Hierarchical regularity encoding and corticofugal pathways

At present, little is known regarding the nature of the relationship between cortical indices of change-detection (i.e., MMN) and those purported to operate at subcortical structures (i.e., modulation of the MLR or FFR). Nor is much understood regarding

how change-detection processes operating at either level are affected by the bottom-up encoding of stimulus features in either spectral (i.e., frequency) or temporal (i.e., amplitude envelope) domains.

The sensitivity of MMN to standard probability (i.e., increased MMN amplitude with greater standard likelihood; Sato *et al.*, 2000) could reflect early adaptation to repetitive standard sounds at subcortical structures. A release from this adaptation through novel stimulation could then feedforward to enhance activity at cortical MMN generators. However, it is also possible that at least part of the SSA at structures of the midbrain and subcortex may be inherited from cortical structures via corticofugal projections (i.e., feedback; Nelken and Ulanovsky, 2007). The largest corticofugal pathways in the human brain project from AC to areas in the MGB and IC (Fig. 1.1.A; for a review see: Winer and Schreiner, 2005). Although the precise role of this corticofugal system remains unknown, physiological studies in mammals suggest that the projections act to selectively filter ascending auditory information according to central demands (He, 2003; Villa *et al.*, 1991). For example, Bäuerle *et al.* (2011) found a substantial reduction in SSA in the gerbil MGB following pharmacological lesioning of AC. Conversely, reversible deactivation of rat AC via cooling reduced but did not abolish SSA in the IC (Anderson and Malmierca, 2013) or MGB (Antunes and Malmierca, 2011), suggesting that corticofugal projections are not essential for generating subcortical SSA but may have a role in its modulation. Still, such pathways could provide the means by which predictive coding mechanisms might adjust model expectations at lower levels of the processing hierarchy.

Corticothalamic projections in particular exhibit target-dependent effects which strongly implicate their involvement in a cortically-mediated sensory filter. Afferents

from ipsilateral ICc project to the MGBv which itself projects primarily to tonotopically organized areas of A1. A dense network of descending fibers also project back to MGBv from cortical layers V and VI (Bajo *et al.*, 1995). Owing to this close reciprocity, the two auditory structures are often considered as a single functional unit. Neurons in the MGBv exhibit sharper frequency tuning curves than those of AN fibers and are able to synchronize to rapid temporal modulations (Bartlett and Wang, 2011). This sharpening can be partly contributed to local inhibition (Suga *et al.*, 1997), though it may also reflect cortical influences because manipulations of AC change the tuning curves of MGBv neurons (Zhang and Suga, 2000). In cats and guinea pigs, corticothalamic projections to MGBv appear to serve a facilitative role in that their activation is associated with an increase in spontaneous firing rates at target neurons. Conversely, activation of neurons projecting to extralemniscal thalamic nuclei (i.e., medial and dorsal divisions of MGB) has an inhibitory effect on activity at target neurons. Cooling the cortical source of these extralemniscal efferents has the opposite effect of increasing spontaneous activity at their target nuclei (Villa *et al.*, 1991). Both excitatory and inhibitory effects operate over relatively long latencies of between 300–1000 ms, though in the case of facilitation, a maximal effect is typically observed between 50–150 ms (He *et al.*, 2002). In this way, the corticothalamic system might be well-suited for rapid selective enhancement of certain frequency channels (i.e., tonotopically matched areas in AC and MGBv) while minimizing interference from non-auditory information integrating in the extralemniscal thalamic nuclei (He, 2003).

In part, the functional relationship between MGBv and A1 might be reflected in the ASSR at 40 Hz. As previously discussed, sources of the 40 Hz ASSR fit to

both cortical and subcortical dipoles (Herdman *et al.*, 2002). Whereas the subcortical source likely reflects the same envelope-following activity supporting generation of the 80 Hz ASSR, higher sources of the 40 Hz ASSR are thought to reflect the activity of a central oscillatory mechanism. One possibility is that cortical and thalamic sources of the 40 Hz ASSR support local neural oscillations via the interaction of excitatory and inhibitory connections within a thalamocortical loop. Indeed, neural network models suggest that such loops resonate maximally to periodic sensory input at frequencies close to 40 Hz (Llinás and Ribary, 2001). Moreover, sleep (Aoyagi *et al.*, 1993; Cohen *et al.*, 1991) and general anesthetic (e.g., propofol; Plourde *et al.*, 2008), both of which result in hyperpolarization of thalamocortical neurons, have been found to attenuate both subcortical and cortical sources of the 40 Hz ASSR equally, but have no effect on the 80 Hz ASSR. Previous work has also demonstrated modulation of the 40 Hz ASSR to changing stimulus parameters. For example, Ross (2008) reported phase advancement of the 40 Hz ASSR to infrequent changes in interaural phase difference resulting in a perceptual change from a focal to a spacious sound. Some researchers have further suggested that deviant-related modulation of the 40 Hz ASSR might result from fresh afferent stimulation (Chakalov *et al.*, 2014).

On the other hand, the corticocollicular system appears specialized for adapting to auditory stimuli over longer timescales by coordinating response plasticity between areas of AC and subcortical nuclei in IC (Suga *et al.*, 2000, 2002; Zhou and Jen, 2007). Both repetitive acoustic stimulation and focal electrical stimulation of the AC have been associated with stimulation-specific plastic changes involving the sharpening/broadening and/or shifting of tuning curves for IC neurons (for a review, see: Suga, 2008). Though IC neurons exhibit selective tuning to a wide variety of stimulus

features including amplitude, duration, and direction, much of the work and discussion regarding the action of corticocollicular networks in mammals tends to focus on stimulus-induced BF shifts. The modulatory effect of efferent fibers depends upon the BF relationship between target IC neurons and the stimulated cortical area. Those IC neurons sharing the same BF as the stimulated cortical neurons (i.e., "matched" neurons) exhibit augmented responsiveness at the BF and concurrent inhibition at higher and/or lower frequencies. Conversely, the responsiveness of "unmatched" IC neurons is inhibited at their intrinsic BFs and augmented at frequencies which approach the BF of the stimulated cortical neurons. In this way, the tuning curves of unmatched IC neurons are said to "shift" towards the BF of the stimulated cortical area. These effects combine to enhance subcortical input to the stimulated cortical neurons and improve the subcortical and cortical representations to which this system is tuned—a function known as ego-centric selection (Suga, 2008).

In animal models the reorganization of IC response properties is short-term, in that it disappears in about 3 hours but can be induced within 30 minutes of cortical stimulation. Associated plastic changes in AC can last for much longer (e.g., >26 hrs), though effects lasting longer than 5 hours typically require that the auditory or cortical stimulation be made behaviourally-relevant (e.g., fear conditioned with electric leg-stimulation; Gao and Suga, 2000). Studies also show that collicular BF shifts lead those in AC suggesting that the real role of corticocollicular fibers is to induce BF shifts at the IC, in turn supporting longer lasting reorganization of cortical fields (Gao and Suga, 2000). The role of the IC in auditory learning has also been

demonstrated in ferrets, where ablation of corticocollicular pathways abolishes plastic processes required for the reorganization of spatial cues following the reversible occlusion of one ear (Bajo *et al.*, 2010).

Recently, the FFR has been the subject of renewed interest as a possible means of indexing the action of the corticocollicular pathway in human listeners (for reviews see: Chandrasekaran and Kraus, 2010; Chandrasekaran *et al.*, 2014). Consistent with the action of such a feedback system, human FFR data reveal patterns of morphological change that correlate with acoustic experience, with music or language (Bidelman *et al.*, 2011; Krishnan *et al.*, 2010b,a, 2009, 2008, 2005; Krizman *et al.*, 2012; Skoe *et al.*, 2013; Wong *et al.*, 2007) as well as with short-term acoustic training (Anderson *et al.*, 2013; Carcagno and Plack, 2011; Russo *et al.*, 2005; Song *et al.*, 2008) and real-time statistical learning (Skoe and Kraus, 2010b; Skoe *et al.*, 2013). Insofar as the FFR primarily reflects the activity of subcortical nuclei, these experience-dependent effects complement a considerable body of research showing similar individual- and group-level effects in the morphology and topography of both transient and sustained components attributed to cortical generators. For example, the obligatory N1 and P2 transients—often localized to generators in secondary AC (Engelien *et al.*, 2000; Godey *et al.*, 2001; Picton *et al.*, 1999)—stay neuroplastic into adulthood and can be enhanced following auditory training (Shahin *et al.*, 2003; Tremblay *et al.*, 2001, 2002) as well as passive exposure to acoustic stimuli (Ross and Tremblay, 2009). Cortical components related to change-detection mechanisms, including the MMN and P3a, also exhibit sensitivity to passive exposure, short-term training, and long-term acoustic experience with music and language (Atienza *et al.*, 2004; Näätänen, 2008;

Nikjeh *et al.*, 2009; Uther *et al.*, 2006). However, the extent to which the morphology of the FFR is related to these cortical transient components is largely unknown.

Assessing the feedforward encoding of acoustic regularity would benefit from a measure which indexes the efficacy of early acoustic information processing. Though some studies have demonstrated reasonably accurate pure tone thresholds from tone-burst ABRs, the similarities are limited to certain frequencies (i.e., 500–4000 Hz) and require noise-masking to achieve a frequency specificity comparable to behavioural pure tone thresholds (for a review see: Stapells and Oates, 1997). As such, the ABR is not an ideal candidate for examining frequency-specific encoding of ascending auditory information. Much of the work involving the 80 Hz ASSR surrounds its usefulness as a frequency-specific alternative to the click-evoked ABR in clinical assessments of nonverbal populations (Stapells, 2008). More recently, the 80 Hz ASSR has also been examined as a means of measuring the cochlear synaptopathic origins of so-called "hidden" hearing loss (e.g., Paul *et al.*, 2017). Animal studies have shown that the 80 Hz ASSR is sensitive to synaptic losses induced by noise exposure that would otherwise stay hidden from behaviourally-derived auditory threshold measurements (Shaheen *et al.*, 2015). In this way the 80 Hz ASSR might serve as a useful index of the health of the auditory brainstem in otherwise normal hearing listeners. Such an index is especially relevant when considering how auditory encoding at structures of the brainstem affects later processing in the midbrain and cortices.

1.3.3 Simultaneously-evoked auditory potential (SEAP)

Chapters 3 and 4 employ a new method I developed for examining the relationship between concurrently-elicited subcortical and cortical components of the human AEP

termed the Simultaneously-evoked Auditory Potential (SEAP). Central to the SEAP method is a stimulus consisting of a pure tone carrier frequency that is amplitude-modulated at the sum of inharmonic frequencies around 40 and 80 Hz. The stimulus is 500 ms in duration with a 500 ms inter-stimulus interval, and is presented to participants monaurally through a calibrated acoustic delay line over 2400 trials at a fixed polarity. The carrier frequency is changed at random for 15% of trials. The SEAP method should be able to evoke all of the AEP components reviewed in Section 1.3.2. First, presentation of a fixed-polarity, pure tone carrier signal over 2070 trials should elicit a detectable FFR at the pure tone frequency (Krishnan, 2007; Thornton, 2007). Second, amplitude-modulation of the pure tone carrier at inharmonic frequencies near 40 and 80 Hz should elicit separable ASSRs at each modulation frequency. Third, passive auditory stimulation with an inter-stimulus interval of 500 ms should evoke the obligatory N1 and P2 components. Fourth, the infrequent transition in carrier frequency should elicit the MMN and, because the carrier frequencies are easily discriminated by normal-hearing adults, they should also elicit the P3a component of the AEP.

In Chapter 3, I demonstrate the viability of the SEAP method to concurrently elicit all of the aforementioned AEP components and collect normative data regarding how the subcortical AEP components (i.e., the 81 Hz ASSR and FFR) relate to cortical components (i.e., N1, P2, MMN, P3a, and the 40 Hz ASSR) within individuals during passive exposure. Chapter 4 presents SEAP stimuli in a forward-reverse oddball paradigm wherein AEPs are recorded in response to both a frequently occurring (i.e., standard) stimulus and an infrequent (i.e., deviant) stimulus that differs

from the standard in both its carrier and modulation frequencies. Stimuli are presented over two blocks with standard and deviant stimulus assignment reversed in the second block. With this approach, data from Chapter 4 are used to make within-subjects comparisons of the neural activity contributing to each aforementioned AEP component as elicited by the same physical stimulus but in different contexts. Further, Chapter 4 assesses whether change-related modulation of subcortical components (i.e., FFR) precedes or follows later cortical change-detection components (i.e., MMN and P3a), and whether the overall strength and timing of these components is related across listeners. If change-detection operates exclusively in a bottom-up (i.e., feedforward) manner then deviant-related modulation of the FFR should precede or coincide with the latency of the MMN and P3a components. On the other hand, if change-detection signals are first generated in the AC and then feedback to modulate subcortical processing in a top-down manner (i.e., model updating), then deviant-related modulation of the FFR should follow elicitation of the MMN and P3a.

1.4 References

Abrams, D. A., Nicol, T., White-Schwoch, T., Zecker, S., and Kraus, N. (2017). Population responses in primary auditory cortex simultaneously represent the temporal envelope and periodicity features in natural speech. *Hearing Research*, **348**, 31–43.

Aiken, S. J. and Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, **245**(1-2), 35–47.

Alain, C., Snyder, J. S., He, Y., and Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cerebral Cortex*, **17**(5), 1074–1084.

Alain, C., Roye, A., and Arnott, S. R. (2013). Middle-and long-latency auditory evoked potentials: what are they telling us on central auditory disorders. *Handbook of Clinical Neurophysiology: Disorders of Peripheral and Central Auditory Processing*, **10**, 177–199.

Althen, H., Grimm, S., and Escera, C. (2011). Fast detection of unexpected sound intensity decrements as revealed by human evoked potentials. *Public Library of Science One*, **6**(12), e28522.

Anderson, L. A. and Malmierca, M. S. (2013). The effect of auditory cortex deactivation on stimulus-specific adaptation in the inferior colliculus of the rat. *European Journal of Neuroscience*, **37**(1), 52–62.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(11), 4357–4362.

Antunes, F. M. and Malmierca, M. S. (2011). Effect of auditory cortex deactivation on stimulus-specific adaptation in the medial geniculate body. *The Journal of Neuroscience*, **31**(47), 17306–17316.

Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., and Koike, Y. (1993). Optimal modulation frequency for amplitude-modulation following response in young children during sleep. *Hearing Research*, **65**(12), 253–261.

Atienza, M., Cantero, J. L., and Stickgold, R. (2004). Posttraining sleep enhances automaticity in perceptual discrimination. *Journal of Cognitive Neuroscience*, **16**(1), 53–64.

Ayala, Y. A. and Malmierca, M. S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Frontiers in Neural Circuits*, **6**, 1–19.

Ayala, Y. A., Pérez-González, D., and Malmierca, M. S. (2016). Stimulus-specific adaptation in the inferior colliculus: the role of excitatory, inhibitory and modulatory inputs. *Biological Psychology*, **116**, 10–22.

Bajo, V. M., Rouiller, E. M., Welker, E., Clarke, S., Villa, A. E. P., Ribaupierre, Y. d., and Ribaupierre, F. d. (1995). Morphology and spatial distribution of corticothalamic terminals originating from the cat auditory cortex. *Hearing Research*, **83**(12), 161–174.

Bajo, V. M., Nodal, F. R., Moore, D. R., and King, A. J. (2010). The descending corticocollicular pathway mediates learning-induced auditory plasticity. *Nature Neuroscience*, **13**(2), 253–260.

Bartlett, E. L. and Wang, X. (2011). Correlation of neural response properties with auditory thalamus subdivisions in the awake marmoset. *Journal of Neurophysiology*, **105**(6), 2647–2667.

Bäuerle, P., Behrens, W. v. d., Kössl, M., and Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *The Journal of Neuroscience*, **31**(26), 9708–9722.

Beckius, G. E., Batra, R., and Oliver, D. L. (1999). Axons from anteroventral cochlear nucleus that terminate in medial superior olive of cat: observations related to delay lines. *The Journal of Neuroscience*, **19**(8), 3146–3161.

Bidelman, G. M. (2015a). Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, **323**, 68–80.

Bidelman, G. M. (2015b). Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials. *Journal of Neuroscience Methods*, **241**, 94–100.

Bidelman, G. M. and Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience*, **35**(3), 1240–1249.

Bidelman, G. M., Krishnan, A., and Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians perceptual advantages with pitch. *European Journal of Neuroscience*, **33**(3), 530–538.

Bidelman, G. M., Moreno, S., and Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage*, **79**, 201–212.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., and Alain, C. (2014a). Age-related changes in the subcorticalcortical encoding and categorical perception of speech. *Neurobiology of Aging*, **35**(11), 2526–2540.

Bidelman, G. M., Weiss, M. W., Moreno, S., and Alain, C. (2014b). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical speech perception in musicians. *European Journal of Neuroscience*, **40**(4), 2662–2673.

Bosnyak, D. J., Eaton, R. A., and Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, **14**(10), 1088–1099.

Brosch, M. and Schreiner, C. E. (1997). Time course of forward masking tuning curves in cat primary auditory cortex. *Journal of Neurophysiology*, **77**(2), 923–943.

Butler, B. E. and Trainor, L. J. (2012). Sequencing the cortical processing of pitch-evoking stimuli using EEG analysis and source estimation. *Frontiers in Psychology*, **3**, 1–13.

Cacciaglia, R., Escera, C., Slabu, L., Grimm, S., Sanjuán, A., Ventura-Campos, N., and Ávila, C. (2015). Involvement of the human midbrain and thalamus in auditory deviance detection. *Neuropsychologia*, **68**, 51–58.

Carcagno, S. and Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, **12**(1), 89–100.

Chakalov, I., Paraskevopoulos, E., Wollbrink, A., and Pantev, C. (2014). Mismatch negativity to acoustical illusion of beat: how and where the change detection takes place? *NeuroImage*, **100**, 337–346.

Chandrasekaran, B. and Kraus, N. (2010). The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology*, **47**(2), 236–246.

Chandrasekaran, B., Skoe, E., and Kraus, N. (2014). An integrative model of subcortical auditory plasticity. *Brain Topography*, **27**(4), 539–552.

Chase, S. M. and Young, E. D. (2008). Cues for sound localization are encoded in multiple aspects of spike trains in the inferior colliculus. *Journal of Neurophysiology*, **99**(4), 1672–1682.

Clifton, R. K., Morrongiello, B. A., Kulig, J. W., and Dowd, J. M. (1981). Newborns' orientation toward sound: possible implications for cortical development. *Child Development*, **52**(3), 833.

Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, **7**, 11070.

Coffey, E. B. J., Musacchia, G., and Zatorre, R. J. (2017). Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *The Journal of Neuroscience*, **37**(4), 830–838.

Cohen, L. T., Rickards, F. W., and Clark, G. M. (1991). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *The Journal of the Acoustical Society of America*, **90**(5), 2467–2479.

Crassini, B. and Broerse, J. (1980). Auditory-visual integration in neonates: a signal detection analysis. *Journal of Experimental Child Psychology*, **29**(1), 144–155.

Davis, H., Davis, P. A., Loomis, A. L., Harvey, E. N., and Hobart, G. (1939). Electrical reactions of the human brain to auditory stimulation during sleep. *Journal of Neurophysiology*, **2**(6), 500–514.

Duque, D., Pérez-González, D., Ayala, Y. A., Palmer, A. R., and Malmierca, M. S. (2012). Topographic distribution, frequency, and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *The Journal of Neuroscience*, **32**(49), 17762–17774.

Eggermont, J. (2006). Electric and magnetic fields of synchronous neural activity propogated to the surface of the head: peripheral and central origins of AEPs. In R. Burkard, M. Don, and J. Eggermont, editors, *Auditory Evoked Potentals*, pages 2–21. Lippincott Williams & Wilkins, Baltimore, MD.

Eggermont, J. (2007). Electric and magnetic fields of synchronous neural activity: peripheral and central origins of auditory evoked potentials. In *Auditory Evoked Potentials: Basic Principles and Clinical Application*. Lippincott Williams & Wilkins, Baltimore, MD.

Eggermont, J. and Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In L. Werner, R. Fay, and A. Popper, editors,

Springer Handbook of Auditory Research: Human Auditory Development, pages 61–105. Springer, New York, NY.

Engelien, A., Schulz, M., Ross, B., Arolt, V., and Pantev, C. (2000). A combined functional in vivo measure for primary and secondary auditory cortices. *Hearing Research*, **148**(12), 153–160.

Escera, C. and Malmierca, M. S. (2014). The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology*, **51**(2), 111–123.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hearing Research*, **44**(23), 99–122.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**(1456), 815–836.

Gao, E. and Suga, N. (2000). Experience-dependent plasticity in the auditory cortex and the inferior colliculus of bats: role of the corticofugal system. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(14), 8081–8086.

Gao, P. P., Zhang, J. W., Cheng, J. S., Zhou, I. Y., and Wu, E. X. (2014). The inferior colliculus is involved in deviant sound detection as revealed by BOLD fMRI. *NeuroImage*, **91**, 220–227.

Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., and Liégeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clinical Neurophysiology*, **112**(10), 1850–1859.

Griffiths, T. D. and Warren, J. D. (2004). What is an auditory object? *Nature Reviews Neuroscience*, **5**(11), 887–892.

Grimm, S., Escera, C., Slabu, L., and Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain: hierarchical organization of auditory change detection. *Psychophysiology*, **48**(3), 377–384.

Grimm, S., Recasens, M., Althen, H., and Escera, C. (2012). Ultrafast tracking of sound location changes as revealed by human auditory evoked potentials. *Biological Psychology*, **89**(1), 232–239.

Grimm, S., Escera, C., and Nelken, I. (2016). Early indices of deviance detection in humans and animal models. *Biological Psychology*, **116**, 23–27.

Hackett, T. A. (2011). Information flow in the auditory cortical network. *Hearing Research*, **271**(12), 133–146.

Hackett, T. A. (2015). Anatomic rganization of the auditory cortex. In G. G. Celesia and G. Hickok, editors, *The Human Auditory System: Fundamental Organization and Clinical Disorders*, volume 129 of *Handbook of Clinical Neurology*, pages 27–53. Elsevier, Edinburgh.

He, C., Hotson, L., and Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, **19**(5), 878–892.

He, J. (2003). Corticofugal modulation of the auditory thalamus. *Experimental Brain Research*, **153**(4), 579–590.

He, J., Yu, Y.-Q., Xiong, Y., Hashikawa, T., and Chan, Y.-S. (2002). Modulatory effect of cortical activation on the lemniscal auditory thalamus of the guinea pig. *Journal of Neurophysiology*, **88**(2), 1040–1050.

Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M., and Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, **15**(2), 69–86.

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

Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews*, **84**(2), 541–577.

Kelly, J. B. and Judge, P. W. (1985). Effects of medial geniculate lesions on sound localization by the rat. *Journal of Neurophysiology*, **53**(2), 361–372.

Kraus, N. (1982). Auditory middle latency responses (MLRs) in patients with cortical lesions. *Electroencephalography and Clinical Neurophysiology*, **54**(3), 275–287.

Krishnan, A. (2007). Frequency-following response. In R. Burkard, J. Eggermont, and M. Don, editors, *Auditory Evoked Potentials: Basic Principles and Clinical Application*, pages 313–335. Lippincott Williams & Wilkins, Philadelphia, PA.

Krishnan, A., Xu, Y., Gandour, J., and Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, **25**(1), 161–168.

Krishnan, A., Swaminathan, J., and Gandour, J. T. (2008). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, **21**(6), 1092–1105.

Krishnan, A., Gandour, J. T., Bidelman, G. M., and Swaminathan, J. (2009). Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, **20**(4), 408–413.

Krishnan, A., Gandour, J. T., and Bidelman, G. M. (2010a). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, **23**(1), 81–95.

Krishnan, A., Bidelman, G. M., and Gandour, J. T. (2010b). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, **268**(12), 60–66.

Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., and Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, **50**(12), 2849–2859.

Krizman, J., Marian, V., Shook, A., Skoe, E., and Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(20), 7877–7881.

Leung, S., Recasens, M., Grimm, S., and Escera, C. (2013). Electrophysiological index of acoustic temporal regularity violation in the middle latency range. *Clinical Neurophysiology*, **124**(12), 2397–2405.

Lieder, F., Stephan, K. E., Daunizeau, J., Garrido, M. I., and Friston, K. J. (2013). A neurocomputational model of the mismatch negativity. *Public Library of Science Computational Biology*, **9**(11), e1003288.

Linden, J. F. (2003). Spectrotemporal structure of receptive fields in areas AI and AAF of mouse auditory cortex. *Journal of Neurophysiology*, **90**(4), 2660–2675.

Lins, O. G., Picton, P. E., Picton, T. W., Champagne, S. C., and Durieux-Smith, A. (1995). Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. *The Journal of the Acoustical Society of America*, **97**(5), 3051–3063.

Llinás, R. and Ribary, U. (2001). Consciousness and the brain. *Annals of the New York Academy of Sciences*, **929**(1), 166–175.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press, Cambridge, MA.

Lumani, A. and Zhang, H. (2010). Responses of neurons in the rat's dorsal cortex of the inferior colliculus to monaural tone bursts. *Brain Research*, **1351**, 115–129.

Malmierca, M. S., Cristaudo, S., Pérez-González, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *The Journal of Neuroscience*, **29**(17), 5483–5493.

May, P. J. C. and Tiitinen, H. (2010). Mismatch negativity (MMN), the deviance-elicited auditory deflection, explained. *Psychophysiology*, **47**(1), 66–122.

Michie, P. T., Malmierca, M. S., Harms, L., and Todd, J. (2016). The neurobiology of MMN and implications for schizophrenia. *Biological Psychology*, **116**, 90–97.

Møller, A. R. and Jannetta, P. J. (1985). Neural generators of the auditory brainstem response. *The Auditory Brainstem Response*, pages 13–31.

Moore, B. (1997). *Introduction to the Psychology of Hearing*. Academic Press, London, UK, 4th edition.

Muir, D. and Clifton, R. K. (1985). Infants' orientation to the location of sound sources. In G. Gottlieb and N. Krasnegor, editors, *The Measurement of Audition and Vision During the First Year of Life: A Methodological Overview*, pages 171–194. Ablex, Norwood, NJ.

Muir, D. and Hains, S. (2004). The U-shaped developmental function for auditory localization. *Journal of Cognition and Development*, **5**(1), 123–130.

Muir, D., Abraham, W., Forbes, B., and Harris, L. (1979). The ontogenesis of an auditory localization response from birth to four months of age. *Canadian Journal of Psychology*, **33**(4), 320–333.

Muir, D. W., Clifton, R. K., and Clarkson, M. G. (1989). The development of a human auditory localization response: a U-shaped function. *Canadian Journal of Psychology*, **43**(2), 199–216.

Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, **241**(12), 34–42.

Näätänen, R. (2008). Mismatch negativity (MMN) as an index of central auditory system plasticity. *International Journal of Audiology*, **47**(sup2), S16–S20.

Näätänen, R. and Alho, K. (1995). Mismatch negativity-a unique measure of sensory processing in audition. *International Journal of Neuroscience*, **80**(1-4), 317–337.

Näätänen, R. and Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, **125**(6), 826–859.

Näätänen, R., Paavilainen, P., Alho, K., Reinikainen, K., and Sams, M. (1989). Do event-related potentials reveal the mechanism of the auditory sensory memory in the human brain? *Neuroscience Letters*, **98**, 217–221.

Näätänen, R., Jacobsen, T., and Winkler, I. (2005). Memory-based or afferent processes in mismatch negativity (MMN): a review of the evidence. *Psychophysiology*, **42**(1), 25–32.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, **118**(12), 2544–2590.

Näätänen, R., Kujala, T., and Winkler, I. (2011). Auditory processing that leads to conscious perception: a unique window to central auditory processing opened by the mismatch negativity and related responses. *Psychophysiology*, **48**(1), 4–22.

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Current Opinion in Neurobiology*, **14**(4), 474–480.

Nelken, I. and Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology*, **21**(3-4), 214–223.

Nikjeh, D. A., Lister, J. J., and Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training. *Ear and Hearing*, **30**(4), 432–446.

Paavilainen, P. (2013). The mismatch-negativity (MMN) component of the auditory event-related potential to violations of abstract regularities: a review. *International Journal of Psychophysiology*, **88**(2), 109–123.

Paul, B. T., Bruce, I. C., and Roberts, L. E. (2017). Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus. *Hearing Research*, **344**, 170–182.

Pérez-González, D. and Malmierca, M. S. (2012). Variability of the time course of stimulus-specific adaptation in the inferior colliculus. *Frontiers in Neural Circuits*, **6**, 1–12.

Pérez-González, D., Malmierca, M. S., and Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience*, **22**(11), 2879–2885.

Pickles, J. O. (2015). Auditory pathways: anatomy and physiology. In G. G. Celesia and G. Hickok, editors, *The Human Auditory System: Fundamental Organization and Clinical Disorders*, volume 129 of *Handbook of clinical neurology*. Elsevier, Edinburgh, UK.

Picton, T. W., Alain, C., Woods, D., John, M., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., and Trujillo, N. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology and Neurotology*, **4**, 64–79.

Picton, T. W., Alain, C., Otten, L., Ritter, W., and Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology*, **5**(3-4), 111–139.

Plourde, G., Garcia-Asensi, A., Backman, S., Deschamps, A., Chartrand, D., Fiset, P., and Picton, T. W. (2008). Attenuation of the 40-Hz auditory steady state response by propofol involves the cortical and subcortical generators. *Anesthesiology*, **108**, 233–242.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–2148.

Ponton, C. W., Eggermont, J., Kwong, B., and Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, **111**(2), 220–236.

Recasens, M., Grimm, S., Capilla, A., Nowak, R., and Escera, C. (2014). Two sequential processes of change detection in hierarchically ordered areas of the human auditory cortex. *Cerebral Cortex*, **24**(1), 143–153.

Regan, D. and Regan, M. P. (1988). The transducer characteristic of hair cells in the human ear: a possible objective measure. *Brain Research*, **438**(12), 363–365.

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

Ross, B. (2008). A novel type of auditory responses: temporal dynamics of 40-Hz steady-state responses induced by changes in sound localization. *Journal of Neurophysiology*, **100**(3), 1265–1277.

Ross, B. and Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, **248**(12), 48–59.

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., and Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, **156**(1), 95–103.

Sato, Y., Yabe, H., Hiruma, T., Sutoh, T., Shinozaki, N., Nashida, T., and Kaneko, S. (2000). The effect of deviant stimulus probability on the human mismatch process. *Neuroreport*, **11**(17), 3703–3708.

Schnupp, J. W., Mrsic-Flogel, T. D., and King, A. J. (2001). Linear processing of spatial cues in primary auditory cortex. *Nature*, **414**, 200–204.

Shaheen, L. A., Valero, M. D., and Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. *Journal of the Association for Research in Otolaryngology*, **16**(6), 727–745.

Shahin, A., Bosnyak, D. J., Trainor, L. J., and Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience*, **23**(13), 5545–5552.

Shahin, A., Roberts, L. E., and Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport*, **15**(12), 1917–1921.

Shamma, S. (2001). On the role of space and time in auditory processing. *Trends in Cognitive Sciences*, **5**(8), 340–348.

Shiga, T., Althen, H., Cornella, M., Zarnowiec, K., Yabe, H., and Escera, C. (2015). Deviance-related responses along the auditory hierarchy: combined FFR, MLR and MMN evidence. *Public Library of Science One*, **10**(9), e0136794.

Skoe, E. and Kraus, N. (2010a). Auditory brainstem response to complex sounds: a tutorial. *Ear and Hearing*, **31**(3), 302–324.

Skoe, E. and Kraus, N. (2010b). Hearing it again and again: on-line subcortical plasticity in humans. *Public Library of Science One*, **5**(10), e13645.

Skoe, E., Krizman, J., Spitzer, E., and Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, **243**, 104–114.

Slabu, L., Escera, C., Grimm, S., and Costa-Faidella, J. (2010). Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials: early auditory change detection in humans. *European Journal of Neuroscience*, **32**(5), 859–865.

Slabu, L., Grimm, S., and Escera, C. (2012). Novelty detection in the human auditory brainstem. *The Journal of Neuroscience*, **32**(4), 1447–1452.

Smith, J. C., Marsh, J. T., and Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, **39**(5), 465–472.

Smith, J. C., Marsh, J. T., Greenberg, S., and Brown, W. S. (1978). Human auditory frequency-following responses to a missing fundamental. *Science*, **201**(4356), 639–641.

Smith, P. H., Joris, P. X., and Yin, T. C. T. (1993). Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *The Journal of Comparative Neurology*, **331**(2), 245–260.

Smith, P. H., Uhlrich, D. J., Manning, K. A., and Banks, M. I. (2012). Thalamocortical projections to rat auditory cortex from the ventral and dorsal divisions of the medial geniculate nucleus. *The Journal of Comparative Neurology*, **520**(1), 34–51.

Sohmer, H. and Feinmesser, M. (1970). Cochlear and cortical audiometry conveniently recorded in the same subject. *Israel Journal of Medical Science*, **6**, 219–223.

Sohmer, H., Pratt, H., and Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology*, **42**(5), 656–664.

Song, J. H., Skoe, E., Wong, P. C., and Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, **20**(10), 1892–1902.

Sonnadara, R. R., Alain, C., and Trainor, L. J. (2006). Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location. *Brain Research*, **1071**(1), 175–185.

Stapells, D. R. (2008). The 80-Hz auditory steady-state response compared with other auditory evoked potentials. In G. Rance, editor, *Auditory Steady-State Responses: Generation, Recording, and Clinical Application*, pages 149–160. Plural Publishing, San Diego, CA.

Stapells, D. R. and Oates, P. (1997). Estimation of the pure-tone audiogram by the auditory brainstem response: a review. *Audiology and Neurotology*, **2**(5), 257–280.

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A*, **194**(2), 169–183.

Suga, N., Zhang, Y., and Yan, J. (1997). Sharpening of frequency tuning by inhibition in the thalamic auditory nucleus of the mustached bat. *Journal of Neurophysiology*, **77**(4), 2098–2114.

Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(22), 11807–11814.

Suga, N., Xiao, Z., Ma, X., and Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, **36**(1), 9–18.

Suzuki, Y. (2000). Contribution of cochlear nucleus to 80 Hz amplitude-modulation following response. *Nippon Jibiinkoka Gakkai Kaiho*, **103**, 177–187.

Thornton, A. (2007). Instrumentation and recording parameters. In R. Burkard, J. Eggermont, and M. Don, editors, *Auditory Evoked Potentials: Basic Principles and Clinical Application*, pages 42–47. Lippincott Williams & Wilkins, Philadelphia, PA.

Tietze, G. (1979). Stimulation methods for simultaneous derivation of acoustically evoked brainstem and cortical responses. *Scandinavian Audiology. Supplementum*, **11**, 97–104.

Trainor, L. J. (2007). Event-related potential (ERP) measures in auditory development research. In L. Schmidt and S. Segalowitz, editors, *Developmental Psychophysiology: Theory, Systems, and Methods*, pages 69–102. Cambridge University Press, New York, NY.

Trainor, L. J. (2012). Musical experience, plasticity, and maturation: issues in measuring developmental change using EEG and MEG. *Annals of the New York Academy of Sciences*, **1252**(1), 25–36.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, B., and others (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and Hearing*, **22**(2), 79–90.

Tremblay, K. L., Piskosz, M., and Souza, P. (2002). Aging alters the neural representation of speech cues. *Neuroreport*, **13**(15), 1865–1870.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, **6**(4), 391–398.

Ulanovsky, N., Las, L., Farkas, D., and Nelken, I. (2004). Multiple time scales of adaptation in auditory cortex neurons. *The Journal of Neuroscience*, **24**(46), 10440–10453.

Uther, M., Kujala, A., Huotilainen, M., Shtyrov, Y., and Näätänen, R. (2006). Training in Morse code enhances involuntary attentional switching to acoustic frequency: evidence from ERPs. *Brain Research*, **1073-1074**, 417–424.

Villa, A. E. P., Rouiller, E. M., Simm, G. M., Zurita, P., Ribaupierre, Y. d., and Ribaupierre, F. d. (1991). Corticofugal modulation of the information processing in the auditory thalamus of the cat. *Experimental Brain Research*, **86**(3), 506–517.

Von Békésy, G. (1960). *Experiments in Hearing*, volume 8. McGraw-Hill New York.

Wenstrup, J. J. (2005). The tectothalamic system. In J. A. Winer and C. E. Schreiner, editors, *The Inferior Colliculus*, pages 200–230. Springer New York.

Wertheimer, M. (1961). Psychomotor coordination of auditory and visual space at birth. *Science*, **134**, 1692.

Winer, J. A. and Schreiner, C., editors (2005). *The Inferior Colliculus*. Springer, New York, NY.

Winkler, I. and Cowan, N. (2005). From sensory to long-term memory: evidence from auditory memory reactivation studies. *Experimental Psychology*, **52**(1), 3–20.

Winkler, I. and Czigler, I. (2012). Evidence from auditory and visual event-related potential (ERP) studies of deviance detection (MMN and vMMN) linking predictive coding theories and perceptual object representations. *International Journal of Psychophysiology*, **83**(2), 132–143.

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, **13**(12), 532–540.

Winkler, I., Denham, S., and Escera, C. (2015). Auditory event-related potentials. In D. Jaeger and R. Jung, editors, *Encyclopedia of Computational Neuroscience*, pages 209–233. Springer Science+Business Media, New York.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, **10**(4), 420–422.

Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Perspectives of Event-Related Potentials Research*, **44**, 102–209.

Woods, D. L., Alain, C., Covarrubias, D., and Zaidel, O. (1993). Frequency-related differences in the speed of human auditory processing. *Hearing Research*, **66**(1), 46–52.

Worden, F. G. and Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology*, **25**(1), 42–52.

Yabe, H., Tervaniemi, M., Reinikainen, K., and Näätänen, R. (1997). Temporal window of integration revealed by MMN to sound omission. *Neuroreport*, **8**(8), 1971–1974.

Yabe, H., Tervaniemi, M., Sinkkonen, J., Huotilainen, M., Ilmoniemi, R. J., and Näätänen, R. (1998). Temporal window of integration of auditory information in the human brain. *Psychophysiology*, **35**(5), 615–619.

Zatorre, R. J., Bouffard, M., and Belin, P. (2004). Sensitivity to auditory object features in human temporal neocortex. *The Journal of Neuroscience*, **24**(14), 3637–3642.

Zhang, Y. and Suga, N. (2000). Modulation of responses and frequency tuning of thalamic and collicular neurons by cortical activation in mustached bats. *Journal of Neurophysiology*, **84**(1), 325–333.

Zhou, X. and Jen, P. H.-S. (2007). Corticofugal modulation of multi-parametric auditory selectivity in the midbrain of the big brown bat. *Journal of Neurophysiology*, **98**(5), 2509–2516.

Chapter 2

Cortical indices of sound localization mature monotonically in early infancy

2.1 Abstract

In human neonates, orienting behaviour in response to an off-midline sound source disappears around the first postnatal month, but re-emerges at 4 months. It is unclear whether sound localization processes continue to operate between postnatal months 1 and 3. Here we used an event-related potential that reflects change detection in the auditory cortices to measure cortical responses elicited by large (±90° relative to midline), infrequent changes in sound source location in 2-, 5-, 8- and 13-month-old infants. Both fast-negative mismatch negativity (MMN; Näätänen *et al.*, 2007) and slow-positive mismatch response (MMR; Trainor *et al.*, 2003) were elicited from all age groups. However, both components were smaller and the fast-negative component

occurred later in the 2-month-old group than in older age groups. Additionally, the slow-positive component diminished in amplitude with increasing age, whereas the fast-negative component grew larger and occurred earlier with increasing age. These results suggest that the cortical representation of sound location matures similarly to representations of pitch and duration. A subsequent investigation of 2-month-old infants confirmed that the observed MMR and MMN were elicited by changes in sound source location, and were not merely attributable to changes in loudness cues. The presence of both MMR and MMN in the 2-month-old group indicates that the cortex is able to detect changes in sound location despite the behavioural silence observed around 1–3 months of age.

2.2 Introduction

Research suggests that the postnatal development of human spatial hearing is best approximated by a U-shaped function (reviewed by: Muir *et al.*, 1989). Neonatal orienting responses, as observed within the first postnatal month (Wertheimer, 1961; Crassini and Broerse, 1980), disappear between 1 and 3 months of age, but re-emerge between 3 and 4 months of age with decreased latency, greater sensitivity, and greater absolute precision (Muir *et al.*, 1979, 1989; Clifton *et al.*, 1981; Muir and Hains, 2004). The dominant explanation for this dip in orienting behaviour posits that neonatal localization is driven by a subcortically mediated reflex that becomes suppressed at $\sim$1 month of age. Suppression is thought to have a facilitative role in the transition to a more volitional and cortically mediated sound localization mechanism (Muir and Clifton, 1985; Muir and Hains, 2004).

It is not clear whether infants 1–3 months-of-age are able to localize sounds and, if they are, whether they are doing so at a cortical level. Here, we addressed this question by measuring electroencephalographic indices of pre-attentive change-detection in the auditory cortices, known as mismatch response (MMR) and mismatch negativity (MMN), in a cross-sectional sample of infants between 2 and 13 months of age. Developmental studies are consistent in observing a maturation of the infant mismatch response. In very young infants a slow-positivity defines the MMR, but this slow-positive component diminishes in amplitude with increasing age. A fast-negative wave, resembling the MMN response of adults and older children, then emerges in older infants and increases in amplitude with increasing age (for reviews see: Trainor, 2007, 2012a,b). In adults, the MMN originates primarily in the auditory cortex (for a review, see Näätänen *et al.* 2007), and its features, such as amplitude and latency, are highly correlated with behavioural measures of discrimination thresholds for many acoustic features, including pitch (e.g., Novitski *et al.* 2004), gap detection (Trainor *et al.*, 2001), and, of relevance to the current study, sound localization (Paavilainen *et al.*, 1989; Schröger, 1996; Schröger and Wolff, 1996; Kaiser *et al.*, 2000a,b; Altman *et al.*, 2005; Altmann *et al.*, 2009; Altman *et al.*, 2010; Deouell *et al.*, 2006; Sonnadara *et al.*, 2006; Pakarinen *et al.*, 2007; Röttger *et al.*, 2007; Spierer *et al.*, 2007; Vasilenko and Shestopalova, 2010; Grimm *et al.*, 2012; Bennemann *et al.*, 2013)

The first experiment in the present study employed an oddball paradigm, such that in 80% of trials a sound was presented from directly in front of the infant at midline, and in the remaining trials it was presented from a location $\pm 90^{\circ}$ re midline. If the infant auditory system is indeed sensitive to changes in sound source location during the period of behavioural silence, then we expect deviant trials to elicit one or both

mismatch components from all age groups. Furthermore, if 2-month-old localization is driven primarily by subcortical processes, then we expect the immature slow-positive MMR to dominate the 2-month-old response to deviant trials and an adult-like MMN response to emerge later. Alternatively, if neural localization processes, along with orienting behaviour, are suppressed between 1 and 3 months of age, then we expect to observe both mismatch responses from 5-, 8-, and 13-month-olds, but neither from 2-month-olds. Finally, we expect the slow-positive response to diminish with age and the fast-negative response to increase in amplitude and decrease in latency with increasing age and cortical maturation.

A second experiment, conducted in a new group of 2-month-old infants, followed the procedures outlined for the first experiment, except that the amplitude of the stimulus was pseudo-randomly varied by ±4 dB to control for changes in stimulus loudness that accompany ±90° shifts in azimuthal sound source origin at higher frequencies (Blauert, 1983).

2.3 Materials and methods

2.3.1 Experiment 1

Participants

We recruited a total of 73 healthy full-term infants, with no known hearing deficits, ranging in age from 2 to 13 months. Prior to the experiment, basic demographic information was obtained from each infant's caregiver. The final sample consisted of 16 infants aged 2 months (six female; mean age = 2.51, SD = 0.20 months), 13 infants aged 5 months (five female; mean age = 5.49, SD = 0.22 months), nine infants aged

8 months (six female; mean age = 8.46, SD = 0.31 months), and 14 infants aged 13 months (five female; mean age = 13.05, SD = 0.44 months). Nineteen infants (four aged 2 months, 13 aged 5 months, and two aged 8 months) were excluded from the analyses because of excessive fussiness (e.g., crying or not facing the center speaker) and/or too few artifact-free trials. Two additional infants (one aged 2 months and one aged 8 months) were excluded from the final sample because they fell asleep during testing. All infants included in the final sample were awake during the course of the experiment. Written consent was obtained from all parents of the infants in compliance with a protocol approved by the McMaster Research Ethics Board. The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* (18 July 1964).

Stimulus

A burst of frozen white noise (duration, 300 ms) was generated with Audacity software (www.audacity.sourceforge.net). The stimulus was presented with a Tucker-Davis Technologies (TDT) RP2 Real Time Processor, relayed through a TDT PM2R Power Multiplexer connected to a Hafler P1000 Trans.Ana 100-W amplifier. Stimulus presentation occurred in a sound-attenuated room through matching AudioVideo Methods speakers (P73) at an onset asynchrony of 500 ms and at an average C-weighted sound pressure level (SPL-C) of 76.5 dB.

Procedure

The infant was seated on the caregiver's lap. Both caregiver and infant faced the center speaker positioned 1 m away. Additional loudspeakers were aligned at ±90°

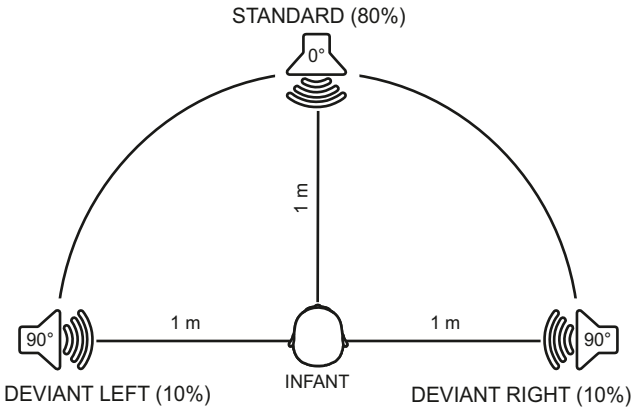

FIGURE 2.1: Diagram of the experimental design. The infant was seated on the caregiver's lap in a sound-attenuating room. Both infant and caregiver were positioned at a distance of 1 m from three identical speakers. Stimulation was provided as bursts of frozen white noise emanating from one of the three speakers in a random order over 2400 trials (center, 80%; deviant left, 10%; and deviant right, 10%).

relative to the center speaker, at a distance of 1 m from the infant and caregiver (FIG. 2.1). A silent video (Baby Einstein) and a live puppet show were used to keep the infant happy and to maintain the infant's attention in the direction of the center speaker. Stimulus presentation followed an oddball paradigm. Standard stimuli, defined as presentations through the center speaker, accounted for 80% of trials. Deviant stimuli, defined as presentations through one of the lateral speakers ($\pm 90°$), accounted for the remaining 20% of trials. Deviant stimulation was further divided such that trials occurred equally often between the left (10%) and right (10%) lateral speakers. The complete test session consisted of 2400 trials, although the experiment was stopped early if the infant became fussy or fell asleep. Of those infants included in the final sample, the number of accepted trials ranged from 868 to 2064 (mean = 1815, SD = 207).

Data acquisition and analysis

During testing, we collected continuous electroencephalograms from 129 channels (referenced to vertex) with an Electrical Geodesics NetAmps 200 amplifier passing the digitized signal to Electrical Geodesics netstation software (v.4.3.1). Signals were digitized at a rate of 1000 Hz and subjected to online filtering between 0.1 and 400 Hz. The electrical impedance of each electrode was maintained below 50 kΩ.

The continuous data were further filtered offline in eeprobe software (Advanced Neuro Technologies) with two different bandpass filter settings: 0.5-20 Hz and 3-18 Hz. Filtering the data between 0.5 and 20 Hz is most effective at revealing the MMR, whereas filtering between 3 and 18 Hz effectively removes the slow-wave MMR and allows for visualization of the faster adult-like MMN. Thus, for each subject, the analysis technique generated two electroencephalographic traces, one in which the original data were filtered to visualize the MMR, and the other in which the data were filtered to visualize the MMN.

The filtered data were then resampled at 200 samples/s and epoched from 100 to 600 ms relative to stimulus onset. Artifact trials were defined, for each electrode, as those epochs where activity exceeded $\pm$100 μV. Trials containing artifact were corrected with an artifact-blocking algorithm applied in MATLAB (Mourad *et al.*, 2007; Fujioka *et al.*, 2011). Corrected data from each subject were re-referenced to the average reference before being averaged across epochs, separately for each filtered trace (MMR and MMN). Baseline activity, defined as the mean amplitude in the 100 ms window prior to stimulus onset, was subtracted from each epoch during averaging. For statistical analysis, 72 electrodes were selected and divided into four groups per hemisphere (Fig. 2.2). The waveforms for all electrodes in each group were averaged

together to represent activity at the frontal (16 electrodes), central (20 electrodes), parietal (18 electrodes) and occipital (18 electrodes) scalp regions. Data representing each scalp region were also grand-averaged across subjects in each age group with respect to trial type. Finally, to visualize the MMR and MMN, difference waves for each deviant type (leftward or rightward) were constructed by subtracting the grand-averaged response to standard stimulation from the grand-averaged response to deviant stimulation.

For each subject, the MMR component in the 0.5-20 Hz bandpass data was characterized by measuring the area under the curve (AUC) of the difference wave in frontal and central scalp region waveforms from 100 to 500 ms after stimulus onset. The AUC was further divided into four 100-ms windows to examine the temporal morphology of the MMR. The MMN component in the 3-18 Hz bandpass data was characterized for each subject by measuring the amplitude and latency of the largest negative deflection in the difference wave in frontal and central scalp regions between 100 and 300 ms after stimulus onset. Peak amplitudes of the MMN and the AUC of the MMR in frontal, central and occipital electrode groups were also tested against zero with a two-tailed one-sample t-test. All measures were extracted in MATLAB. Repeated-measures ANOVAs were used to test whether the MMRs and/or MMNs varied across the between-subjects factor Age Group (2, 5, 8 and 13 months), and the within-subject factors Deviant Side (left/right), Electrode Hemisphere (left/right), Electrode Group (frontal/central), and, only for AUC measures, Time Bin (100-200, 200-300, 300-400 and 400-500 ms after stimulus onset). Mauchly's test of sphericity was used to evaluate the assumption of equal variance for the differences between

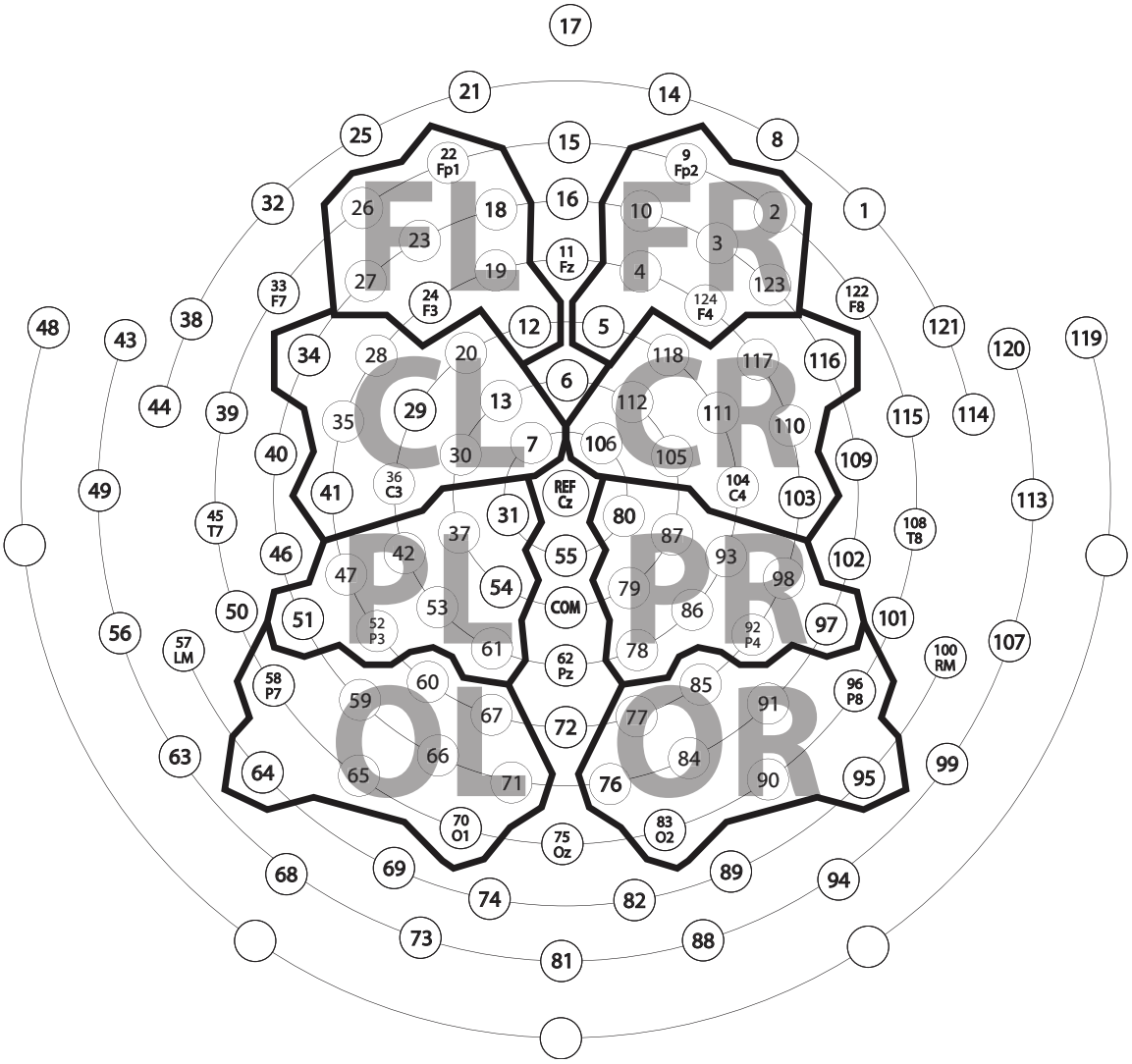

FIGURE 2.2: A schematic of the 129-channel HyrdoCel Geodesic sensor net. Bordered regions denote electrode groupings used to represent frontal (FL/FR), central (CL/CR), parietal (PL/PR) and occipital (OL/OR) scalp-recorded activity. Each grouping contains between 8 and 10 electrodes per hemisphere. The remaining 57 electrodes were excluded from the analysis to reduce artefacts and allow for comparison between hemispheres. COM: Isolated Common; REF: Reference.

all combinations of related groups (levels). Huynh-Feldt correction was applied to critical F-values for levels found to violate assumptions of sphericity.

2.3.2 Experiment 2

Participants

We recruited 15 healthy full-term 2-month-old infants (12 female; mean age = 2.61, SD = 0.19 months) with no known hearing deficits. Prior to the experiment, basic demographic information was obtained from each infant's caregiver. The final sample consisted of 13 infants aged 2 months (10 female; mean age = 2.62, SD = 0.21 months). Two infants were excluded from the analysis because of too few artifact-free trials. All infants included in the final sample were awake during the course of the experiment. Written consent was obtained from all parents of the infants in compliance with a protocol approved by the McMaster Research Ethics Board. The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* (18 July 1964).

Stimulus and procedure

Frozen white noise bursts were presented in an oddball paradigm identical to that detailed for Experiment 1, except that stimulus amplitude was pseudo-randomly varied around 76.5 dB SPL-C by ±4 dB in 1 dB increments. Amplitude assignment for each presentation also followed the restriction that no two subsequent presentations could be of equal amplitude, thereby ensuring that no sensory memory trace could be established for loudness cues associated with standard presentation from the center speaker. In this manner, the modified presentation paradigm controlled for the

possibility that the MMN and MMR seen in Experiment 1 were evoked solely by the changes in loudness cues that accompany ±90° shifts in azimuthal sound source origin, rather than the change in sound source location.

To address the added possibility that the caregivers were in some manner influencing the infants, all caregivers holding the infants wore sound-attenuating ear muffs during the experiment. Following the experimental session, researchers confirmed that caregivers could not identify the location of the sound source while wearing the ear muffs.

Data acquisition and analysis

The data were filtered and averaged with the same procedure outlined for Experiment 1. Peak amplitudes of the MMN and the AUC of the MMR in frontal, central and occipital electrode groups were tested against zero with a two-tailed one-sample t-test ($\alpha = 0.05$). To examine whether loudness cues contributed to the MMR and MMN recorded from 2-month-old infants, a two-tailed independent-samples t-test ($\alpha = 0.05$) was used to evaluate the null hypothesis of no difference in peak amplitude/latency (MMN) or AUC (MMR) between the groups of 2-month-olds tested in the roving and in the fixed-amplitude conditions. Repeated-measures ANOVAs, similar to those detailed for Experiment 1, were used to evaluate whether Deviant Side, Electrode Group or Electrode Hemisphere had any effect on MMN amplitude, MMN latency, and, with the inclusion of a Time Bin factor, MMR AUC.

2.4 Results

2.4.1 Experiment 1

The grand-averaged MMR (bandpass, 0.5-20 Hz) and MMN (bandpass, 3-18 Hz) of each age group are shown in FIGURES 2.3 and 2.4, respectively. Grand-averaged difference waves for each age group are shown for leftward and rightward deviants in FIGURE 2.5 (MMR) and FIGURE 2.6 (MMN).

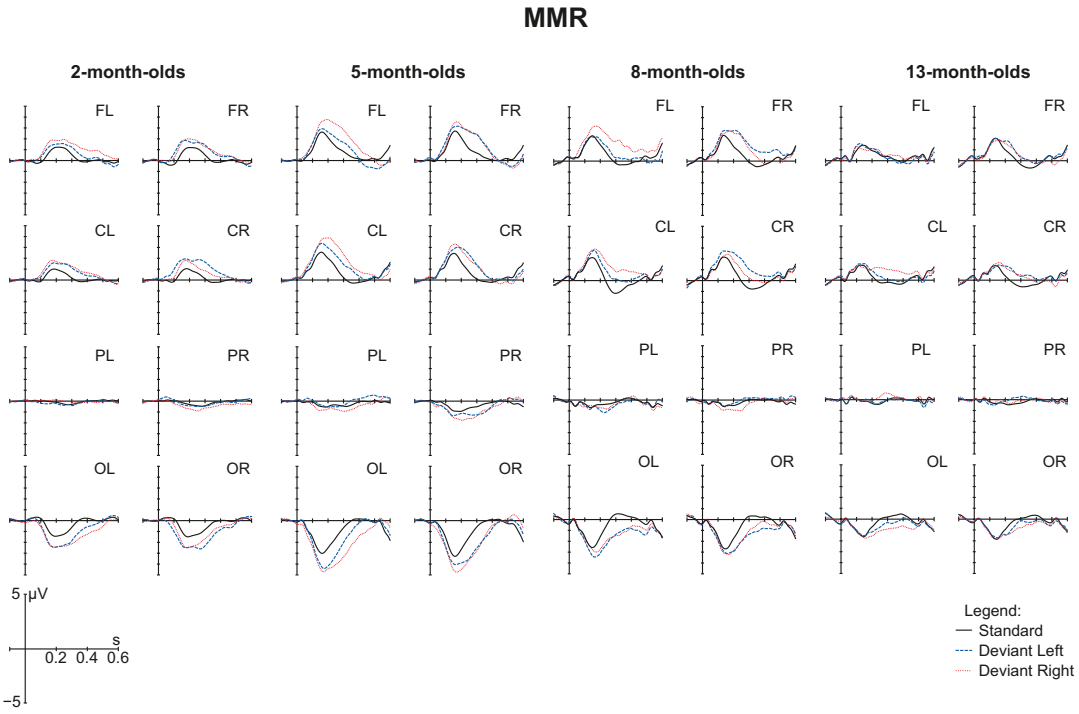

FIGURE 2.3: Grand-averaged standard (solid), deviant left (dashed) and deviant right (dotted) responses recorded from four different age groups (2-, 5-, 8-, and 13-month-olds). The data shown here were bandpass-filtered between 0.5 and 20 Hz. Responses represent averaged activity in eight scalp regions: frontal (FL/FR), central (CL/CR), parietal (PL/PR), and occipital (OL/OR), as grouped in FIG. 2.2.

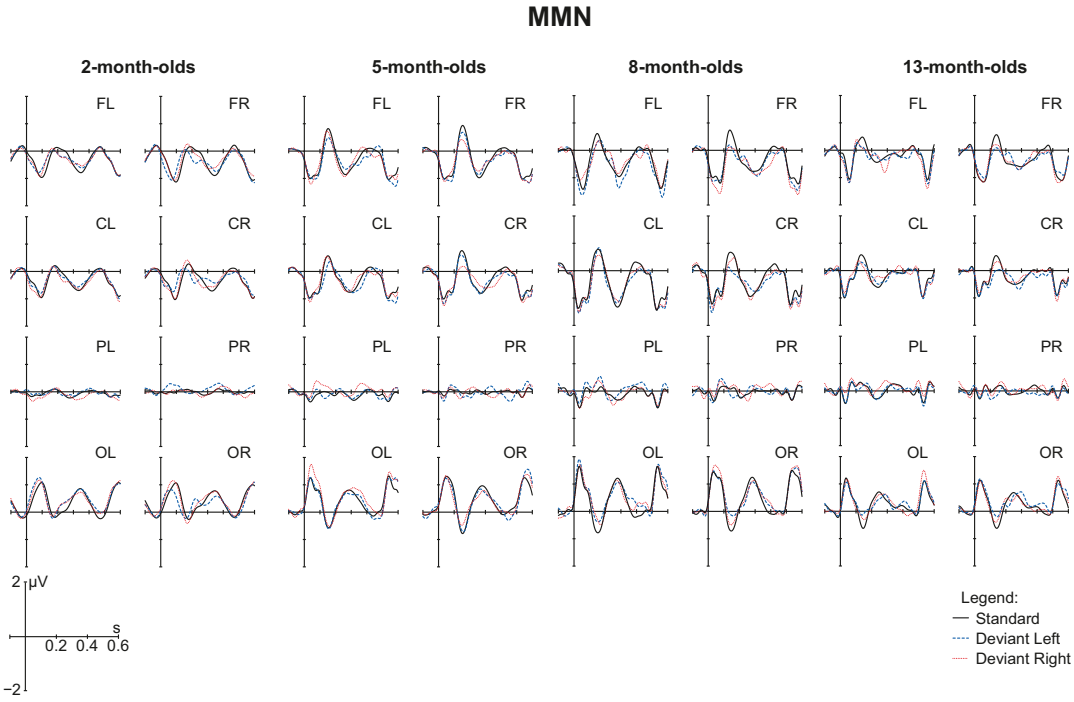

FIGURE 2.4: Grand-averaged standard (solid), deviant left (dashed) and deviant right (dotted) responses recorded from four different age groups (2-, 5-, 8-, and 13-month-olds). The data shown here were bandpass-filtered between 3 and 18 Hz. Responses represent averaged activity in eight scalp regions: frontal (FL/FR), central (CL/CR), parietal (PL/PR), and occipital (OL/OR), as grouped in FIG. 2.2.

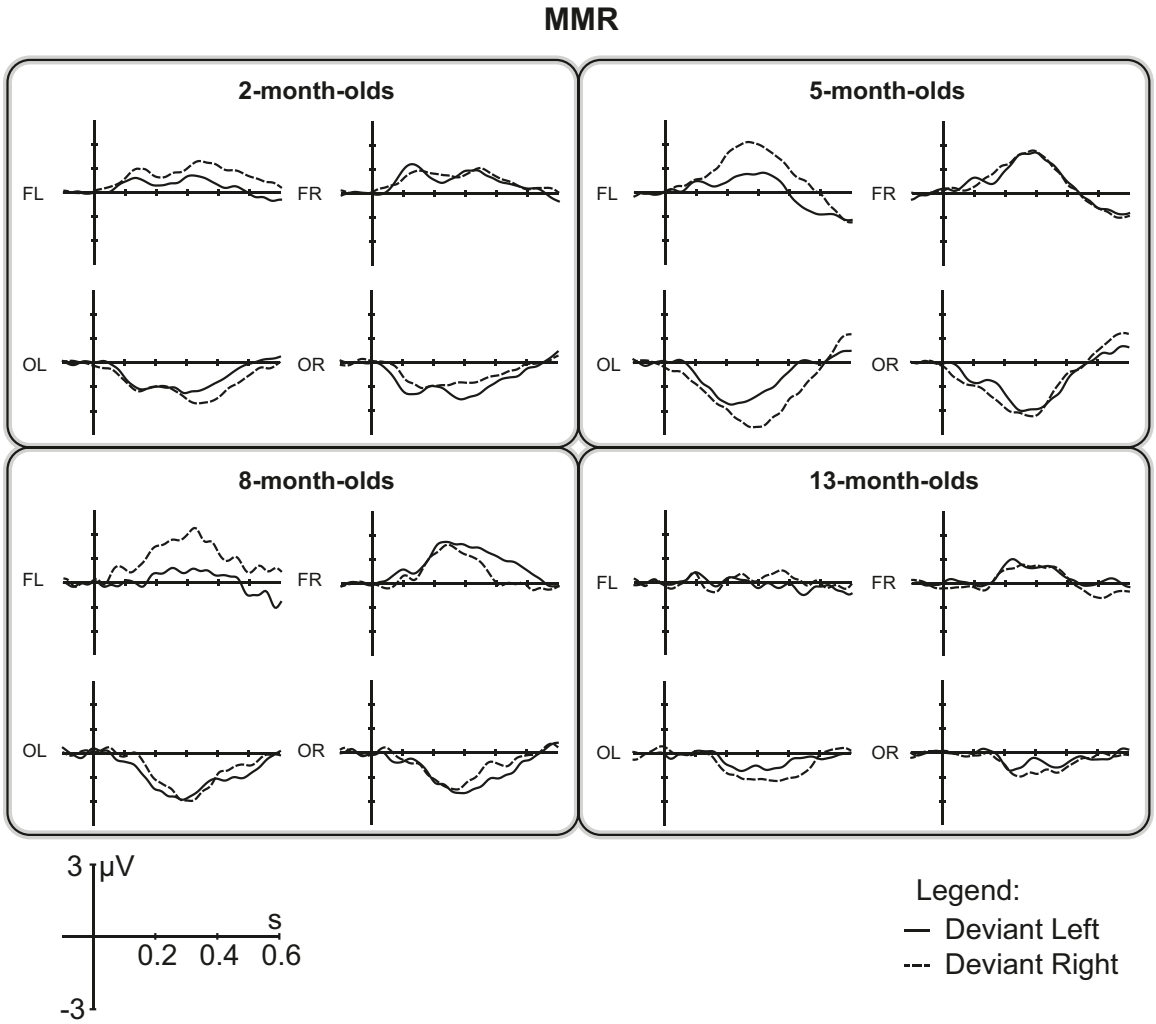

FIGURE 2.5: Grand-averaged difference waves (deviant minus standard; bandpass, 0.5–20 Hz) in frontal (FL/FR) and occipital (OL/OR) scalp regions as elicited in 2-, 5-, 8-, and 13-month-olds. Responses to leftward and rightward deviants are shown as solid and dashed traces, respectively.

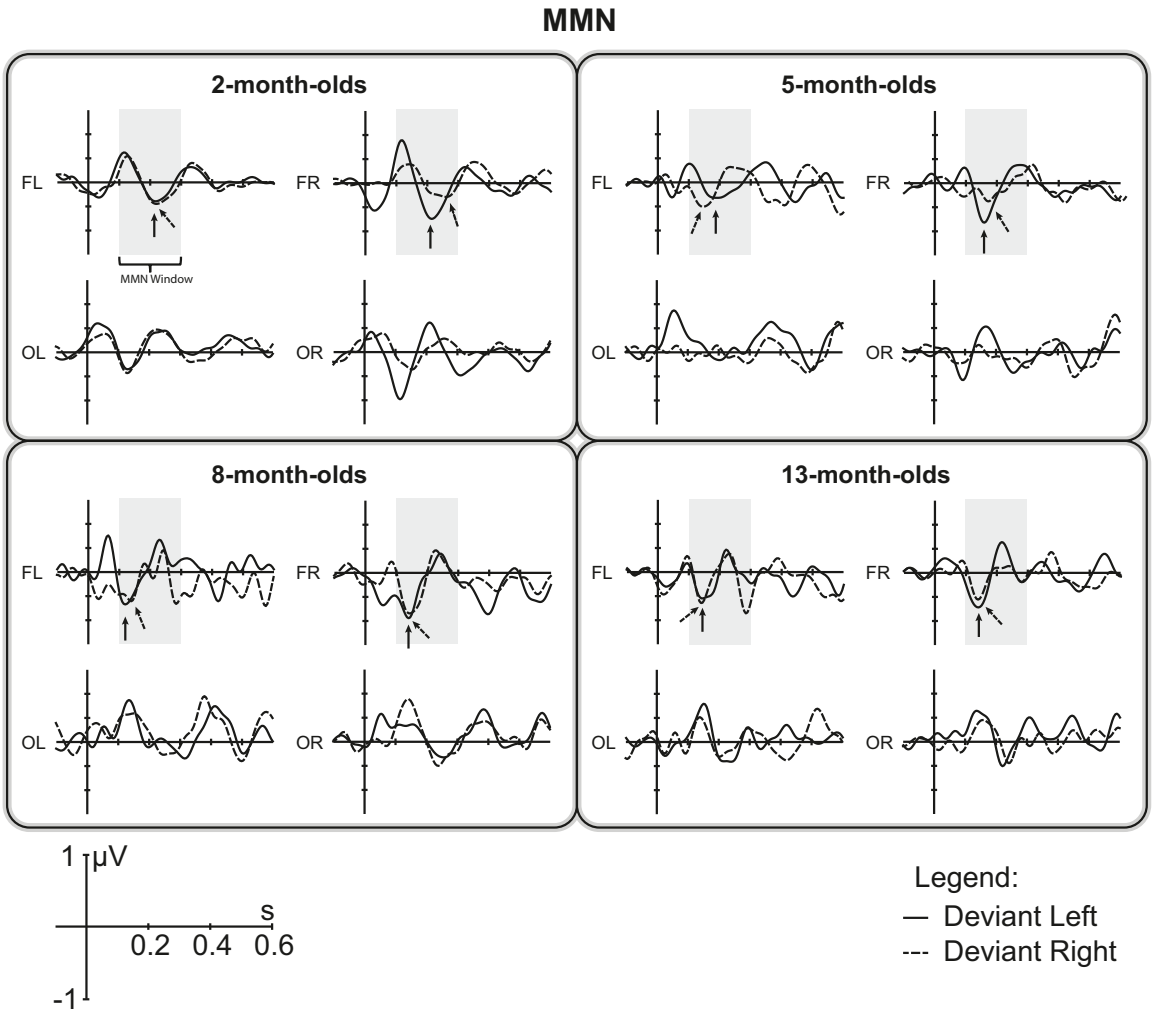

FIGURE 2.6: Grand-averaged difference waves (deviant minus standard; bandpass, 3–18 Hz) in frontal (FL/FR) and occipital (OL/OR) scalp regions as elicited in 2-, 5-, 8-, and 13-month-olds. Responses to leftward and rightward deviants are shown as solid and dashed traces, respectively. Gray windows denote the time after stimulus onset over which the MMN peak was extracted. Solid and dashed arrows denote grand-averaged MMN peaks in response to leftward and rightward deviants, respectively.

MMRs

The slow-wave event-related potentials (ERPs) were dominated by broad positivity in the frontal (FL/FR) and central (CL/CR) electrode sites that spanned between 200 and 400 ms after stimulus onset. The fronto-central positivity was accompanied by reciprocal negativity in the occipital (OL/OR) regions (Fig. 2.3). In each age group, except for 13 months, deviant trials elicited significantly more positive responses in frontal electrode groups than did standard trials. In all age groups, deviant trials elicited significantly more positive responses at central electrode groups, and more negative responses in occipital electrode groups, than did standard trials (Table 2.1; Fig. 2.5).

Table 2.1: Average AUC (μV*ms) of the MMR recorded from 2-, 5-, 8-, and 13-month-old infants in response to changes in sound source location. Data are also shown for a 2-month-old group tested in Experiment 2 using a roving condition (2-roving) where stimulus amplitude was randomly varied by ±4 dB on each trail.

	Electrode Group					
Age(months)	**Frontal**		**Central**		**Occipital**	
2	13.63	$\pm 10.81^{***}$	14.28	$\pm 6.40^{***}$	-19.43	$\pm 11.00^{***}$
2-roving	12.56	$\pm 11.78^{**}$	11.87	$\pm 6.81^{***}$	-13.99	$\pm 11.22^{**}$
5	15.67	$\pm 15.49^{**}$	17.56	$\pm 9.86^{***}$	-25.21	$\pm 13.61^{***}$
8	19.02	$\pm 12.91^{**}$	17.51	$\pm 8.12^{***}$	-20.61	$\pm 7.91^{***}$
13	5.78	± 14.17	10.15	$\pm 6.79^{***}$	-9.93	$\pm 7.11^{***}$

$p < 0.01^{**}$ $p < 0.001^{***}$

The morphology and scalp topography of the deviant response was consistent with previous reports of infant ERPs elicited by other deviant classes, e.g., pure tones (Leppänen *et al.*, 1997), harmonic tones (Fellman *et al.*, 2004), piano tones (He *et al.*,

2007, 2009), and speech syllables (Friederici *et al.*, 2002; Friedrich *et al.*, 2004). The polarity inversion along the anteriorposterior axis further suggests bilateral generators around the temporal lobes, probably reflecting activity within the auditory cortices (Dehaene-Lambertz and Dehaene, 1994).

The omnibus ANOVA revealed a significant effect of Age Group on the AUC measured from the MMR ($F_{3,48} = 3.28$, $p < 0.05$) (Fig. 2.7A). Subsequent pairwise comparisons (Tukey HSD) found the AUC of the 13-month-old group to be significantly smaller than that of the 8-month-old group. No other age differences were significant. The AUC measures were also significantly affected by Time Bin ($F_{2.74,131.67} = 55.78$, $p < 0.001$), and an interaction between Time Bin and Age Group ($F_{8.23,131.67} = 4.82$, $p < 0.001$) (Fig. 2.7B). In the 5-, 8- and 13-month-old groups, the MMR appeared to be most positive between 200 and 400 ms after stimulus onset, with a peak at 200-300 ms, whereas the MMR of the 2-month-old group appeared to be more spread out over time. These data suggest that the neural processes responsible for generating the slow-positive MMR in response to acoustic spatial deviants undergo a transition between 8 and 13 months of age that is accompanied by diminution and/or suppression of the MMR.

The ANOVA revealed several additional effects that probably reflected the locations of the MMR generators. First, there was a significant effect of a Deviant Side by Electrode Hemisphere interaction ($F_{1,48} = 18.06$, $p < 0.001$) (Fig. 2.7C). Deviant trials elicited larger responses in the electrode hemisphere contralateral to the deviant sound's spatial position. Enhancement of the spatial MMN over the contralateral hemisphere has been observed in adult subjects (Kaiser and Lutzenberger, 2001; Sonnadara *et al.*, 2006), and is thought to reflect the contralateral predominance

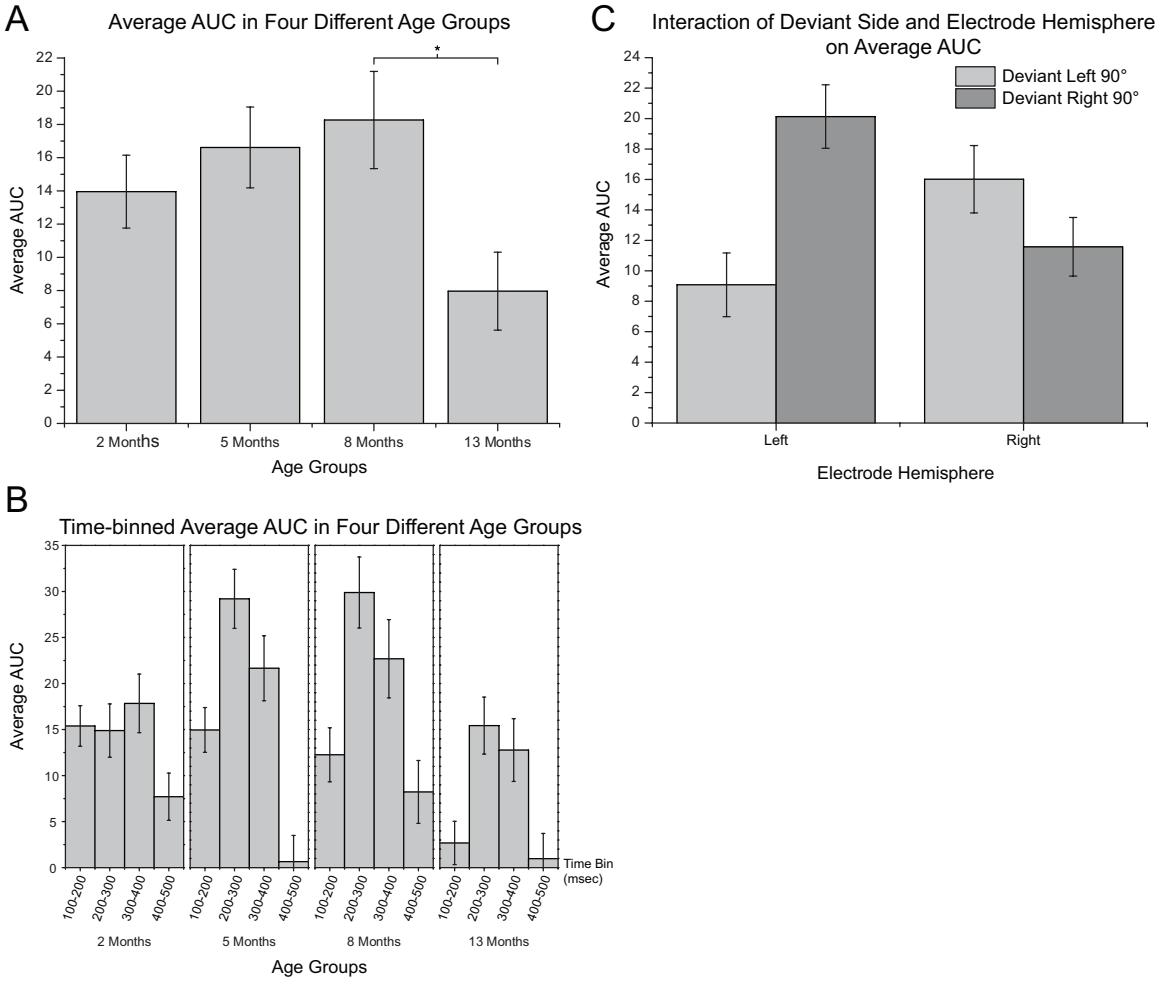

FIGURE 2.7: (A) Average AUC, from 100 to 500 ms after stimulus onset, in each age group measured from the difference waves in the frontal (FL/FR) and central (CL/CR) scalp regions. Pairwise comparisons (Tukey HSD) revealed a significant difference between the AUCs of the 8- and 13-month-old groups ($p < 0.05$). No other significant differences were found. (B) Significant effect of an interaction between Age Group and Time Bin on average AUC measures taken from the slow-wave MMR data (bandpass, 0.5-20 Hz) in four different age groups ($p < 0.001$, Huynh-Feldt correction). (C) Significant interaction effect between electrode hemisphere and Deviant Side on the average AUC measured from 100 to 500 ms across all age groups ($p < 0.001$). In all figures, error bars represent standard errors.

of cortical neurons in the auditory pathway (Phillips and Irvine, 1983). Second, the ANOVA revealed significant effects of a Time Bin by Electrode Hemisphere interaction ($F_{2.95,141.71} = 4.66$, $p < 0.01$) and a three-way interaction between Time Bin, Electrode Hemisphere, and Deviant Side ($F_{2.17,191.96} = 6.40$, $p < 0.01$). From 100 to 300 ms after stimulus onset, MMRs were larger over right hemisphere electrodes than over left hemisphere electrodes. This pattern was reversed from 300 to 500 ms after stimulus onset. When examined separately for each deviant stimulus, the AUC was larger over the hemisphere contralateral to the deviant source in every time bin. Third, there were also effects of a four-way interaction between Deviant Side, Time Bin, Electrode Group, and Electrode Hemisphere ($F_{2.33,112.04} = 8.49$, $p < 0.001$), and a five-way interaction between Deviant Side, Time Bin, Electrode Group, Electrode Hemisphere, and Age ($F_{7.00,112.04} = 2.26$, $p < 0.05$). The AUC measured from central electrodes was always larger than from frontal electrodes; however, the difference in AUC measured from the left and right hemisphere electrodes varied according to the location of the deviant source, time bin, electrode group, and age. These data indicate that the generators of the MMR signal do not respond identically to leftward and rightward deviants, and that this asymmetry is affected by development.

MMNs

Bandpass filtering of the data between 3 and 18 Hz was effective in removing the slow-wave MMR (Fig. 2.4). In all age groups, deviant trials elicited significantly more negative responses in the frontal and central electrode groups, and significantly more positive responses in the occipital electrode groups, than did standard trials (Table 2.2). Difference waves (Fig. 2.6) showed a negative deflection peaking

TABLE 2.2: Average peak amplitude (μV) of the MMN recorded from 2-, 5-, 8-, and 13-month-old infants in response to changes in sound source location. Data are also shown for a 2-month-old group tested in Experiment 2 using a roving condition (2-roving) where stimulus amplitude was randomly varied by ±4 dB on each trail.

	Electrode Group					
Age(months)	**Frontal**		**Central**		**Occipital**	
2	-0.58	$\pm 0.26^{***}$	-0.53	$\pm 0.34^{***}$	0.61	$\pm 0.31^{***}$
2-roving	-0.56	$\pm 0.28^{***}$	-0.56	$\pm 0.23^{***}$	0.67	$\pm 0.20^{***}$
5	-0.85	$\pm 0.33^{***}$	-0.78	$\pm 0.23^{***}$	0.77	$\pm 0.23^{***}$
8	-0.88	$\pm 0.28^{***}$	-0.84	$\pm 0.22^{***}$	0.95	$\pm 0.25^{***}$
13	-0.86	$\pm 0.28^{***}$	-0.71	$\pm 0.28^{***}$	0.85	$\pm 0.25^{***}$

$p < 0.001^{***}$

between 100 and 300 ms after stimulus onset accompanied by reciprocal positivity in the occipital electrode groups. Again, this topography is suggestive of generators around the auditory cortices.

Analysis of the MMN-like response revealed a significant effect of Age Group on the peak amplitude of the negative component ($F_{3,48} = 4.49$, $p < 0.01$) (FIG. 2.8A). Pairwise comparisons (Tukey HSD) indicated that the effect of Age Group was driven primarily by smaller MMN amplitude in the 2-month-old group than in all other age groups. The omnibus ANOVA also revealed a significant effect of Electrode Group ($F_{1,48} = 8.28$, $p < 0.01$) on MMN amplitude, with peak amplitude larger in frontal than in central electrode groupings. As in the AUC data, there was also a significant effect of a Deviant Side by Electrode Hemisphere interaction on MMN amplitude ($F_{1,48} = 12.37$, $p < 0.01$) (FIG. 2.8C).

Peak latency was significantly affected by Age Group ($F_{3,48} = 8.80$, $p < 0.001$) (FIG. 2.8B), with a longer latency in the 2-month-old group than in all other age

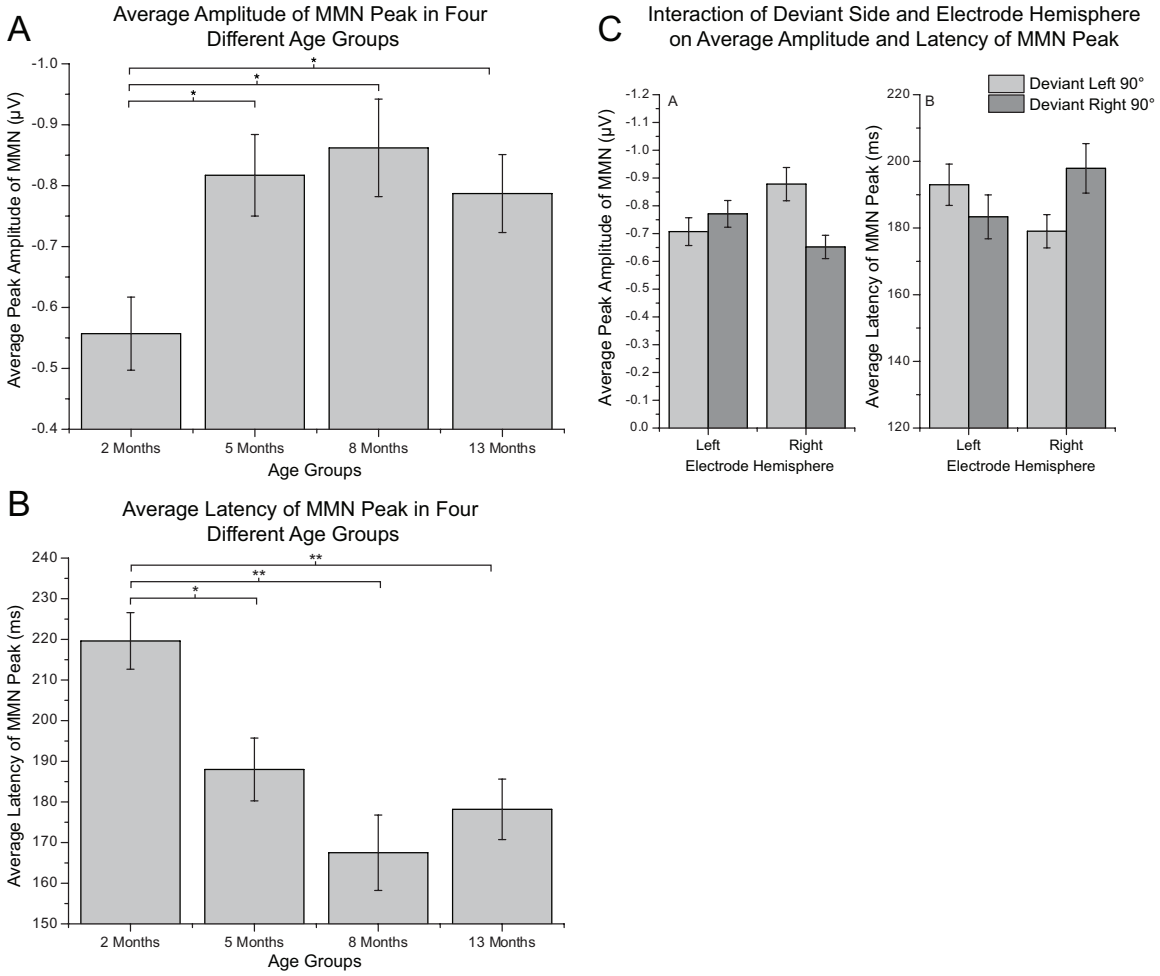

FIGURE 2.8: (A) Average peak MMN amplitude (100-300 ms after stimulus onset) and (B) latency measured in the frontal (FL/FR) and central (CL/CR) scalp regions in each age group. Pairwise comparisons (Tukey HSD) revealed a significant difference between the peak MMN amplitude and latency of the 2-month-old group and those of all other age groups ($*p < 0.05$; $**p < 0.01$). No other significant differences were found. (C) Significant effect of an interaction between electrode hemisphere (x-axis) and Deviant Side (light gray, deviant left; dark gray, deviant right) on the average amplitude (left side) and latency (right side) of the MMN across all age groups ($p < 0.01$). In all figures, error bars represent standard errors.

groups (pairwise comparisons, Tukey HSD). There were also significant effects of a Deviant Side by Electrode Hemisphere interaction ($F_{1,48} = 7.67$, $p < 0.01$) (Fig. 2.8C) and a three-way interaction between Deviant Side, Electrode Group and Electrode Hemisphere ($F_{1,48} = 12.39$, $p < 0.01$) on the peak latency of the MMN. Leftward deviants elicited shorter latency MMN-like negativities in frontal electrode groups, over both hemispheres, than rightward deviants. Conversely, a contralateral bias was observed in the latency measured from central electrode groups, where deviants elicited shorter latency MMN-like responses over the contralateral hemisphere.

2.4.2 Experiment 2

Grand-averaged difference waves for leftward and rightward deviants are shown in Figure 2.9A (MMR) and Figure 2.9B (MMN).

MMRs

The slow-wave ERPs showed a morphology consistent with the infant MMR recorded in Experiment 1. The AUCs of responses elicited by deviant stimuli were significantly more positive at frontal and central electrode sites and significantly more negative at occipital electrode sites than those elicited by standard stimuli (Table 2.1). An independent-samples t-test found no significant difference in the AUCs of the difference waves in any electrode group between data collected in Experiment 2 and those collected in Experiment 1 (Fig. 2.10).

The repeated-measures ANOVA revealed a significant effect of Time Bin ($F_{2.292,27.504} = 7.483$, $p < 0.01$) and an interaction between Time Bin and Electrode Group ($F_{1.982,23.783} = 4.116$, $p < 0.01$) on the AUC of the MMR. Overall, the measured

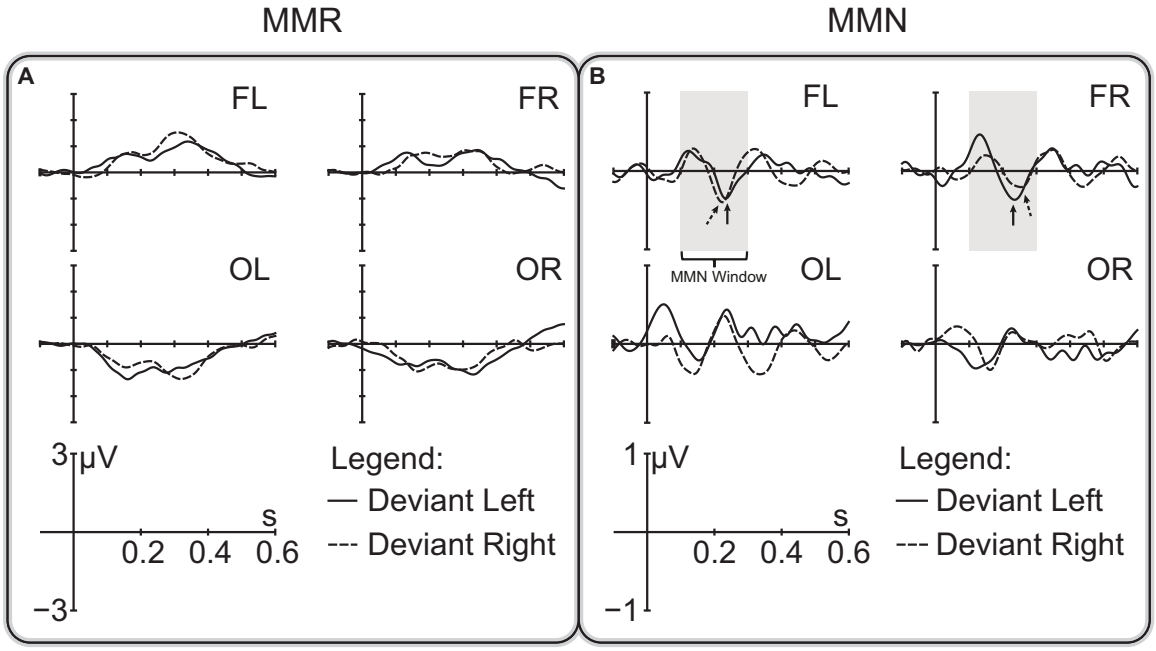

FIGURE 2.9: Grand-averaged difference waves (deviant - standard) in frontal (FL/FR) and occipital (OL/OR) scalp regions elicited in 2-month-old infants in Experiment 2. Responses to leftward and rightward deviants are shown as solid and dashed traces, respectively. The MMR dominates the difference waves when the data are filtered between 0.5 and 20 Hz (A). Filtering the data between 3 and 18 Hz reveals an MMN-like response (B). Gray windows denote the time after stimulus onset over which the MMN peak was extracted. Solid and dashed arrows denote grand-averaged MMN peaks in response to leftward and rightward deviants, respectively.

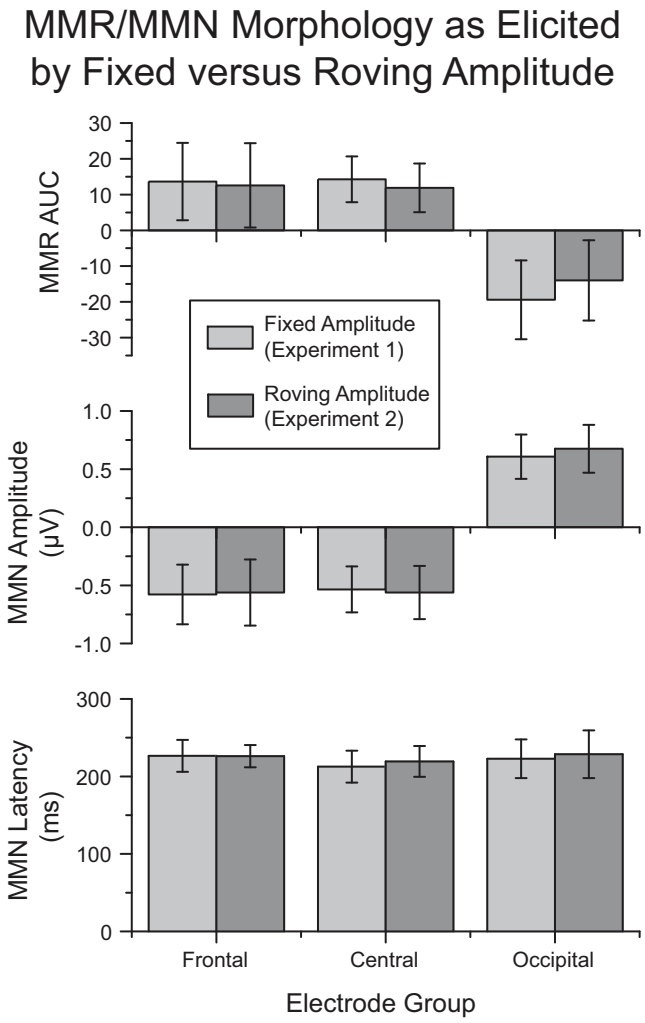

FIGURE 2.10: Comparison of MMR and MMN morphology as elicited in two different groups of 2-month-old infants. Light gray bars represent data collected in Experiment 1, in which the amplitude of deviants and standards presentations was fixed. Dark gray bars represent data collected in Experiment 2, in which the amplitude of deviants and standards trials was pseudo-randomly roved by ±4 dB. Independent-samples t-tests (two-tailed, $\alpha = 0.05$) revealed no significant differences between the fixed and roving amplitude experiments in the AUC of the MMR (top) or the peak amplitude (middle) and latency (bottom) of the MMN, as measured in frontal, central or occipital electrode groups.

AUC was largest between 100 and 400 ms after stimulus onset, and was greatly diminished between 400 and 500 ms. The distribution of MMR signal strength across the electrode groups also varied over time. Between 100 and 200 ms after stimulus onset, the MMR signal was stronger at central electrode sites than at frontal electrode sites, after which this pattern was reversed.

MMNs

Filtering the data between 3 and 18 Hz revealed an MMN-like response, peaking at 230 ms after stimulus onset, in the difference wave that was obtained by subtracting the averaged standard response from the averaged deviant response. One-sample t-tests conducted on the peak amplitude of the MMN confirmed that deviant responses were significantly more negative than standard responses at frontal and central electrode sites and significantly more positive at occipital electrode sites (Table 2.2). The repeated-measures ANOVA revealed no significant effects of Deviant Side, Electrode Group, or Hemisphere, or an interaction between them, on either the amplitude or latency of the MMN. As for the MMR, an independent-samples t-test failed to show any significant difference in the peak MMN amplitude or latency between data collected in Experiment 2 and those collected in Experiment 1 (Fig. 2.10).

2.5 Discussion

2.5.1 Experiment 1

The results provide evidence that the auditory system of infants is sensitive sound location changes, in that pre-attentive mismatch responses were evident in both the

MMR and MMN in all age groups. However, changes in sound location produce interaural changes in loudness (Blauert, 1983). Specifically, with respect to one ear, sounds in contralateral space will be slightly less intense than sounds in ipsilateral space. To rule out the possibility that the MMR and MMN responses that we observed here were attributable to monaural sound level differences, Experiment 2 repeated Experiment 1 with the youngest age group using a roving stimulus level.

2.5.2 Experiment 2

The results of Experiment 2 strongly suggest that the change-detection responses found in Experiment 1 were not elicited by monaural loudness cues, but rather primarily reflected responses to changes in sound location.

2.5.3 General

The pre-attentive mismatch response was present as both a slow-positive (MMR) and a fast-negative (MMN) component in 2-, 5-, 8-, and 13-month-olds. The effects of Deviant Side by Electrode Hemisphere interactions observed on the AUC measures of the MMR and peak amplitude/latency measures of the MMN are consistent with previous studies showing the adult MMN to be larger and to occur earlier in electrodes that are contralateral to the deviant sound source (e.g., Kaiser *et al.* 2000a,b; Kaiser and Lutzenberger 2001; Nager *et al.* 2003; Sonnadara *et al.* 2006; Richter *et al.* 2009). Whole-cell recordings (*in vivo*) of pyramidal cells in superficial layers (II–IV) of the rat auditory cortex show a similar bias to contralateral sound presentation in free-field – faster-rising excitatory postsynaptic potentials, shorter spike latency, and greater likelihood of spike occurrence (Chadderton *et al.*, 2009). Along with the reversal

in response polarity across frontal/central and occipital electrode groups (Figs 2.5 and 2.6), the observed contralateral response bias (Figs 2.7C and 2.8C) strongly implies generators situated in the auditory cortices. Moreover, the MMRs and MMNs recorded from 2-month-olds in Experiment 2 did not differ significantly from those recorded in Experiment 1, which rules out the possibility that these ERPs reflect primarily neural activity elicited by a change in stimulus loudness at each ear rather than the change in sound source spatial location. Thus, despite the well-documented behavioural insensitivity to sound location observed between 1 and 3 months of age (e.g., Muir *et al.* 1979, 1989; Clifton *et al.* 1981; Muir and Hains 2004), our data indicate that infants in all age groups respond to changes in sound source location at a cortical level.

According to Muir and colleagues (Muir and Clifton, 1985; Muir and Hains, 2004), the subcortically generated reflex to off-midline sounds should undergo suppression between 1 and 4 months of age as the auditory system transitions to a cortical localization processor. We can look at this hypothesis by examining the MMR, because it reflects cortical activity directly driven by subcortical input. Positive obligatory evoked potentials, including the positive MMR, probably reflect a depolarization in layer IV of the auditory cortex (Eggermont and Moore, 2012). Layer IV receives activity from the ventral medial geniculate nucleus via thalamocortical projections that do not undergo substantial myelination until 3 months of age. If suppression of a subcortically mediated localization process is taking place, our findings indicate that it does not affect input to the auditory cortex in a way that would significantly perturb the average amplitude of the MMR from 2 to 8 months of age. Although the AUC measured in 2-month-olds was more spread out in time than that of the 5-

and 8-month-olds, this difference is not likely to be a reflection of a suppressed MMR generator(s), but rather reflects poor temporal synchrony in the population of cortical neurons stimulated by largely unmyelinated thalamocortical afferents (Eggermont and Moore, 2012). Furthermore, significant diminution of the MMR is not seen until between 8 and 13 months of age, which is at least 4 months later than the onset of the mature localization behaviour thought to reflect engagement of cortical auditory spatial processing.

Interestingly, the lack of MMR amplitude diminution before 8-13 months of age highlights a protracted developmental trajectory as compared with that described for other acoustic deviants (Trainor *et al.*, 2003) [pitch deviants are reviewed in He *et al.* (2007)]. It is difficult to ascertain whether the sustained presence of a slow-positive MMR is attributable to the nature of the experimental stimulation or to delayed development in auditory pathways conveying sound location information to the cortex. Kushnerenko *et al.* (2002) have suggested that the slow-positive component reflects a neural correlate of the infant orienting response, similarly to how the adult P3a response is thought to index an involuntary orienting of attention (Escera *et al.*, 2000). In this case, spatial deviants might have acted as maximally attention-grabbing stimuli, thus eliciting MMR in all age groups where neurophysiology still allowed for propagation of a slow-positive component from the cortex.

After the reported onset of thalamocortical myelination, a shorter-latency negative component is observed to precede the longer-latency positive MMR. The negativity is probably generated in layer I, which begins to receive excitatory projections from the medial division of the medial geniculate nucleus at $\sim$4 months of age. According to Eggermont and Moore (2012), the persistence of a negative mismatch after 8 months

of age requires a generator in layer I and/or layer II, and excitatory modulation from pyramidal cells with bodies in layers II, III, and V. In this neurophysiological framework, if the auditory cortex is at all sensitive to subcortically extracted spatial information, lemniscal (thalamocortical) projections should broadly activate layer IV and produce the dominant slow-positive MMR seen in all age groups. After sufficient thalamocortical myelination, cortical sensitivity to spatial information might additionally modulate non-lemniscal activity from afferents projecting from the medial division of the medial geniculate to layer I of the auditory cortex; however, this modulation does not necessarily preclude thalamocortical stimulation of layer IV from continuing to produce broad positive MMRs later in infancy, especially if the evoking stimulus is sufficiently salient to induce strong activation through the auditory pathway. In the case of our experiment, 90° spatial deviants in the azimuth represent maximally discrepant stimuli relative to the midline standards, and so might have continued to promote broad activation of layer IV in age groups in which positive MMR is rarely observed in response to other acoustic deviants. Of course, these proposed mechanisms are highly speculative, and the myriad of sources and sinks in various cortical layers that overlap in time and that give rise to surface-measured potentials is extremely complex (e.g., Steinschneider *et al.* 2008). Nonetheless, it is useful to consider how differential patterns of MMR/MMN maturation, as elicited by different acoustic deviants, might inform us with respect to the general and/or specific development of early acoustic discriminative capabilities. This is especially relevant to the current experiment, in which the developmental course of the MMR is atypical as compared with that observed for other acoustic deviants, and therefore suggests a

degree of within-domain specificity in the maturation of mismatch processes and/or generators.

As predicted, we observed significant changes in the MMN between the 2- and 5-month-old groups that manifested as a larger average peak amplitude and shorter average latency. Enhancement of the MMN coincides with the reported onset of mature, probably cortically mediated, sound localization abilities; however, MMN emergence is consistent with other infant studies of pitch discrimination (He *et al.*, 2007), gap detection (Trainor *et al.*, 2003), and resolution of the missing fundamental (He *et al.*, 2009). It is thus difficult to ascertain whether the age-related changes in the MMN elicited by our spatial deviants reflect something unique to the development of sound localization abilities or rather the general maturation of the auditory cortex. Follow-up studies could reveal the extent to which the adult-like MMN reflects behavioural sensitivity to spatial deviants presented within a hemifield. Insofar as such discriminations are thought to require an intact/functional auditory cortex, evidence that infant MMN indexes these changes after, but not before, 4 months of age would further imply that the MMN enhancement observed between 2 and 5 months of age is tied to the onset of cortically mediated localization processes.

After 2 months of age (i.e., in the 5-, 8- and 13-month-old groups), the amplitude and latency of the MMN-like fast negativity remain relatively stable. Behavioural studies indicate substantial improvements in minimum audible angle (MAA) discrimination between 5 and 13 months of age (Litovsky and Ashmead, 1997), with the MAA dropping from 19.8° at 5 months to 8.0° at 12 months. We did not record corresponding enhancements in MMN amplitude or latency over the same age range; however, assessment of the MAA was not performed in this study, and the MMN

elicited by spatial deviants in free-field is known to saturate when the discrimination is greater than the MAA at the standard location (Paavilainen *et al.*, 1989). The large ±90° shift used in the present study likely far exceeded the MAA in the 5-, 8- and 13-month-old age groups, probably causing saturation of the MMN generators. Subsequent studies might use smaller angular differences in azimuthal location to reveal potential differences in the MMN across age.

Given that the central auditory system of 2-month-olds is responding automatically to changes in sound source location, it is still unclear why reliable orienting behaviour is absent between postnatal months 1 and 3. Furthermore, there is good evidence to suggest that neonatal orienting to sound does not reflect the same sound localization mechanism observed in adults or older children (Litovsky and Ashmead, 1997; Muir and Hains, 2004; Litovsky, 2012). Neonatal localization abilities appear to be largely limited to left/right hemifield discriminations (Muir *et al.*, 1989), and require longer-duration transients than those necessary for localization in older infants and adults (Clarkson *et al.*, 1989). Neonates also lack sensitivity to the precedence effect–an auditory spatial illusion thought to be part of an echo suppression mechanism requiring neurons in the primary auditory cortex (Cranford *et al.*, 1971; Fitzpatrick *et al.*, 1999; Muir and Hains, 2004; Mickey, 2005). Sensitivity to precedence effect stimuli is not observed until between 4 and 5 months of age (Clifton *et al.*, 1984; Muir *et al.*, 1989), a window that tightly coincides with the reinstatement of orienting behaviour (Clifton *et al.*, 1981). It is unlikely that the period of behavioural silence reflects heighted sensitivity to visual interference or habituation, as infants of this age continually fail to orient towards a sound source when tested in complete darkness (Muir *et al.*, 1979) or in response to a series of novel auditory

tokens (Muir, 1985). Moreover, infants never lose the ability to turn their heads, and also continue to show robust orienting towards visual targets in both the temporal and nasal visual fields (Johnson, 1990). Therefore, the absence of an auditory orienting response cannot simply be attributed to physical limitations in the muscles of the neck or immaturities in the developing motor system.

One possibility concerns the sensitivity of head-turn or gaze-shift measures in assessing the auditory spatial abilities of infants. In a single-interval yes–no task, Morrongiello *et al.* (1990) asked observers to judge whether an infant showed any behavioural response (e.g., alerting, quieting, head turns towards or away from the sound source, or changes in pacifier-sucking behaviour) to sound presentations that either did or did not include a shift in sound location (see the observer-based psychoacoustic procedure in: Olsho *et al.*, 1987). Under these test conditions, infants between 2 and 5 months of age show above-chance responses to sounds that shift laterally relative to a control sound source located at the midline. Moreover, the change in the MAA required to elicit some form of response decreased linearly from 2 to 5 months, and is accompanied by a complementary linear increase in the proportion of "yes" responses coinciding with head turns towards the correct direction of the sound shift (Morrongiello *et al.*, 1990).

Another possibility is that, at this age, the developing motor system cannot make use of auditory spatial information in the same way that it can use visual information. Expanding on this idea, Muir and Hains (2004) propose that the neonatal auditory orienting response is suppressed to facilitate the integration of audio and visual spatial fields. As the visual field is very narrow at birth (Lewis and Maurer, 1992), the infant orienting response to sound sources might be delayed until the visual and auditory

spatial fields become aligned at approximately 3-4 months of age. This age is coincident with the onset of reaching behaviour in response to auditory targets presented in the dark (Clifton *et al.*, 1993). Behavioural sensitivity to the illusory percept known as the McGurk effect (requiring fusion of auditory speech stimuli and facial cues) also emerges after 4 months of age (Burnham and Dodd, 2004). Furthermore, adult data show a multisensory advantage in reaction time during localization of bimodal (auditory and visual) as compared with unimodal (auditory or visual) targets (Hughes *et al.*, 1994). Neil *et al.* (2006) found that a similar processing advantage is only observed after 8 months of age; however, it is unknown whether multisensory cues regarding spatial location elicit larger and/or earlier mismatch components in early infancy.

There is still much to be understood about how auditory space is represented in early infancy. Developing such an understanding is particularly important given how sound localization processes are known to direct the infant's sensory and attentional systems towards potentially meaningful sound-producing objects in the environment (King, 2009), to integrate spatially congruent audio-visual cues (Lewkowicz, 2002; Neil *et al.*, 2006), and to facilitate auditory stream segregation (Hawley *et al.*, 2004; Hollich *et al.*, 2005), stimulus detection (Nozza, 1987; Nozza *et al.*, 1988), speech intelligibility in noise (Hirsh, 1950; Litovsky *et al.*, 2006), and language acquisition (Lieu J, 2004).

2.6 Conclusion

In this study, we tested the cortical responsiveness of 2-, 5-, 8- and 13-month-old infants to large spatial changes in the azimuthal position of a sound source by using the

mismatch response. Despite the reported absence of overt behavioural orienting, it appears that the auditory systems of 2-month-old infants detect large ($\pm 90^{\circ}$ relative to midline) changes in sound location. Moreover, the changes in MMR/MMN morphology for auditory spatial deviants generally follow the trajectory found for other auditory abilities, such as pitch discrimination and gap detection, over the first 13 postnatal months. One exception is that, relative to these other auditory discriminations, there is a protracted diminution of the slow-positive MMR. This may reflect enhanced salience of spatial deviants in all age groups. Future work should investigate whether and/or which of these mismatch components accurately reflects behavioural MAAs at different stages in early development and the early developmental trajectory for audio-visual spatial integration.

2.7 References

Altman, J. A., Vaitulevich, S. P., Shestopalova, L. B., and Varfolomeev, A. L. (2005). Mismatch negativity evoked by stationary and moving auditory images of different azimuthal positions. *Neuroscience Letters*, **384**(3), 330–335.

Altman, J. A., Vaitulevich, S. P., Petropavlovskaia, E. A., and Shestopalova, L. B. (2010). Discrimination of the dynamic properties of sound source spatial location in humans: electrophysiology and psychophysics. *Human Physiology*, **36**(1), 72–79.

Altmann, C. F., Wilczek, E., and Kaiser, J. (2009). Processing of auditory location changes after horizontal head rotation. *The Journal of Neuroscience*, **29**(41), 13074–13078.

Bennemann, J., Freigang, C., Schröger, E., Rbsamen, R., and Richter, N. (2013). Resolution of lateral acoustic space assessed by electroencephalography and psychoacoustics. *Frontiers in Psychology*, **4**, 1–12.

Blauert, J. (1983). *Spatial Hearing: The Psychophysics of Human Sound localization*. MIT Press, Cambridge, MA.

Burnham, D. and Dodd, B. (2004). Auditory-visual speech integration by prelinguistic infants: perception of an emergent consonant in the McGurk effect. *Developmental Psychobiology*, **45**(4), 204–220.

Chadderton, P., Agapiou, J. P., McAlpine, D., and Margrie, T. W. (2009). The synaptic representation of sound source location in auditory cortex. *The Journal of Neuroscience*, **29**(45), 14127–14135.

Clarkson, M. G., Clifton, R. K., Swain, I. U., and Perris, E. E. (1989). Stimulus duration and repetition rate influence newborns' head orientation toward sound. *Developmental Psychobiology*, **22**(7), 683–705.

Clifton, R. K., Morrongiello, B. A., Kulig, J. W., and Dowd, J. M. (1981). Newborns' orientation toward sound: possible implications for cortical development. *Child Development*, **52**(3), 833.

Clifton, R. K., Morrongiello, B. A., and Dowd, J. M. (1984). A developmental look at an auditory illusion: the precedence effect. *Developmental Psychobiology*, **17**(5), 519–536.

Clifton, R. K., Muir, D. W., Ashmead, D. H., and Clarkson, M. G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, **64**(4), 1099–1110.

Cranford, J., Ravizza, R., Diamond, I. T., and Whitfield, I. C. (1971). Unilateral ablation of the auditory cortex in the cat impairs complex sound localization. *Science*, **172**(3980), 286–288.

Crassini, B. and Broerse, J. (1980). Auditory-visual integration in neonates: a signal detection analysis. *Journal of Experimental Child Psychology*, **29**(1), 144–155.

Dehaene-Lambertz, G. and Dehaene, S. (1994). Speed and cerebral correlates of syllable discrimination in infants. *Nature*, **370**(6487), 292–295.

Deouell, L. Y., Parnes, A., Pickard, N., and Knight, R. T. (2006). Spatial location is accurately tracked by human auditory sensory memory: evidence from the mismatch negativity. *European Journal of Neuroscience*, **24**(5), 1488–1494.

Eggermont, J. and Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In L. Werner, R. Fay, and A. Popper, editors, *Springer Handbook of Auditory Research: Human Auditory Development*, pages 61–105. Springer, New York, NY.

Escera, C., Alho, K., Schröger, E., and Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neurotology*, **5**(3-4), 151–166.

Fellman, V., Kushnerenko, E., Mikkola, K., Ceponiene, R., Leipälä, J., and Näätänen, R. (2004). Atypical auditory event-related potentials in preterm infants during the first year of life: a possible sign of cognitive dysfunction? *Pediatric Research*, **56**(2), 291–297.

Fitzpatrick, D. C., Kuwada, S., Kim, D. O., Parham, K., and Batra, R. (1999). Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. *The Journal of the Acoustical Society of America*, **106**(6), 3460–3472.

Friederici, A. D., Friedrich, M., and Weber, C. (2002). Neural manifestation of cognitive and precognitive mismatch detection in early infancy. *Neuroreport*, **13**(10), 1251–1254.

Friedrich, M., Weber, C., and Friederici, A. D. (2004). Electrophysiological evidence for delayed mismatch response in infants at-risk for specific language impairment. *Psychophysiology*, **41**(5), 772–782.

Fujioka, T., Mourad, N., He, C., and Trainor, L. J. (2011). Comparison of artifact correction methods for infant EEG applied to extraction of event-related potential signals. *Clinical Neurophysiology*, **122**(1), 43–51.

Grimm, S., Recasens, M., Althen, H., and Escera, C. (2012). Ultrafast tracking of sound location changes as revealed by human auditory evoked potentials. *Biological Psychology*, **89**(1), 232–239.

Hawley, M. L., Litovsky, R. Y., and Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: effect of location and type of interferer. *The Journal of the Acoustical Society of America*, **115**(2), 833–843.

He, C., Hotson, L., and Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, **19**(5), 878–892.

He, C., Hotson, L., and Trainor, L. J. (2009). Maturation of cortical mismatch responses to occasional pitch change in early infancy: effects of presentation rate and magnitude of change. *Neuropsychologia*, **47**(1), 218–229.

Hirsh, I. J. (1950). The relation between localization and intelligibility. *The Journal of the Acoustical Society of America*, **22**(2), 196–200.

Hollich, G., Newman, R. S., and Jusczyk, P. W. (2005). Infants' use of synchronized visual information to separate streams of speech. *Child Development*, **76**(3), 598–613.

Hughes, H. C., Reuter-Lorenz, P. A., Nozawa, G., and Fendrich, R. (1994). Visual-auditory interactions in sensorimotor processing: saccades versus manual responses.

Journal of Experimental Psychology: Human Perception and Performance, **20**(1), 131–153.

Johnson, M. H. (1990). Cortical maturation and the development of visual attention in early infancy. *Journal of Cognitive Neuroscience*, **2**(2), 81–95.

Kaiser, J. and Lutzenberger, W. (2001). Location changes enhance hemispheric asymmetry of magnetic fields evoked by lateralized sounds in humans. *Neuroscience Letters*, **314**(12), 17–20.

Kaiser, J., Lutzenberger, W., Preissl, H., Ackermann, H., and Birbaumer, N. (2000a). Right-hemisphere dominance for the processing of sound-source lateralization. *The Journal of Neuroscience*, **20**(17), 6631–6639.

Kaiser, J., Lutzenberger, W., and Birbaumer, N. (2000b). Simultaneous bilateral mismatch response to right-but not leftward sound lateralization. *Neuroreport*, **11**(13), 2889–2892.

King, A. J. (2009). Visual influences on auditory spatial learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **364**(1515), 331–339.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., and Näätänen, R. (2002). Maturation of the auditory change detection response in infants: a longitudinal ERP study. *Neuroreport*, **13**(15), 1843–1848.

Leppänen, P. H. T., Eklund, K. M., and Lyytinen, H. (1997). Eventrelated brain potentials to change in rapidly presented acoustic stimuli in newborns. *Developmental Neuropsychology*, **13**(2), 175–204.

Lewis, T. L. and Maurer, D. (1992). The development of the temporal and nasal visual fields during infancy. *Vision Research*, **32**(5), 903–911.

Lewkowicz, D. J. (2002). Heterogeneity and heterochrony in the development of intersensory perception. *Cognitive Brain Research*, **14**(1), 41–63.

Lieu J (2004). Speech-language and educational consequences of unilateral hearing loss in children. *Archives of Otolaryngology-Head & Neck Surgery*, **130**(5), 524–530.

Litovsky, R. (2012). Development of binaural and spatial hearing. In L. Werner, R. Fay, and A. Popper, editors, *Springer Handbook of Auditory Research: Human Auditory Development.*, pages 163–195. Springer, New York, NY.

Litovsky, R. and Ashmead, D. H. (1997). Developmental aspects of binaural and spatial hearing. In H. Gilkey and T. Anderson, editors, *Binaural and Spatial Hearing*, pages 571–592. Lawrence Erlbaum, Hillsdale, NJ.

Litovsky, R. Y., Johnstone, P. M., and Godar, S. P. (2006). Benefits of bilateral cochlear implants and/or hearing aids in children. *International Journal of Audiology*, **45**(Supplement), 78–91.

Mickey, B. J. (2005). Sensitivity of auditory cortical neurons to the locations of leading and lagging sounds. *Journal of Neurophysiology*, **94**(2), 979–989.

Morrongiello, B. A., Fenwick, K. D., and Chance, G. (1990). Sound localization acuity in very young infants: an observer-based testing procedure. *Developmental Psychology*, **26**(1), 75–84.

Mourad, N., Reilly, J. P., de Bruin, H., Rasey, G., and MacCrimmon, D. (2007). A simple and fast algorithm for automatic suppression of high-amplitude artifacts in

EEG data. In *2007 Ieee International Conference on Acoustics, Speech, and Signal Processing, Vol I, Pts 1-3, Proceedings*, pages 393–396. Ieee, New York.

Muir, D. and Clifton, R. K. (1985). Infants' orientation to the location of sound sources. In G. Gottlieb and N. Krasnegor, editors, *The Measurement of Audition and Vision During the First Year of Life: A Methodological Overview*, pages 171–194. Ablex, Norwood, NJ.

Muir, D. and Hains, S. (2004). The U-shaped developmental function for auditory localization. *Journal of Cognition and Development*, **5**(1), 123–130.

Muir, D., Abraham, W., Forbes, B., and Harris, L. (1979). The ontogenesis of an auditory localization response from birth to four months of age. *Canadian Journal of Psychology*, **33**(4), 320–333.

Muir, D. W. (1985). The development of infants auditory spatial sensitivity. In S. E. Trehub and B. A. Schneider, editors, *Auditory Development in Infancy*, number 10 in Advances in the Study of Communication and Affect, pages 51–83. Springer US.

Muir, D. W., Clifton, R. K., and Clarkson, M. G. (1989). The development of a human auditory localization response: a U-shaped function. *Canadian Journal of Psychology*, **43**(2), 199–216.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, **118**(12), 2544–2590.

Nager, W., Kohlmetz, C., Joppich, G., Mbes, J., and Mnte, T. F. (2003). Tracking of multiple sound sources defined by interaural time differences: brain potential evidence in humans. *Neuroscience Letters*, **344**(3), 181–184.

Neil, P. A., Chee-Ruiter, C., Scheier, C., Lewkowicz, D. J., and Shimojo, S. (2006). Development of multisensory spatial integration and perception in humans. *Developmental Science*, **9**(5), 454–464.

Novitski, N., Tervaniemi, M., Huotilainen, M., and Näätänen, R. (2004). Frequency discrimination at different frequency levels as indexed by electrophysiological and behavioral measures. *Cognitive Brain Research*, **20**(1), 26–36.

Nozza, R. J. (1987). The binaural masking level difference in infants and adults: developmental change in binaural hearing. *Infant Behavior and Development*, **10**(1), 105–110.

Nozza, R. J., Wagner, E. F., and Crandell, M. A. (1988). Binaural release from masking for a speech sound in infants, preschoolers, and adults. *Journal of Speech, Language, and Hearing Research*, **31**(2), 212–218.

Olsho, L. W., Koch, E. G., Halpin, C. F., and Carter, E. A. (1987). An observer-based psychoacoustic procedure for use with young infants. *Developmental Psychology*, **23**(5), 627–640.

Paavilainen, P., Karlsson, M.-L., Reinikainen, K., and Näätänen, R. (1989). Mismatch negativity to change in spatial location of an auditory stimulus. *Electroencephalography and Clinical Neurophysiology*, **73**(2), 129–141.

Pakarinen, S., Takegata, R., Rinne, T., Huotilainen, M., and Näätänen, R. (2007). Measurement of extensive auditory discrimination profiles using the mismatch negativity (MMN) of the auditory event-related potential (ERP). *Clinical Neurophysiology*, **118**(1), 177–185.

Phillips, D. P. and Irvine, D. R. (1983). Some features of binaural input to single neurons in physiologically defined area AI of cat cerebral cortex. *Journal of Neurophysiology*, **49**(2), 383–395.

Richter, N., Schrger, E., and Rübsamen, R. (2009). Hemispheric specialization during discrimination of sound sources reflected by MMN. *Neuropsychologia*, **47**(12), 2652–2659.

Röttger, S., Schröger, E., Grube, M., Grimm, S., and Rübsamen, R. (2007). Mismatch negativity on the cone of confusion. *Neuroscience Letters*, **414**(2), 178–182.

Schröger, E. (1996). Interaural time and level differences: integrated or separated processing? *Hearing Research*, **96**(12), 191–198.

Schröger, E. and Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. *NeuroReport*, **7**(18), 3005–3008.

Sonnadara, R. R., Alain, C., and Trainor, L. J. (2006). Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location. *Brain Research*, **1071**(1), 175–185.

Spierer, L., Tardif, E., Sperdin, H., Murray, M. M., and Clarke, S. (2007). Learning-induced plasticity in auditory spatial representations revealed by electrical neuroimaging. *The Journal of Neuroscience*, **27**(20), 5474–5483.

Steinschneider, M., Fishman, Y. I., and Arezzo, J. C. (2008). Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. *Cerebral Cortex*, **18**(3), 610–625.

Trainor, L. J. (2007). Event-related potential (ERP) measures in auditory development research. In L. Schmidt and S. Segalowitz, editors, *Developmental Psychophysiology: Theory, Systems, and Methods*, pages 69–102. Cambridge University Press, New York, NY.

Trainor, L. J. (2012a). Musical experience, plasticity, and maturation: issues in measuring developmental change using EEG and MEG. *Annals of the New York Academy of Sciences*, **1252**(1), 25–36.

Trainor, L. J. (2012b). Predictive information processing is a fundamental learning mechanism present in early development: evidence from infants. *International Journal of Psychophysiology*, **83**(2), 256–258.

Trainor, L. J., Samuel, S. S., Desjardins, R. N., and Sonnadara, R. R. (2001). Measuring temporal resolution in infants using mismatch negativity. *NeuroReport*, **12**(11), 2443–2448.

Trainor, L. J., McFadden, M., Hodgson, L., Darragh, L., Barlow, J., Matsos, L., and Sonnadara, R. (2003). Changes in auditory cortex and the development of mismatch negativity between 2 and 6 months of age. *International Journal of Psychophysiology*, **51**(1), 5–15.

Vasilenko, Y. A. and Shestopalova, L. B. (2010). Moving sound source discrimination in humans: mismatch negativity and psychophysics. *Human Physiology*, **36**(2), 139–146.

Wertheimer, M. (1961). Psychomotor coordination of auditory and visual space at birth. *Science*, **134**, 1692.

Chapter 3

The simultaneously-evoked auditory potential (SEAP): a new method for concurrent measurement of cortical and subcortical auditory-evoked activity

3.1 Abstract

Recent electrophysiological work has evinced a capacity for plasticity in subcortical auditory nuclei in human listeners. Similar plastic effects have been measured

in cortically-generated auditory potentials but it is unclear how the two interact. Here we present a new method for recording the Simultaneously-Evoked Auditory Potential (SEAP) to concurrently elicit electrophysiological brain potentials from the inferior colliculus, thalamus, and primary and secondary auditory cortices. Twenty-six normal-hearing adult subjects (mean age = 19.26 years, 17 female) were exposed to 2400 monaural (right-ear) presentations of a specially-designed stimulus which consisted of a pure tone carrier (500 or 600 Hz) that had been amplitude-modulated (depth 100%) at the sum of 37 and 81 Hz. Presentation followed an oddball paradigm wherein the pure tone carrier was set to 500 Hz for 85% of presentations and pseudo-randomly changed to 600 Hz for 15% of the presentations. Single-channel electroencephalographic data were recorded from each subject using a vertical montage referenced to the right earlobe. We show that this stimulation elicits SEAP: a 500 Hz frequency-following response (FFR; generated in inferior colliculus), 80 (subcortical) and 40 (primary auditory cortex) Hz auditory steady-state responses (ASSRs), mismatch negativity (MMN) and P3a (when there is an occasional change in carrier frequency; secondary auditory cortex) in addition to the obligatory N1-P2 complex (secondary auditory cortex). Analyses showed that subcortical and cortical processes are linked as (i) the latency of the FFR predicts the phase delay of the 40 Hz ASSR, (ii) the phase delays of the 40 and 80 Hz ASSRs are correlated, and (iii) the fidelity of the FFR predicts the latency of the N1 component. The SEAP method offers a new approach for measuring the dynamic encoding of acoustic features at multiple levels of the auditory pathway. As such, SEAP is a promising tool with which to study how relationships between subcortical and cortical processes change through early development and auditory learning as well as by hearing loss and aging.

3.2 Introduction

Auditory-evoked potentials (AEPs) are often used to non-invasively examine auditory processing dynamics in human listeners (Chandrasekaran *et al.*, 2014). Acoustic transduction via the inner ear initiates volleys of synchronous neural depolarization along the auditory pathway. Under certain conditions, the post-synaptic potentials generated by populations of depolarizing neurons will summate and create propagating fluctuations in voltage that can be measured at the listener's scalp using electroencephalographic (EEG) electrodes. Averaging the EEG signal evoked over multiple stimulus presentations effectively attenuates spontaneous neural activity and increases the signal-to-noise ratio for EEG activity specifically evoked by the acoustic stimulus: the AEP. A number of so-called "components" can be identified in the human AEP, the exact combination of which depends largely on the properties of the acoustic stimulus, presentation parameters, and task of the listener (Luck, 2005).

Of relevance to the current study, different AEP components have been attributed to functionally discrete stages of auditory processing and, most notably, localized to different subcortical and cortical generator sites along the auditory pathway (Eggermont, 2007). There is thus potential for AEP methods to measure how sound information is represented and transformed as it ascends the auditory neuraxis. However, to date relatively few studies have examined subcortical and cortical AEP components from the same listeners and only some have made efforts to limit between-subject and/or between-session variability by recording subcortical and cortical components simultaneously in individuals (Bidelman and Alain, 2015; Bidelman *et al.*, 2014a,b, 2013; Bidelman, 2015,b; Krishnan *et al.*, 2012; Musacchia *et al.*, 2008; Shiga *et al.*,

2015; Sohmer and Feinmesser, 1970; Tietze, 1979; Woods *et al.*, 1993). Therefore, the relationship between cortical and subcortical AEP components remains unclear.

Developing methods for concurrent subcortical-cortical AEP measurement would be particularly useful experimentally for elucidating the neural dynamics involved in auditory learning as well as clinically for lesion detection and differential diagnosis of hearing loss. For example, recent research suggests that auditory learning results in changes to the way sound is encoded at subcortical nuclei below the level of the auditory cortex (AC). In humans, much of this research is focused on the frequency-following response (FFR) owing to its potential as an index of acoustic feature representation in human subcortex (Skoe and Kraus, 2010). The FFR is thought to reflect population-level synaptic activity from nuclei primarily within the rostral brainstem, namely the inferior colliculus (IC), which phase-locks to the periodicity of the evoking acoustic stimulus (Chandrasekaran and Kraus, 2010; Smith *et al.*, 1975, 1978; Sohmer *et al.*, 1977; Worden and Marsh, 1968), though it may also contain significant activity from auditory nerve (Bidelman, 2015) and, at least for FFRs below 100 Hz, activity from cortical generators (Coffey *et al.*, 2016). As a convergence hub, IC is the target of many ascending and descending projections within the auditory system (Winer, 2005). A wide variety of neural cell types with unique discharge patterns populate the IC and, together, are capable of representing complex auditory signals with high temporal precision (Peruzzi *et al.*, 2000). This precision is partly reflected in the fidelity of the FFR signal, which can represent periodicity in the evoking stimulus up to at least 1000 Hz (Chandrasekaran *et al.*, 2014; Kraus and Nicol, 2005). Crucially, human FFR data reveal patterns of morphological change which correlate with acoustic experience, such as with music or

language (Bidelman *et al.*, 2011; Krishnan *et al.*, 2005, 2008, 2009, 2010b,a; Krizman *et al.*, 2012; Skoe *et al.*, 2013; Wong *et al.*, 2007) as well as with short-term acoustic training (Anderson *et al.*, 2013; Carcagno and Plack, 2011; Russo *et al.*, 2005; Song *et al.*, 2008) and real-time statistical learning (Skoe *et al.*, 2013; Skoe and Kraus, 2010b). Insofar as the FFR primarily reflects the activity of subcortical nuclei, these experience-dependent effects complement a considerable body of research showing similar individual- and group-level effects in the morphology and topography of both transient and sustained components attributed to cortical generators. For example, the obligatory N1 and P2 transients—often localized to generators in secondary AC (Engelien *et al.*, 2000; Godey *et al.*, 2001; Picton *et al.*, 1999)—stay neuroplastic into adulthood and can be enhanced following auditory training (Shahin *et al.*, 2003; Tremblay *et al.*, 2001, 2002) as well as passive exposure to acoustic stimuli (Ross and Tremblay, 2009). Additional transient components generated outside of primary AC, such as the mismatch negativity (MMN) and P3a, also exhibit sensitivity to passive exposure, short-term training, and long-term acoustic experience such as with music and language (Atienza *et al.*, 2004; Näätänen, 2008; Nikjeh *et al.*, 2009; Uther *et al.*, 2006). These so-called novelty-detection potentials are typically associated with the auditory-oddball paradigm wherein an infrequent sound or sequence of sounds is presented pseudo-randomly in an otherwise repetitive sound stream. The MMN is thought to reflect automatic auditory deviance-detection processes localized to secondary AC (Näätänen *et al.*, 2007; Picton *et al.*, 2000), whereas the P3a appears to reflect alerting processes in frontal lobe triggered when novel auditory features cause an involuntary reorienting of attention (for reviews see: Polich, 2007; Yamaguchi and Knight, 1991). Responses from primary AC are also affected by experience. Cortical

auditory steady-state responses (ASSRs), elicited by amplitude modulating a carrier signal at or around 40 Hz (Galambos *et al.*, 1981; Ross *et al.*, 2000), exhibit phase advancement in adult listeners following both auditory training and passive exposure (Bosnyak *et al.*, 2004, 2007; Gander *et al.*, 2010). The 40 Hz ASSR predominantly reflects activity in primary AC, with additional contributions from thalamus, thalamocortical circuits, and brainstem (Herdman *et al.*, 2002).

One possibility is that corticofugal projections allow AC to modulate activity and/or plastic processes at subcortical nuclei to enhance biologically-relevant spectrotemporal features of the impinging sound (Suga *et al.*, 2000, 2002; Suga, 2008; Suga and Ma, 2003; Wu and Yan, 2007; Zhou and Jen, 2007). An emerging literature questions whether such a feedback network can be measured in human listeners, but current evidence from concurrent recording of subcortical and cortical AEP components, though promising, remains inconclusive. For example, Musacchia *et al.* (2008) found that experience (i.e., musicianship) not only affected the morphology of components from both levels of the auditory hierarchy, but also modulated how strongly subcortical (FFR) and cortical (P1-N1) components were related. Similarly, Krishnan *et al.* (2012) found that a measure of pitch saliency in the subcortical FFR correlated with the magnitude of a late cortical potential elicited by stimuli containing cues to pitch. Further, both subcortical and cortical representations of pitch processing predicted behavioural pitch discrimination limens better than either component alone. Bidelman *et al.* (2014b) failed to find a correlation between the morphology of early cortical components (i.e., P1-N1) and the subcortical FFR, but did find that the amplitude of the later cortical P2 component was significantly correlated with FFR encoding of the evoking speech signal's first formant frequency. The correlation

between brainstem FFR and P2 was also stronger in young adult musicians compared to age-matched non-musicians. However, a subsequent study in older adult musicians found that the combined morphology of early cortical components (N1-P2 peak-to-peak amplitude) together with the strength of the subcortical FFR predicted performance on a categorical speech task better than the morphology of either level alone (Bidelman and Alain, 2015).

Some ambiguity in the aforementioned studies may be attributable to differences in recording methodology. Concurrent measurement of AEPs from multiple levels of the auditory pathway involves finding a compromise between the optimal parameters for eliciting cortical and subcortical components. Cortical AEPs require slow presentation rates due to stark refractory effects (Davis *et al.*, 1966; Picton *et al.*, 1977). For example, N1 amplitude increases by $\sim$5.6 μV for every tenfold increase in inter-stimulus interval (ISI) from 0.5 to 3 s (Nelson and Lassman, 1968). Conversely, brainstem generators are relatively unaffected by reductions in ISI down to about 10 ms (Ballachanda *et al.*, 1992), though an ISI of 50 ms is considered ideal for recording subcortical FFR (Bidelman, 2015b). Subcortical AEP components also need to propagate through more tissue relative to cortical AEPs before reaching the scalp. Hence, the subcortical components are smaller in amplitude than cortical components and require averaging over a higher number (2000) of trials to achieve an acceptable signal-to-noise ratio (SNR). Accurate recording of the FFR signal up to 1500 Hz further requires an EEG amplifier capable of sampling scalp voltage at a much higher rate than required to faithfully capture cortical AEPs whose response morphology can be represented using sampling rates as low as 100 Hz (Burkard *et al.*,

2007). This raises concerns about computer processing power and data storage when stimulus duration and ISI are extended for optimal recording of cortical components.

Some studies (e.g., Krishnan *et al.*, 2012; Musacchia *et al.*, 2008) have used a relatively short ISI (500 ms) to find a balance between the optimal ISI for recording cortical responses and the time required to simultaneously collect a sufficient number of trials to visualize subcortical components. This approach cuts down on overall recording time, but attenuates early cortical components such as P1, N1, and P2 relative to recordings using longer ISIs. Others (e.g., Bidelman *et al.*, 2014a,b; Bidelman and Alain, 2015) have used a blocked design wherein cortical responses are collected in one block with a long ISI and subcortical responses are collected in a separate block where stimulus presentation is rapid. Clustered presentation further refines the blocked design by presenting stimulus clusters—where each stimulus is separated by a short ISI—at longer intervals in order to measure a robust cortical response to the first stimulus in each cluster (Bidelman, 2015b). Though blocked and clustered designs allow for optimal recording of subcortical and cortical components within a single experimental session, neither approach results in true simultaneous measurement.

Here we propose a new method termed the Simultaneously-evoked Auditory Potential (SEAP) to examine the relationship between concurrently-elicited subcortical and cortical AEP components. Central to the SEAP method is a stimulus consisting of a pure tone carrier frequency (500 Hz) that is amplitude-modulated at the sum of 37 and 81 Hz. The stimulus is 512 ms in duration with a 500 ms ISI and presented to participants monaurally through a calibrated acoustic transmission line over 2400 trials at a fixed polarity. The carrier frequency is randomly changed from 500 to

600 Hz in 15% of trials. The SEAP method should be able to evoke all AEP components reviewed above. First, presentation of a fixed polarity stimulus with a 500 Hz carrier frequency over 2070 trials should suffice to elicit a detectable FFR at 500 Hz in the averaged AEP (Krishnan, 2007; Thornton, 2007). Further, stimulation occurs via an acoustic transmission line that increases the separation of the transducer from the recording electrodes. Hence, the FFR trace should offer an artifact-free index of spectro-temporal feature representation at the subcortical level. Second, the amplitude-modulation rates of 37 and 81 Hz should elicit ASSRs at 37 and 81 Hz, from cortical and subcortical generators respectively. Third, passive auditory stimulation with an ISI of 500 ms should evoke the obligatory N1 and P2 components. Fourth, the infrequent transition in carrier frequency from 500 to 600 Hz should elicit the MMN, and as the carrier frequencies are easily discriminated by normal-hearing adults (Wier *et al.*, 1977), the infrequent transition should also elicit the P3a.

This experiment intends, first, to demonstrate the viability of the SEAP method as a means of concurrently eliciting all aforementioned AEP components and, second, to collect normative data regarding how the subcortical AEP components elicited by SEAP (i.e., 81 Hz ASSR and FFR) are related to cortical components (i.e., N1, P2, MMN, P3a, and 40 Hz ASSR) within individuals during passive exposure. We expect the morphology of the subcortical FFR to predict the morphology of cortical components (e.g., 37 Hz ASSR, N1, P2), as the generators of these AEPs rely on the acoustic information being relayed forward from IC. Further, the relative onset of steady-state components (FFR and 37/81 Hz ASSRs) may be correlated across listeners in part because these components may share overlapping generators and

thus be similarly affected by individual differences in the conduction velocity and response properties of neurons along the ascending auditory pathway.

3.3 Materials and methods

3.3.1 Participants

A total of 26 adult undergraduate subjects (17 female; mean = 19.26, SD = 1.76 years) were recruited from the McMaster Undergraduate Psychology subject pool. After obtaining informed consent, subjects were asked to complete a brief hearing history questionnaire and all subjects self-reported normal hearing. Participation was remunerated in partial course credit. The research protocol was approved by the McMaster Research Ethics Board in accordance with World Medical Association Declaration of Helsinki.

3.3.2 Stimuli

Stimuli were created by amplitude modulating pure tone carriers at the sum of 37 and 81 Hz to a depth of 100% (Fig. 3.1) according to Eq. 3.1:

$$y(t) = \sin(2\pi f_{carrier} t) \times ([\sin(2\pi f_{AM1} t) + \sin(2\pi f_{AM2} t)] + 1) \tag{3.1}$$

Modulation rates of 37 and 81 Hz were chosen so as to be non-harmonically related but close to the modulation rates evoking maximum amplitude ASSRs (i.e., 40 and 80 Hz) without creating additional modulation of the carrier signal at f2-f1 and allowing for later response extraction and analysis in the frequency domain. Stimuli

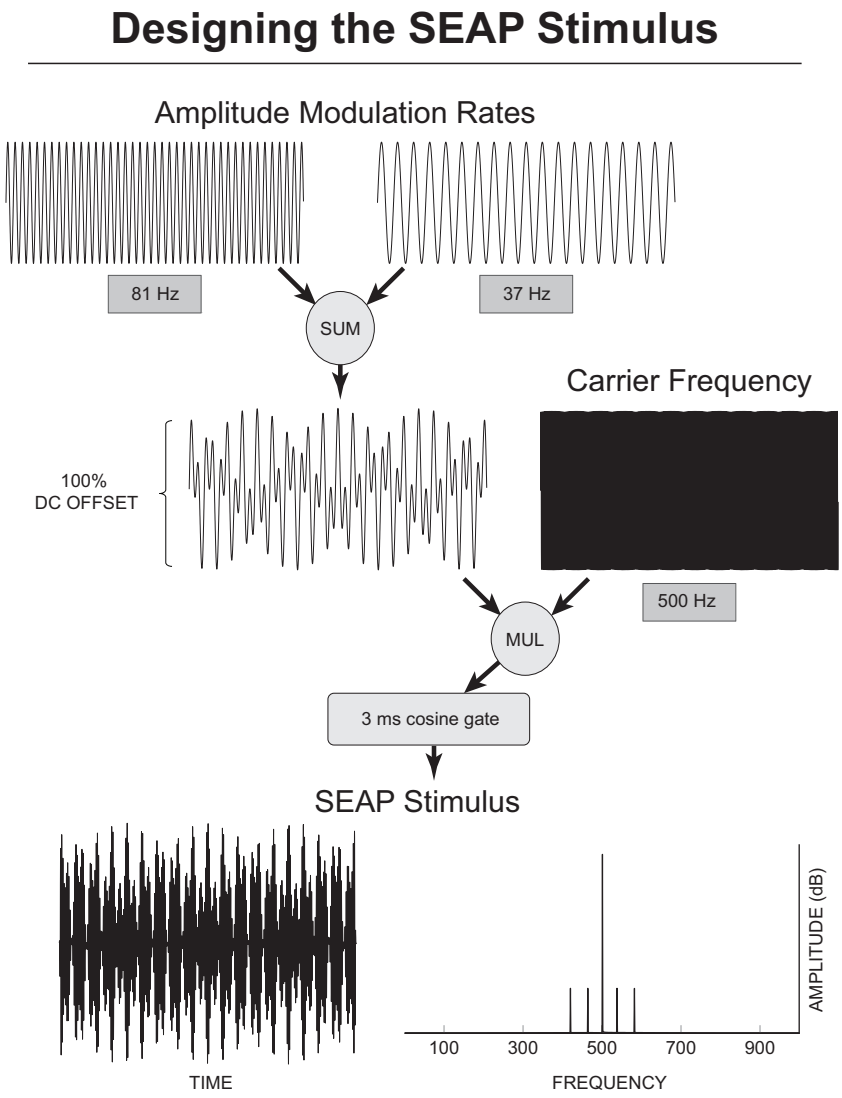

FIGURE 3.1: Schematic representation of the SEAP stimulus generation process. Two amplitude modulation rates are chosen near 40 and 80 Hz such that they do not share a harmonic relationship. The amplitude modulation waves are summed and DC-offset by 100%. A carrier frequency (represented here by 500 Hz) is then multiplied by the summed complex modulation wave. Stimulus onset and offset are gated by a 3 ms cosine gate. The resultant stimulus is depicted in the time (bottom left) and frequency (1–1000 Hz; bottom right) domains.

were generated online using a Tucker-Davis Technologies (TDT) RP2.1 Enhanced Real-time Processor controlled by TDT RPvdsEX (v.5.4) software. The software ran on a Compaq Evo D51C (Intel P4 @ 2.4 Ghz, 1 GB RAM, Windows XP x86 SP2) and controlled the TDT RP2.1 via a USB interface. Signal output from the TDT RP2.1 was routed through a TDT P5A Programmable Attenuator set at 23.9 dB to present the stimulus at a C-weighted sound pressure level (SPL-C) of 83 dB (normalized hearing level of 70 dB as determined from thresholds measured in a different group of subjects). Attenuator output was passed to a TDT HB7 Headphone Driver (0 dB gain) which drove a single Etymotic ER-3A (10 Ohm) ear-insert transducer. Presentation levels were calibrated using a Bruël & Kjær Artificial Ear (Type 4152) connected to a Bruël & Kjær 2260 Investigator sound level meter (SLM).

To prevent electromagnetic contamination of the EEG data, the ER-3A transducer was kept outside of the sound-attenuated testing chamber. Sound was delivered to each subject through a 292 cm length of black flexible polyvinyl chloride tubing (6 mm inside-diameter) coupled to the stock 27.6 cm length of tubing shipped with the ER-3A unit (1.93 mm inside-diameter). Inside the testing chamber, the flexible polyvinyl chloride tube terminated in a 13 mm disposable adult foam ear-insert tip (ERI-14A; 1.93 mm inside-diameter) which was seated in the subject's right-ear canal for stimulus delivery. The entire length of the tube assembly was 320.9 cm, resulting in a theoretical acoustic delay of approximately 9.43 ms (assuming speed of sound = 340.29 m/s at sea level). Separate testing sessions were conducted weekly over the course of the experiment to ensure no stimulus artifact contaminated the EEG recordings. In these sessions the foam tip of the stimulus tube was clamped shut and attached to the collar of a lab member. We then recorded EEG following the experimental

paradigm described below. No energy was found at the stimulus frequency in the averaged EEG responses from any of the artifact-testing sessions.

A digital room correction algorithm (http://drc-fir.sourceforge.net/) was used to correct the frequency response of the modified tube assembly. The algorithm produced a 1000 tap finite impulse response (FIR) filter which was applied online via TDT RPvdsEX software to correct the distribution of spectral energy for all sound output. Analysis of the impulse response generated by the filtered output revealed nearly uniform energy (±3 dB) in the frequency domain from 20–4000 Hz.

A stimulus delay of 44.6 ms was measured as the time between the transistor-transistor logic pulse in the EEG trace and the time of sound arrival at the ear insert. The line level output of the SLM was recorded through a bipolar channel on the SynAmps RT amplifier. All subsequent measures of component latency, as well as markers of stimulus onset, were adjusted to reflect this delay.

3.3.3 Recording paradigm

During the experiment, subjects sat on a comfortable chair in a sound-attenuating booth. To maintain subject arousal, a silent DVD movie (subtitled) was displayed on a computer monitor positioned 1 m directly in front of the chair. Subjects were instructed to remain as still as possible and focus their attention on the movie.

The experimental session consisted of a single block of 2400 trials. Trials were defined by the presentation of a stimulus (512 ms duration; 3 ms rise/fall cosine gate; fixed polarity) followed by a 500 ms silent inter-stimulus interval. The pure tone carrier of the stimulus was set to 500 Hz for 85% of trials (standard trials). To elicit an auditory mismatch response, the frequency of the pure tone carrier was changed to 600

Hz on 15% of trials (deviant trials). Deviant trials were interspersed pseudo-randomly into the stimulus train with the restriction that any deviant trial be separated by at least 2 standard trials (Picton *et al.*, 2000). Standard trials immediately following deviants were eliminated from further analysis as they physically differed from the preceding stimulus. The recording block lasted approximately 40.5 min and the total experimental session, including subject orientation, electrode placement, and debriefing was <60 min. This paradigm collected a total of 2070 FFR trials, 2070 ASSR trials, 1740 N1/P2/MMN standard trials, and 330 MMN/P3a deviant trials per subject, prior to artifact rejection.

We collected EEG data using a Compumedics Neuroscan SynAmps RT amplifier (Model: 9032) and Compumedics SCAN 4.5 Acquire software running on an Intel PC (Intel Core i5 @ 3.33 GHz, 4 GB RAM, Windows 7 x64). The amplifier sampled voltage from the electrodes at a rate of 20,000 Hz using a 24-bit A/D converter, operating in a range of ±200 mV with a least significant bit (LSB) resolution of 23.84 nV (0.400 V/224) in DC-mode. Stimulus onset was denoted by a transistor-transistor logic pulse sent to the SynAmps RT amplifier from the TDT RP.2.1. Three Ag/AgCl sintered electrodes were filled with a conductive gel (Signa Gel) and attached to subjects via double-sided tape washers. Electrode placement followed the vertical montage used in auditory brainstem recording. A ground electrode was applied to the center of the forehead, a reference electrode to the back of the right earlobe (ipsilateral to stimulus delivery), and a recording electrode (Cz) to the vertex of the skull. The vertex was defined as the cranial intersection of the midway point between the ear canals and the midway point between the bridge of the nose and the inion. The recording electrode was connected to a bipolar channel on the SynAmps RT

headbox and measured against the reference channel. This montage is optimal for recording evoked responses of both subcortical and cortical origin (Bidelman, 2015b; Musacchia *et al.*, 2008). Recordings were subject to a hardware band-pass filter of 0.5–3000 Hz and a software notch filter at 60 Hz. Electrode impedance of all subjects was kept below 50 kΩ. Four subjects were excluded from further analysis for failing to maintain electrode impedance below 50 kΩ during the duration of the recording session; the mean impedance for the remaining sample was 25.20 kΩ (SD = 16.27).

3.3.4 Data analysis

Recorded data were segmented into epochs between −94.6 and 655.4 ms, relative to stimulus onset, using BESA software (v.5.1.8). Baseline correction was applied by subtracting the mean of the pre-stimulus window from each epoch. Prior to artifact rejection, the standard and deviant response data were band-pass filtered between 0.5–1000 Hz (12 dB/octave; zero-phase). Epochs were rejected through BESA's artifact rejection tool if they contained activity that exceeded ±80 μV. Accepted trials were averaged according to trial type (standard or deviant) before being transferred to MATLAB R2009a (v.7.8.0.347) for further analysis.

Discrete Fourier transforms (DFTs) of the standard-evoked sustained AEPs (i.e., 37 and 81 Hz ASSRs, 500 Hz FFR) were calculated on a portion of the response ranging from 50 to 450 ms post stimulus onset. A spectral resolution of 1 Hz was interpolated by zero-padding the extracted portion of the response epoch up to 20,000 samples. Amplitudes corresponding to the 500 Hz FFR as well as the 37 and 81 Hz ASSR components were extracted from the DFTs of standard-evoked responses. Amplitude values for each component, as well as 40 neighboring non-response frequency

bins (20 higher and 20 lower), were extracted from each subject's standard DFT data and converted into power for F-test verification as described in Zurek (1992). Significance of the F ratio was then evaluated against critical values of F with 2 and 80 degrees of freedom (Zurek, 1992). This test revealed significant sustained components in the standard-evoked AEPs of all but 2 subjects. Visual inspection confirmed that both subjects were missing a peak in their response power spectra at 37 Hz. One of the two subjects was also missing a peak at 81 Hz and so their data were eliminated from further analysis for failing both the F-test and visual inspection at 2 out of 3 ASSR components. Therefore, the final sample consisted of 21 subjects (15 female; mean age = 19.44, SD = 1.93 years). The actual sample used in each statistical measure consisted of some subset of these 21 subjects that survived outlier elimination in each morphological category, where outliers were defined as values which fell more than 1.5 times the interquartile range above the third quartile or below the first quartile. Sustained AEP analysis was limited to standard-evoked AEPs because the greater trial count compared to deviants allowed for more reliable signal detection using the F-test method.

Unwrapped phase values were also extracted for the 37 and 81 Hz ASSRs. Phase unwrapping makes use of an algorithm to correct the radian phase angles in a vector by adding multiples of $\pm 2\pi$ when absolute jumps between consecutive elements of the vector are greater than or equal to π radians. In this way the instantaneous phase represented by the complex-valued DFT was made continuous allowing for calculation of phase delay as described in John and Picton (2000). In determining phase delay from unwrapped phase values, first, the measured phase was advanced by 90° to reflect the difference between the DFT phase, computed as cosine, and the phase

of the stimulus, which was amplitude modulated with sine functions. Second, the adjusted phase was subtracted from 360° to represent the delay between measured phase and the leading stimulus phase (i.e., phase delay grows larger as separation increases). In order to maintain phase values between 0 and 360°, values below or above this range were adjusted by adding or subtracting 360°, respectively. Means and standard deviations for phase delay data were computed using CircStat, a circular statistics toolbox for MATLAB (Berens, 2009). The Rayleigh test was also used to measure the degree to which phase data polarized for each ASSR component in our sample of listeners. This analysis uses the length of the mean resultant vector of phase angles to test the null hypothesis that data are uniformly distributed about the unit circle (Mardia and Jupp, 2009). Significant polarization would suggest that oscillatory activity at 37 and 81 Hz was phase-locked to the amplitude modulation rates of the evoking stimulus.

Phase delays were not computed for the 500 Hz FFR because the circular nature of these measures makes it difficult to interpret correlations between phases of high frequency signals and those of low frequency signals such as the 37 and 81 Hz ASSRs. The phase-locking capacity of neurons contributing to the FFR was instead assessed by cross-correlating the FFR (standard-evoked response filtered around 500 Hz) with the stimulus waveforms as recorded at the ear-insert (see 3.3.2 Stimuli). Standard-evoked AEPs were band-pass filtered using Chebyshev Type II infinite impulse response filters (IIR) (Table 3.1). Filtering was applied in the forward and reverse direction to ensure zero phase distortion. The latency of each subject's FFR was inferred from the time shift at the maximum peak of the cross-correlation function.

TABLE 3.1: Summary of Chebyshev Type II IIR band-pass filters.

	Frequency (Hz)			**Attenuation (dB)**		
AEP	*Low-stop*	*Passband*	*High-stop*	*Low-stop*	*Passband*	*High-stop*
37 Hz ASSR	3	25–55	200	−20	1	−30
81 Hz ASSR	40	70–92	190	−20	1	−30
500 Hz ASSR	200	400–600	800	−20	1	−30
Transients	0.5	2–20	60	−20	1	−30

To assess transient AEPs, standard and deviant responses from all subjects were forward-reverse band-pass filtered between 2 and 20 Hz using a Chebyshev Type II IIR filter (TABLE 3.1). This filter was used to remove any interfering low-frequency noise as well as energy contributed from the ASSR components. Amplitude and latency measures of the N1 and P2 components were extracted from the filtered standard responses of each subject. The peak of the N1 component was defined as the minimum peak value between 80 and 150 ms post stimulus onset. The peak of the P2 component was defined as the maximum peak value between 120 and 300 ms post stimulus onset. Difference waves were calculated by subtracting the filtered standard response from the filtered deviant response (Picton *et al.*, 2000). The grand mean of all averaged difference waves was tested against zero at each time point using a single sample t-test (two-tailed, $\alpha = 0.05$, Bonferroni corrected). The MMN was defined as the largest negative peak in the difference wave occurring between 115 and 200 ms post stimulus onset where the single sample t-test indicated a significant negative deviation from zero (FIG. 3.2).

Amplitude and latency values of the MMN were recorded for later statistical analyses. Similarly, amplitude and latency measures of the P3a component were

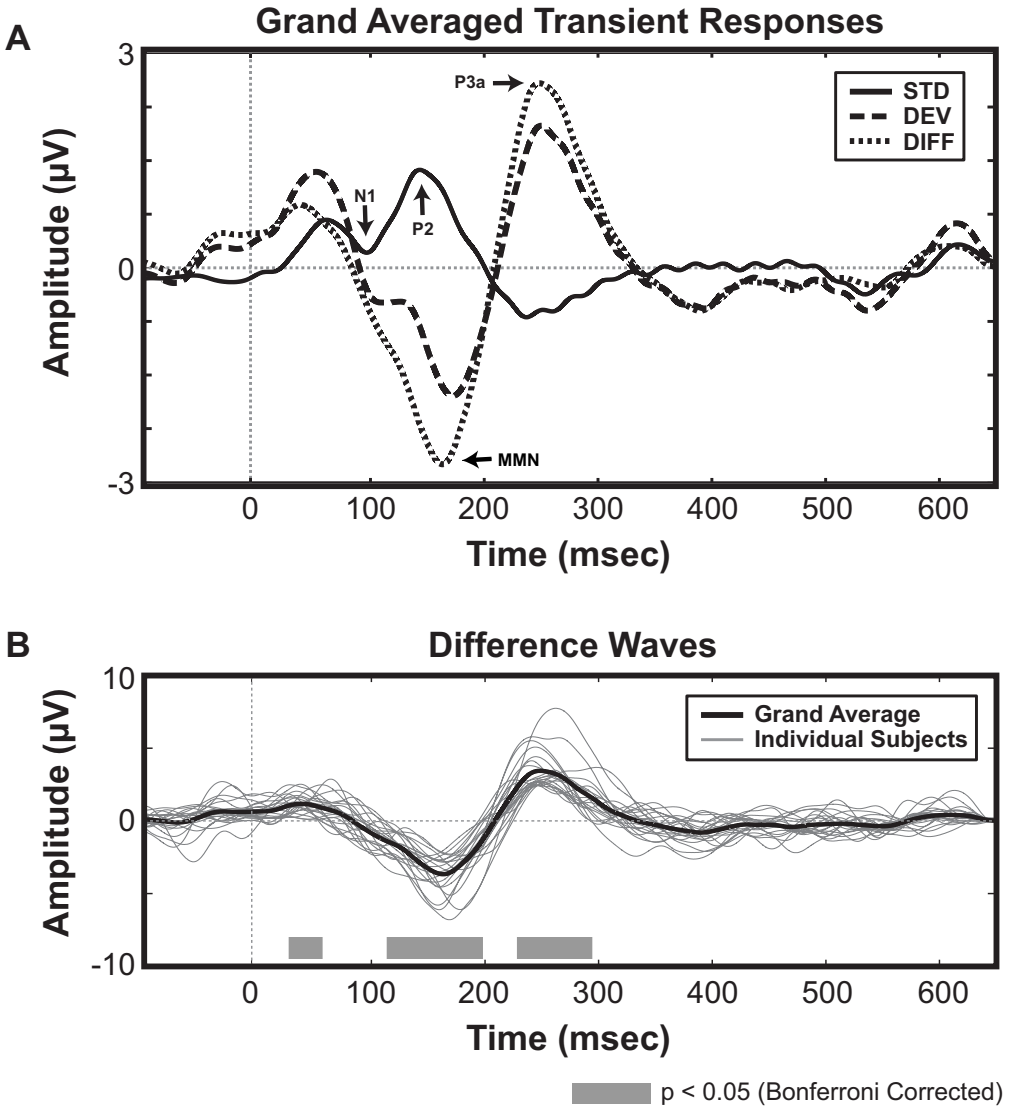

FIGURE 3.2: (A) Grand-averaged ($n = 21$) transient responses elicited by standard (solid trace) and deviant (dashed trace) SEAP stimuli. The difference wave (standard response subtracted from deviant response) is also shown (dotted trace). Labels identify transient component peaks in the grand averaged waves. (B) Differences waves from individual subjects are plotted in light grey with the resultant grand average plotted in black. The grand mean of all averaged difference waves was tested against zero at each time point using a single sample t-test (two-tailed, $\alpha = 0.05$, Bonferroni corrected). Grey bars indicate significant deflections from zero.

extracted from the difference waves. The peak of the P3a component was defined as the maximum peak value between 225 and 295 ms post stimulus onset where the single sample t-test indicated a significant positive deviation from zero (Fig. 3.2). All peaks were selected using a peak picking algorithm in MATLAB and confirmed with visual inspection (Mathworks File Exchange: pickpeaks.m). Grand averages of the standard, deviant, and difference AEP traces were computed by averaging each response type across all participants.

Correlation analyses (2-tailed, $\alpha = 0.05$) were conducted in IBM's SPSS software (v.19.0.0) to examine relationships among the morphological features of the SEAP components. The analyses divided morphological features into timing (latency or phase delay), strength (transient peak and DFT amplitude), and fidelity (FFR-to-stimulus cross-correlation coefficient) information. Pairwise exclusion was applied to cases where data values were missing due to outlier elimination. Normality of the data was verified using the Shapiro-Wilk test ($\alpha = 0.05$). The significance of correlation coefficients computed on data not meeting normality assumptions was assessed using Spearman's ρ; otherwise, all correlations were assessed using Pearson's r. The 95% confidence interval (CI) of each correlation coefficient was estimated by bootstrap sampling the data 1000 times with replacement. Only those correlations whose bootstrapped confidence intervals did not contain zero were considered significant.

Relationships in response timing and strength were measured by separately correlating timing and strength features. To assess whether stimulus fidelity at the level of the IC predicted response morphology at higher levels of the auditory pathway, FFR-to-stimulus cross-correlation coefficients were correlated against peak amplitude and

latency values of the transient components and the DFT amplitude of the 37 Hz ASSR.

3.4 Results

3.4.1 Transient components

The grand averaged standard, deviant, and difference transient AEP waves elicited by the SEAP method are presented in FIGURE 3.2. Average peak amplitude and latency data for the MMN as well as the N1, P2, and P3a components are summarized in TABLE 3.2. The morphology of the MMN response is typical of MMN components elicited by pure tone deviants (Näätänen *et al.*, 2007). Similarly, the average N1-P2 peak-to-peak amplitude (1.88 ±1.22 V, N = 20) corresponds well with values established in the literature (e.g., Tremblay *et al.*, 2001) as does the P3a peak amplitude (e.g., Horváth *et al.*, 2008), suggesting that presentation parameters used in the SEAP method were effective for eliciting known AEP components.

TABLE 3.2: Summary of peak amplitude and latency data from transient AEPs.

Transient AEP	**N**	**Peak Amplitude (μV)**	**Peak Latency (ms)**
MMN	21	-3.76 ± 1.23	166.02 ± 12.13
N1	20	0.05 ± 0.83	106.09 ± 11.82
P2	21	1.98 ± 0.68	153.40 ± 12.31
P3a	21	3.68 ± 1.36	258.81 ± 13.71

3.4.2 Sustained components

Time and frequency domain representations of the grand averaged sustained AEPs (ASSRs/FFRs), as evoked by standard stimulation, are shown in FIGURE 3.3. DFT amplitudes and phase delays of the sustained components as well as stimulus-response cross-correlation coefficients and latency measures for the FFR are summarized in TABLE 3.3. The DFT amplitudes of 37 and 81 Hz ASSR components are in agreement with previously published figures (Bosnyak *et al.*, 2007; D'haenens *et al.*, 2008) as is the DFT amplitude of the 500 Hz FFR (Skoe and Kraus, 2010). Polar plots of the phase delay of 37 and 81 Hz ASSR components are shown in FIGURE 3.4. Combined with the significant results of Rayleigh's tests, these figures illustrate that the observed phase data are non-uniformly distributed around the unit circle, suggesting that neural activity is phase-locked to the amplitude modulation rates of the evoking stimulus.

TABLE 3.3: Summary of steady-state AEP morphology.

Sustained AEP	**N**	**Amplitude (μV$\cdot 10^{-2}$)**	**Phase Delay (deg)**
37 Hz ASSR	21	7.10 ± 3.07	234.61 ± 27.00
81 Hz ASSR	20	3.16 ± 1.82	106.09 ± 31.97
500 Hz FFR	21	1.14 ± 0.71	153.40 ± 67.57
		X-Corr. Coefficient	**X-Corr. Latency (ms)**
500 Hz FFR	21	0.67 ± 0.19	4.23 ± 2.14

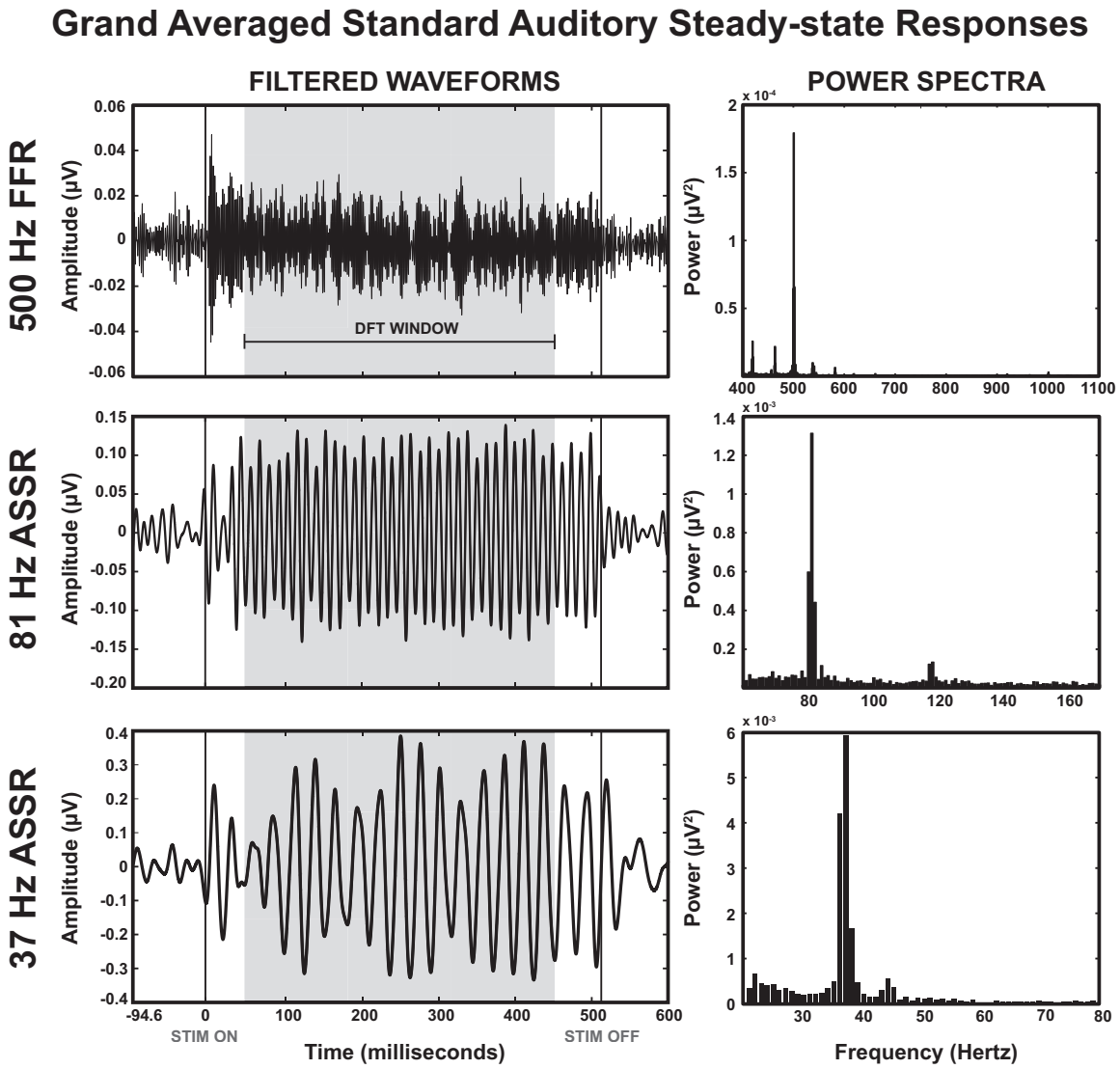

FIGURE 3.3: Grand-averaged ($n = 21$) auditory steady-state (ASSR/FFR) responses elicited by standard SEAP stimuli. The left column displays the grand averaged response in the time-domain as filtered to visualize the 500 Hz FFR (top), 81 Hz ASSR (middle), and 37 Hz ASSR (bottom). Vertical lines at 0 and 512 ms denote stimulus onset and offset, respectively. Light grey blocks in the time-domain indicate the window (50–450 ms post stimulus-onset) over which the DFT was measured. The right column displays the power spectrum of each response as measured from the DFT window.

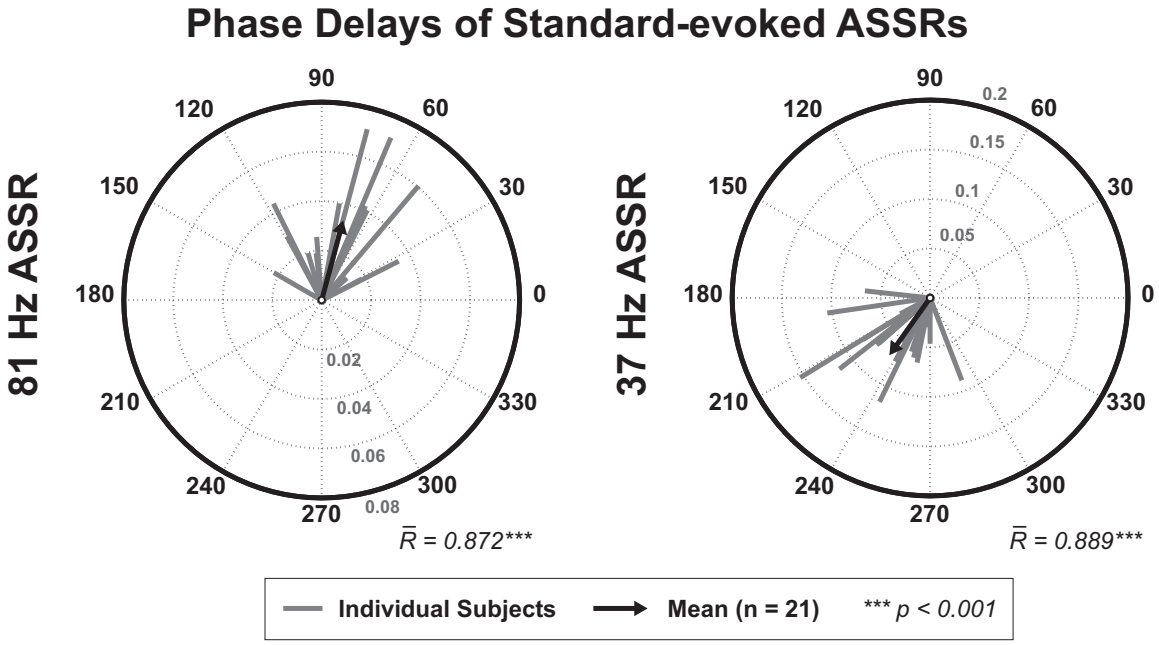

FIGURE 3.4: Polar plots of the phase delay for each ASSR component elicited by standard presentations of the SEAP stimulus. Grey vectors represent the phase delay of individual subject responses. Black arrowhead vectors represent the average phase delay for each component ($n = 21$). Amplitudes are represented by vector length. The mean resultant vector length ($\bar{R}$) of the phase angles with uniform amplitude was significant for both 37 and 81 Hz ASSR components ($p < 0.001$).

3.5 Correlations

Analysis of timing information from the standard-evoked AEP components revealed significant positive correlations between: (1) latency of the 500 Hz FFR and phase delay of the 37 Hz ASSR, and (2) phase delays of the 37 and 81 Hz ASSRs (FIG. 3.5). Thus, the standard-evoked data suggest that the relative onset of 37 Hz ASSR activity is affected by the neural processes reflected in the 81 Hz ASSR and the FFR at 500 Hz, albeit through different mechanisms as the 81 Hz ASSR and FFR were not themselves correlated in time.

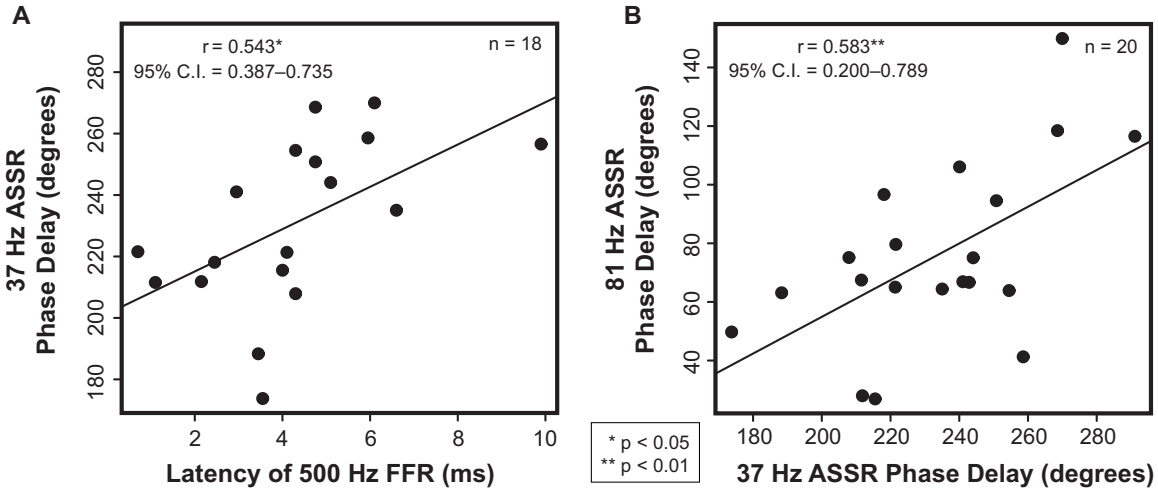

FIGURE 3.5: Scatterplot of significant correlations (two-tailed Pearson's Correlation Analysis, $p < 0.05$) between timing information in the SEAP components as elicited by presentation of the standard stimulus. (A) Latency of the standard 500 Hz FFR component (as measured through cross-correlation with the stimulus waveform) and phase delay of the 37 Hz ASSR component. (B) Phase delays of the 37 and 81 Hz ASSR components.

Analysis of the steady-state spectral amplitude revealed a significant positive correlation between the standard-evoked 500 Hz FFR and 81 Hz ASSR (FIG. 3.6 LEFT). This correlation might reflect proximal or overlapping generators in subcortex. Cross-correlation of the entire FFR signal with the eliciting stimulus waveform produced coefficients that were negatively correlated with N1 Latency (FIG. 3.6 RIGHT), which suggests that greater stimulus fidelity at the level of the IC reduced the time required for N1 generation. There were no other significant correlations among response strength data, nor did other morphological features of the remaining transient or sustained AEPs significantly correlate with FFR signal fidelity.

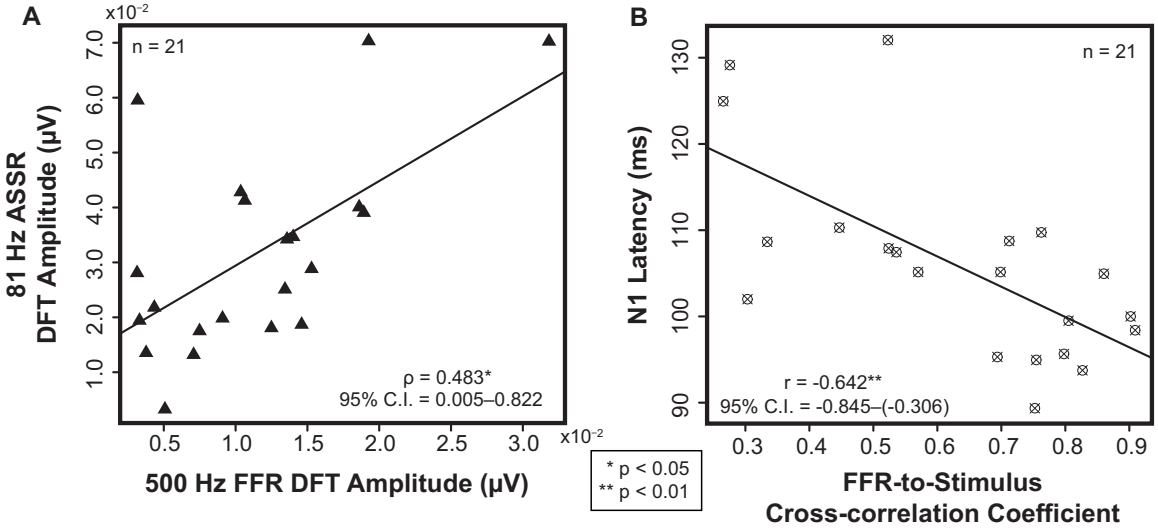

FIGURE 3.6: (A) Scatterplot of a significant correlation (two-tailed Spearman's Correlation Analysis, $p < 0.05$) between amplitude information in the 500 Hz FFR and the 81 Hz ASSR as elicited by frequently-presented (standard) stimuli. (B) Scatterplot of the significant correlation ($p < 0.01$) between the FFR-stimulus cross-correlation coefficient (standard response) and the latency of the transient N1 component.

3.6 Discussion

3.6.1 General

To our knowledge, SEAP is the first stimulation method to simultaneously record cortical N1, P2, MMN, P3a, and 40 Hz ASSR as well as subcortical FFR and 80 Hz ASSR components from normal-hearing adults. Relative to separate recordings (Bidelman *et al.*, 2014a,b; Bidelman and Alain, 2015; Krishnan *et al.*, 2012) or clustered presentation approaches (Bidelman, 2015b), simultaneous measurement has the advantage of controlling for between-subject and between-session variance which might otherwise obfuscate relationships among AEP component morphologies. In this way, the

SEAP method offers a new way to study how subcortical and cortical auditory processing stages interact in different populations of listeners as well as under different task demands or in response to different listening conditions.

The subcortical mammalian auditory system is uniquely complicated relative to other sensory systems, such as vision or somatosensation. As the AEP represents aggregate activity of many neuron types from the many nuclei of the central auditory system, it is impossible to know precisely which nuclei or cell types contribute to each AEP component or measure their relative contribution. In the discussion that follows, we accept the most widely purported generator sites for each AEP component as discussed in this paper's introduction, and adopt a largely anatomically- and physiologically-derived interpretation. However, as our data reflect the activity recorded from a single channel, all inferences made about the relationship between cortical and subcortical generators are based only on the waveform of the signal and thus remain speculative. In the future, multi-channel recordings might be useful in further isolating the source activity of each AEP component in order to make more confident assertions about the dynamics of auditory processing along the ascending auditory hierarchy.

3.6.2 Correlations between the timing of subcortical sustained components and the phase of cortical 40 Hz ASSR

Using SEAP, we found that the timing of neural activity in primary auditory cortex is correlated with the timing of neural activity of at least two subcortical generators. Specifically, during passive listening, the phase delay of the cortical 40 Hz ASSR was related to both the phase delay of the subcortical 80 Hz ASSR as well as the latency

of the FFR at 500 Hz. However, the onsets of the two subcortical AEPs were not related to one another, so they appear to contribute independent input to the 40 Hz ASSR. Current models of the 40 Hz ASSR consider the signal to reflect the combined activity of a peripherally-driven envelope-following mechanism and a centrally-driven oscillatory mechanism (Draganova *et al.*, 2008). One possibility is that the 80 Hz ASSR contributes mainly to the envelope-following mechanism and the FFR to the oscillatory mechanism.

Common subcortical sources of envelope-following activity might partly explain the correlation in timing between 40 and 80 Hz ASSR. Envelope information is first extracted through a demodulation process related to the non-linearity of hair cells in the cochlea (Regan and Regan, 1988). Therefore, the envelope-following activity of the auditory nerve can be preserved in either subcortical or cortical sources insofar as neurons at these sources are capable of phase-locking to the envelope periodicity (Lins *et al.*, 1995). As one ascends the auditory pathway fewer neurons are able to follow higher modulation rates and the limits of observable phase-locking decrease from about 1 to 2 kHz at the auditory nerve to less than 70 Hz in cortex (for a review see: Joris *et al.*, 2004). Dipole solutions for human ASSR data reflect this phase-locking gradient. Sources of the 40 Hz ASSR fit to both cortical and subcortical dipoles, whereas the largest source of the 80 Hz ASSR fits to brainstem generators with minimal contribution from cortical dipoles (Herdman *et al.*, 2002). Intracellular recordings in animal models suggest cochlear nucleus (CN) as the primary subcortical source of scalp-recorded 80 Hz ASSR activity as neurons there respond best to amplitude-modulation rates above 80 Hz (Frisina *et al.*, 1990; Suzuki, 2000). The CN is the terminus of Type I auditory nerve fibers (ANFs), which carry information about

inner hair cell activity via bipolar spiral ganglion cells. The onset of activity in ANFs is, among other factors, affected by the structure (e.g., ear canal and basilar membrane length) and function (e.g., cochlear filter delay) of peripheral hearing organs. The phase delay measured from both 40 and 80 Hz ASSRs is demonstrably sensitive to variation in the timing of acoustic transduction and, for example, increases systematically with increasing carrier frequency as a result of the frequency-dependent travelling wave mechanics of the basilar membrane (Greenberg *et al.*, 1998; John and Picton, 2000; Ross *et al.*, 2000). To the extent that 40 Hz ASSR also reflects subcortical envelope-following activity, the onset of both 40 and 80 Hz ASSR components could be similarly affected by individual differences in the physiology and overall health of peripheral auditory structures or neural conductance velocity across our sample of listeners.

On the other hand, the correlation observed between the timing of the 40 Hz ASSR and the FFR might reflect neural synchrony between activity in the IC and the central oscillatory mechanism of the 40 Hz ASSR rather than an envelope-following mechanism. Cortical and thalamic sources of the 40 Hz ASSR are thought to support local neural oscillations via interactions between excitatory and inhibitory connections within a thalamocortical loop. Neural network models suggest such loops resonate maximally to periodic sensory input at frequencies close to 40 Hz (Llinás and Ribary, 2001). Neurons in the central nucleus of the IC constitute the primary lemniscal input to the ventral division of the medial geniculate body of the thalamus (Malmierca, 2015). Extracellular recordings in the rabbit IC has shown that neurons respond best to modulation rates between 20 and 40 Hz (Batra *et al.*, 1989), implicating the IC as a major subcortical contributor to 40 Hz ASSR. It follows then that the onset of

activity at the IC, as indexed by FFR latency, might strongly influence the rate at which these thalamocortical oscillations enter into a resonant state at 40 Hz.

Further, the timing of activity at the IC appears to be more strongly related to the activity of these thalamocortical circuits than to ascending input as we did not observe a timing correlation between the FFR and the 80 Hz ASSR. It is possible that coordinated timing between the subcortical 80 Hz ASSR and FFR components might have been obscured by individual differences in neuroanatomy and neural conductance velocity. Assuming the primary generator of the 80 Hz ASSR is indeed the CN, then a greater number of synaptic relays separate the 80 Hz ASSR activity from FFR activity at the IC compared to those that separate FFR activity from thalamocortical generators of the 40 Hz ASSR. On the other hand, the correlation between the 80 Hz ASSR and FFR amplitudes suggests these components originate from proximal sources or even partially overlapping sources.

Another possibility is that the latency of processing at the IC might be affected by descending corticofugal projections in a way that does not affect the envelope-following mechanism which generates the 80 Hz ASSR. Indeed, reversible ablation of the auditory cortex in rats prolongs the latency of evoked responses generated in the IC (Nwabueze-Ogbo *et al.*, 2002). Sleep (Aoyagi *et al.*, 1993; Cohen *et al.*, 1991) and general anesthetic (e.g., propofol; Plourde *et al.*, 2008), both of which result in hyperpolarization of thalamocortical neurons, have been found to attenuate both subcortical and cortical sources of the 40 Hz ASSR equally, but have no effect on the 80 Hz ASSR. Moreover, the amplitude of scalp-recorded low-frequency ASSR (around 40 Hz) in rabbit is diminished by administration of pentobarbital anesthesia and by potassium chloride-induced cortical depression. Conversely, low-frequency

ASSR is enhanced through arousal (cocaine administration and tactile stimulation). No such modification is observed in the ASSR elicited by higher (e.g., >80 Hz) modulation rates (Kuwada *et al.*, 2002). Presumably, sleep should not affect subcortical envelope-following responses, though whether sleeping subjects show enhanced timing correlations between the 40 and 80 Hz ASSR components and diminished timing correlations between the 40 Hz ASSR and FFR remains an open question for future research.

One significant limitation inherent in using fixed polarity presentation to measure the FFR is that one cannot rule out possible peripheral contributions from auditory nerve (Bidelman, 2015). Indeed some measures of FFR latency are unrealistically short (<2 ms) for sources in upper brainstem. However, if our FFR signal mainly reflects activity at the AN then it is peculiar that we did not observe a strong timing relationship between FFR latency and the phase of the 80 Hz ASSR whose generators are largely limited to lower brainstem (Herdman *et al.*, 2002). As we did not limit the lag of the largest stimulus-to-FFR cross-correlation coefficient, it is possible that the shortest of these FFR latencies reflects diminishing fidelity of the FFR signal over time in certain individuals. Alternatively, our method of recording the stimulus waveform from the line level output of the SLM into the bipolar channels of our EEG amplifier might have introduced a delay that was unaccounted for when measuring the lag of maximum cross-correlation coefficient. Further, the average amplitudes of many FFR components at 500 Hz and the 80 Hz ASSR fall below the resolution of the EEG system in DC-mode ($\sim 24nV$). These data are thus extrapolated from averaging and any correlations involving the morphology of these components are, in fact, between approximated values. Future studies might use AC-coupled amplifier input to improve

the resolution of the system to 3 nV/bit at the cost of attenuating/distorting low-frequency information in the signal.

3.6.3 Subcortical acoustic feature representation and cortical processing

Our data suggest that the fidelity of acoustic feature representation at subcortical levels is related to the time required for acoustic feature integration at the AC. Specifically, we found the strength of the cross-correlation between the FFR signal and the acoustic stimulus to be negatively associated with the latency of the N1 component. This result adds to the distinction between the 40 Hz ASSR and N1 by demonstrating that the underlying generators of these components differ in their interaction with auditory-evoked activity from subcortical generators. Though both the 40 Hz ASSR and the N1 are often localized to sources in Heschl's gyrus, these components reflect distinct neural networks which differ in both their tuning properties (Ross *et al.*, 2003) and specific dipole solutions (Draganova *et al.*, 2008). Sources of the 40 Hz ASSR reportedly lie deep in Heschl's gyrus bilaterally (Draganova *et al.*, 2002; Engelien *et al.*, 2000; Ross *et al.*, 2002), in regions widely accepted as the location of primary ACs (Da Costa *et al.*, 2011; Penhune *et al.*, 1996). The 40 Hz ASSR also implicates source activity from a midbrain structure likely around the thalamus or IC (Herdman *et al.*, 2002). In contrast, sources of the N1 are exclusively cortical and localized to the lateral part of Heschl's gyrus and the planum temporale (Godey *et al.*, 2001; Pantev *et al.*, 1995). As such, the 40 Hz ASSR probably reflects the earliest activation of primary AC via thalamocortical circuits, whereas N1 involves more synapses spread across primary and secondary AC (Eggermont and Ponton, 2003; Godey *et al.*, 2001).

Unlike the FFR, the N1 is not sensitive to pitch salience per se (Krishnan *et al.*, 2012; Winkler *et al.*, 1997), though it is thought to reflect an indexing of exogenous stimulus features at the level of AC (Alain *et al.*, 2007; Bidelman *et al.*, 2013). Given that auditory information must first be represented at subcortical nuclei before reaching cortex, it is reasonable to speculate that the quality of the subcortical representation might also impact N1 morphology. Consistent with this idea, learning disabilities involving phonological awareness, reading comprehension, and speech-in-noise discrimination are associated both with poor subcortical stimulus representation in the FFR component (Banai *et al.*, 2009, 2005; Hornickel *et al.*, 2009; King *et al.*, 2002) and a delayed N1 latency (Tonnquist-Uhlén *et al.*, 1996). Similar changes to FFR and N1 morphology are observed in senescence and are also accompanied by speech processing deficits (Clinard and Tremblay, 2013; Tremblay *et al.*, 2002). Consistently, auditory training improves stimulus fidelity in the FFR signal (Anderson *et al.*, 2013; Carcagno and Plack, 2011; Song *et al.*, 2008) and advances the onset of N1 (Bosnyak *et al.*, 2004). Extensive auditory training (i.e., musicianship) has even been shown to offset age-related declines in speech sound processing and preserve FFR and N1 morphology in older listeners (Bidelman *et al.*, 2014a; Bidelman and Alain, 2015; Parbery-Clark *et al.*, 2012; Zendel and Alain, 2014).

As relatively few studies have directly compared the FFR and N1 components in individual listeners, the nature of the relationship between FFR generators and early composite measures of cortical activity is still largely unclear. Musacchia *et al.* (2008) found the FFR spectral amplitude of a fundamental frequency component elicited by the speech syllable *da* to strongly predict the slope of the P1-N1 complex. Moreover, the authors found this relationship to be stronger in musicians than non-musicians,

and both the FFR amplitude and P1-N1 slope correlated with the number of years that musicians engaged in musical training. Musicians also exhibited larger and earlier P1 and N1 peaks in response to the *da* stimulus relative to their non-musician counterparts. In contrast, Bidelman *et al.* (2014b) failed to find a difference between young adult musicians and non-musicians in either N1 amplitude or latency when evoked by vowel sounds in a categorical perception task. Additionally, the authors did not find a correlation between the morphology of the N1 and FFR components in either group. However, in older listeners, the same stimulation paradigm evoked earlier N1 components in musicians relative to non-musician controls (Bidelman and Alain, 2015). Further, N1-P2 peak-to-peak amplitudes as well as FFR fundamental amplitudes were robust predictors of musicians' performance on a categorical perception task, whereas only the FFR amplitude predicted non-musicians' performance suggesting that musicianship facilitated coordinated processing in subcortical and cortical generators, perhaps through enhanced corticofugal feedback. Of course, we must keep in mind that all AEPs reflect an aggregate of neural activity and several overlapping subcomponents of the N1 have been identified to respond differentially across different conditions. Differences in stimulation paradigms and stimulus features may engage cortical and subcortical processes differently. Our results, together with earlier studies, suggest that faithful encoding of stimulus features at the level of subcortex is associated with faster cortical processing.

3.6.4 Novelty detection along the auditory hierarchy

Contrary to our expectations, FFR fidelity did not predict the morphology of either the MMN or P3a components. However, change-detection in this experiment was

easily perceptible (pure tone change from 500 to 600 Hz) and so could have saturated MMN generators despite variable spectro-temporal representations at the level of the IC. Other studies have shown change-detection to be a property of multiple stages along ascending auditory neuraxis (for reviews, see: Escera and Malmierca, 2014; Grimm *et al.*, 2016). However, the auditory signal becomes increasingly abstracted as it ascends the auditory neuraxis (Imaizumi and Lee, 2014). The fidelity of spectro-temporal information at subcortex, then, might not strongly influence the morphology of components generated in secondary AC, particularly for stimuli that are readily discriminated. Manipulating subject attention through task demands might engage the corticofugal circuitry required to coordinate the activity reflected in subcortical and cortical AEP component morphology. We are currently using SEAP in a paradigm similar to that of Cacciaglia *et al.* (2015) to examine whether automatic attentional responses to occasional deviants in a stream of standard stimuli modify the relationships we observed here under passive-listening.

3.7 Conclusion

We demonstrate that the SEAP method can viably record several well-studied subcortical and cortical AEP components simultaneously in human adult listeners. In a passive-listening context, better stimulus fidelity at subcortical FFR generators was related to earlier onset of the cortical N1, suggesting that the spectro-temporal fidelity of ascending auditory information promotes neural synchrony and temporal integration at N1 generators. The correlation between the FFR latency and the 40 Hz ASSR phase delay further supports a link between neural activity at IC and early cortical processing. However, we failed to find a relationship between stimulus fidelity at FFR

generators and the morphology of change-detection or attention-orienting components from secondary auditory cortices (i.e., MMN and P3a, respectively). Because a large and easily perceptible stimulus change was used, ceiling effects may have precluded seeing such potential relationships. Future studies might use less perceptible stimulus differences to probe whether MMN elicitation in individual subjects is related to the quality of subcortical acoustic feature representation.

As this experiment only recorded AEPs evoked under passive listening conditions, we make no assertions regarding the causality of these relationships, be they bottom-up, or top-down, or some combination thereof. Presumably, the influence of top-down mechanisms would be relatively limited because this experiment presented pure tones of no behavioural relevance to the listeners. It is also possible that some bottom-up learning occurred during the course of the experiment (for a review of possible bottom-up and top-down influences on subcortical FFR, see: Chandrasekaran *et al.*, 2014). Nevertheless, concurrent measurement of activity from multiple auditory nuclei in individual subjects opens an avenue for assessing dynamic relationships between auditory processing stages in a way that is unaffected by differences in recording paradigms, or by between-subject and between-session variability. Further, true simultaneous AEP measurement permits careful examination of the temporal dynamics between subcortical and cortical AEPs as they might be affected by attention, development, and experience-dependent plastic processes. In this way, the SEAP method should prove particularly useful in delineating the influence of corticofugal projections throughout early infancy as well as in assessing the efficacy of hearing-related interventions on acoustic feature representation at the level of cortex and subcortex.

3.8 References

Alain, C., Snyder, J. S., He, Y., and Reinke, K. S. (2007). Changes in auditory cortex parallel rapid perceptual learning. *Cerebral Cortex*, **17**(5), 1074–1084.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(11), 4357–4362.

Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., and Koike, Y. (1993). Optimal modulation frequency for amplitude-modulation following response in young children during sleep. *Hearing Research*, **65**(12), 253–261.

Atienza, M., Cantero, J. L., and Stickgold, R. (2004). Posttraining sleep enhances automaticity in perceptual discrimination. *Journal of Cognitive Neuroscience*, **16**(1), 53–64.

Ballachanda, B. B., Moushegian, G., and Stillman, R. D. (1992). Adaptation of the auditory brainstem response: effects of click intensity, polarity, and position. *Journal of the American Academy of Audiology*, **3**(4), 275–82.

Banai, K., Nicol, T., Zecker, S. G., and Kraus, N. (2005). Brainstem timing: implications for cortical processing and literacy. *The Journal of Neuroscience*, **25**(43), 9850–9857.

Banai, K., Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Reading and subcortical auditory function. *Cerebral Cortex*, **19**(11), 2699–2707.

Batra, R., Kuwada, S., and Stanford, T. R. (1989). Temporal coding of envelopes and their interaural delays in the inferior colliculus of the unanesthetized rabbit. *Journal of Neurophysiology*, **61**(2), 257–268.

Bidelman, G. M. (2015a). Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, **323**, 68–80.

Bidelman, G. M. (2015b). Towards an optimal paradigm for simultaneously recording cortical and brainstem auditory evoked potentials. *Journal of Neuroscience Methods*, **241**, 94–100.

Bidelman, G. M. and Alain, C. (2015). Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *The Journal of Neuroscience*, **35**(3), 1240–1249.

Bidelman, G. M., Krishnan, A., and Gandour, J. T. (2011). Enhanced brainstem encoding predicts musicians perceptual advantages with pitch. *European Journal of Neuroscience*, **33**(3), 530–538.

Bidelman, G. M., Moreno, S., and Alain, C. (2013). Tracing the emergence of categorical speech perception in the human auditory system. *NeuroImage*, **79**, 201–212.

Bidelman, G. M., Villafuerte, J. W., Moreno, S., and Alain, C. (2014a). Age-related changes in the subcorticalcortical encoding and categorical perception of speech. *Neurobiology of Aging*, **35**(11), 2526–2540.

Bidelman, G. M., Weiss, M. W., Moreno, S., and Alain, C. (2014b). Coordinated plasticity in brainstem and auditory cortex contributes to enhanced categorical

speech perception in musicians. *European Journal of Neuroscience*, **40**(4), 2662–2673.

Bosnyak, D. J., Eaton, R. A., and Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, **14**(10), 1088–1099.

Bosnyak, D. J., Gander, P. E., and Roberts, L. E. (2007). Does auditory discrimination training modify representations in both primary and secondary auditory cortex? *International Congress Series*, **1300**, 25–28.

Burkard, R., Eggermont, J., and Don, M., editors (2007). *Auditory Evoked Potentials: Basic Principles and Clinical Application.* Lippincott Williams & Wilkins, Philadelphia, PA.

Cacciaglia, R., Escera, C., Slabu, L., Grimm, S., Sanjuán, A., Ventura-Campos, N., and Ávila, C. (2015). Involvement of the human midbrain and thalamus in auditory deviance detection. *Neuropsychologia*, **68**, 51–58.

Carcagno, S. and Plack, C. J. (2011). Subcortical plasticity following perceptual learning in a pitch discrimination task. *Journal of the Association for Research in Otolaryngology*, **12**(1), 89–100.

Chandrasekaran, B. and Kraus, N. (2010). The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology*, **47**(2), 236–246.

Chandrasekaran, B., Skoe, E., and Kraus, N. (2014). An integrative model of subcortical auditory plasticity. *Brain Topography*, **27**(4), 539–552.

Clinard, C. G. and Tremblay, K. L. (2013). Aging degrades the neural encoding of simple and complex sounds in the human brainstem. *Journal of the American Academy of Audiology*, **24**(7), 590–599; quiz 643–644.

Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, **7**, 11070.

Cohen, L. T., Rickards, F. W., and Clark, G. M. (1991). A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *The Journal of the Acoustical Society of America*, **90**(5), 2467–2479.

Da Costa, S., van der Zwaag, W., Marques, J. P., Frackowiak, R. S. J., Clarke, S., and Saenz, M. (2011). Human primary auditory cortex follows the shape of Heschl's gyrus. *The Journal of Neuroscience*, **31**(40), 14067–14075.

Davis, H., Mast, T., Yoshie, N., and Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: recovery process. *Electroencephalography and Clinical Neurophysiology*, **21**(2), 105–113.

D'haenens, W., Vinck, B. M., Vel, E. D., Maes, L., Bockstael, A., Keppler, H., Philips, B., Swinnen, F., and Dhooge, I. (2008). Auditory steady-state responses in normal hearing adults: a test-retest reliability study. *International Journal of Audiology*, **47**(8), 489–498.

Draganova, R., Ross, B., Borgmann, C., and Pantev, C. (2002). Auditory cortical response patterns to multiple rhythms of AM sound. *Ear and Hearing*, **23**(3), 254–265.

Draganova, R., Ross, B., Wollbrink, A., and Pantev, C. (2008). Cortical steady-state responses to central and peripheral auditory beats. *Cerebral Cortex*, **18**(5), 1193–1200.

Eggermont, J. (2007). Electric and magnetic fields of synchronous neural activity: peripheral and central origins of auditory evoked potentials. In *Auditory Evoked Potentials: Basic Principles and Clinical Application*. Lippincott Williams & Wilkins, Baltimore, MD.

Eggermont, J. and Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: correlations with changes in structure and speech perception. *Acta Oto-Laryngologica*, **123**(2), 249–252.

Engelien, A., Schulz, M., Ross, B., Arolt, V., and Pantev, C. (2000). A combined functional in vivo measure for primary and secondary auditory cortices. *Hearing Research*, **148**(12), 153–160.

Escera, C. and Malmierca, M. S. (2014). The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology*, **51**(2), 111–123.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hearing Research*, **44**(23), 99–122.

Galambos, R., Makeig, S., and Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *Proceedings of the National Academy of Sciences of the United States of America*, **78**(4), 2643–2647.

Gander, P. E., Bosnyak, D. J., and Roberts, L. E. (2010). Acoustic experience but not attention modifies neural population phase expressed in human primary auditory cortex. *Hearing Research*, **269**(12), 81–94.

Godey, B., Schwartz, D., de Graaf, J. B., Chauvel, P., and Liégeois-Chauvel, C. (2001). Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: a comparison of data in the same patients. *Clinical Neurophysiology*, **112**(10), 1850–1859.

Greenberg, S., Poeppel, D., and Roberts, T. (1998). A space-time theory of pitch and timbre based on cortical expansion of the cochlear traveling wave delay. In A. Palmer, A. Rees, Q. Summerfield, and R. Meddis, editors, *Psychophysical and Physiological Advances in Hearing*, pages 293–300. Whurr, London.

Grimm, S., Escera, C., and Nelken, I. (2016). Early indices of deviance detection in humans and animal models. *Biological Psychology*, **116**, 23–27.

Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M., and Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, **15**(2), 69–86.

Hornickel, J., Skoe, E., Nicol, T., Zecker, S., and Kraus, N. (2009). Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(31), 13022–13027.

Horváth, J., Winkler, I., and Bendixen, A. (2008). Do N1/MMN, P3a, and RON form a strongly coupled chain reflecting the three stages of auditory distraction? *Biological Psychology*, **79**(2), 139–147.

Imaizumi, K. and Lee, C. C. (2014). Frequency transformation in the auditory lemniscal thalamocortical system. *Frontiers in Neural Circuits*, **8**, 1–9.

John, M. S. and Picton, T. W. (2000). Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. *Hearing Research*, **141**(12), 57–79.

Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiological Reviews*, **84**(2), 541–577.

King, C., Warrier, C. M., Hayes, E., and Kraus, N. (2002). Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. *Neuroscience Letters*, **319**(2), 111–115.

Kraus, N. and Nicol, T. (2005). Brainstem origins for cortical what and where pathways in the auditory system. *Trends in Neurosciences*, **28**(4), 176–181.

Krishnan, A. (2007). Frequency-following response. In R. Burkard, J. Eggermont, and M. Don, editors, *Auditory Evoked Potentials: Basic Principles and Clinical Application*, pages 313–335. Lippincott Williams & Wilkins, Philadelphia, PA.

Krishnan, A., Xu, Y., Gandour, J., and Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, **25**(1), 161–168.

Krishnan, A., Swaminathan, J., and Gandour, J. T. (2008). Experience-dependent enhancement of linguistic pitch representation in the brainstem is not specific to a speech context. *Journal of Cognitive Neuroscience*, **21**(6), 1092–1105.

Krishnan, A., Gandour, J. T., Bidelman, G. M., and Swaminathan, J. (2009). Experience dependent neural representation of dynamic pitch in the brainstem. *Neuroreport*, **20**(4), 408–413.

Krishnan, A., Gandour, J. T., and Bidelman, G. M. (2010a). The effects of tone language experience on pitch processing in the brainstem. *Journal of Neurolinguistics*, **23**(1), 81–95.

Krishnan, A., Bidelman, G. M., and Gandour, J. T. (2010b). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, **268**(12), 60–66.

Krishnan, A., Bidelman, G. M., Smalt, C. J., Ananthakrishnan, S., and Gandour, J. T. (2012). Relationship between brainstem, cortical and behavioral measures relevant to pitch salience in humans. *Neuropsychologia*, **50**(12), 2849–2859.

Krizman, J., Marian, V., Shook, A., Skoe, E., and Kraus, N. (2012). Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(20), 7877–7881.

Kuwada, S., Anderson, J. S., Batra, R., Fitzpatrick, D. C., Teissier, N., and D'Angelo, W. R. (2002). Sources of the scalp-recorded amplitude-modulation following response. *Journal of the American Academy of Audiology*, **13**(4), 188–204.

Lins, O. G., Picton, P. E., Picton, T. W., Champagne, S. C., and Durieux-Smith, A. (1995). Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz. *The Journal of the Acoustical Society of America*, **97**(5), 3051–3063.

Llinás, R. and Ribary, U. (2001). Consciousness and the brain. *Annals of the New York Academy of Sciences*, **929**(1), 166–175.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique*. MIT Press, Cambridge, MA.

Malmierca, M. S. (2015). Anatomy and physiology of the mammalian auditory system. In D. Jaeger and R. Jung, editors, *Encyclopedia of Computational Neuroscience*, pages 155–186. Springer New York.

Mardia, K. V. and Jupp, P. E. (2009). *Directional Statistics*. John Wiley & Sons.

Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, **241**(12), 34–42.

Näätänen, R. (2008). Mismatch negativity (MMN) as an index of central auditory system plasticity. *International Journal of Audiology*, **47**(sup2), S16–S20.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, **118**(12), 2544–2590.

Nelson, D. A. and Lassman, F. M. (1968). Effects of intersignal interval on the human auditory evoked response. *The Journal of the Acoustical Society of America*, **44**(6), 1529–1532.

Nikjeh, D. A., Lister, J. J., and Frisch, S. A. (2009). Preattentive cortical-evoked responses to pure tones, harmonic tones, and speech: influence of music training. *Ear and Hearing*, **30**(4), 432–446.

Nwabueze-Ogbo, F. C., Popelar, J., and Syka, J. (2002). Changes in the acoustically evoked activity in the inferior colliculus of the rat after functional ablation of the auditory cortex. *Physiological Research*, **51**, S95–S104.

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., and Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology*, **94**(1), 26–40.

Parbery-Clark, A., Anderson, S., Hittner, E., and Kraus, N. (2012). Musical experience offsets age-related delays in neural timing. *Neurobiology of Aging*, **33**(7), 1483.e1–1483.e4.

Penhune, V. B., Zatorre, R. J., MacDonald, J. D., and Evans, A. C. (1996). Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex*, **6**(5), 661–672.

Peruzzi, D., Sivaramakrishnan, S., and Oliver, D. L. (2000). Identification of cell types in brain slices of the inferior colliculus. *Neuroscience*, **101**(2), 403–416.

Picton, T. W., Woods, D., Baribaeu-Braun, J., and Healy, T. (1977). Evoked potential audiometry. *Journal of Otolaryngology*, **6**(2), 90–119.

Picton, T. W., Alain, C., Woods, D., John, M., Scherg, M., Valdes-Sosa, P., Bosch-Bayard, J., and Trujillo, N. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology and Neurotology*, **4**, 64–79.

Picton, T. W., Alain, C., Otten, L., Ritter, W., and Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology*, **5**(3-4), 111–139.

Plourde, G., Garcia-Asensi, A., Backman, S., Deschamps, A., Chartrand, D., Fiset, P., and Picton, T. W. (2008). Attenuation of the 40-Hz auditory steady state response by propofol involves the cortical and subcortical generators. *Anesthesiology*, **108**, 233–242.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–2148.

Regan, D. and Regan, M. P. (1988). The transducer characteristic of hair cells in the human ear: a possible objective measure. *Brain Research*, **438**(12), 363–365.

Ross, B. and Tremblay, K. (2009). Stimulus experience modifies auditory neuromagnetic responses in young and older listeners. *Hearing Research*, **248**(12), 48–59.

Ross, B., Borgmann, C., Draganova, R., Roberts, L. E., and Pantev, C. (2000). A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *The Journal of the Acoustical Society of America*, **108**(2), 679–691.

Ross, B., Picton, T. W., and Pantev, C. (2002). Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hearing Research*, **165**(1), 68–84.

Ross, B., Draganova, R., Picton, T. W., and Pantev, C. (2003). Frequency specificity of 40-Hz auditory steady-state responses. *Hearing Research*, **186**(12), 57–68.

Russo, N. M., Nicol, T. G., Zecker, S. G., Hayes, E. A., and Kraus, N. (2005). Auditory training improves neural timing in the human brainstem. *Behavioural Brain Research*, **156**(1), 95–103.

Shahin, A., Bosnyak, D. J., Trainor, L. J., and Roberts, L. E. (2003). Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *The Journal of Neuroscience*, **23**(13), 5545–5552.

Shiga, T., Althen, H., Cornella, M., Zarnowiec, K., Yabe, H., and Escera, C. (2015). Deviance-related responses along the auditory hierarchy: combined FFR, MLR and MMN evidence. *Public Library of Science One*, **10**(9), e0136794.

Skoe, E. and Kraus, N. (2010a). Auditory brainstem response to complex sounds: a tutorial. *Ear and Hearing*, **31**(3), 302–324.

Skoe, E. and Kraus, N. (2010b). Hearing it again and again: on-line subcortical plasticity in humans. *Public Library of Science One*, **5**(10), e13645.

Skoe, E., Krizman, J., Spitzer, E., and Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, **243**, 104–114.

Smith, J. C., Marsh, J. T., and Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, **39**(5), 465–472.

Smith, J. C., Marsh, J. T., Greenberg, S., and Brown, W. S. (1978). Human auditory frequency-following responses to a missing fundamental. *Science*, **201**(4356), 639–641.

Sohmer, H. and Feinmesser, M. (1970). Cochlear and cortical audiometry conveniently recorded in the same subject. *Israel Journal of Medical Science*, **6**, 219–223.

Sohmer, H., Pratt, H., and Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology*, **42**(5), 656–664.

Song, J. H., Skoe, E., Wong, P. C., and Kraus, N. (2008). Plasticity in the adult human auditory brainstem following short-term linguistic training. *Journal of Cognitive Neuroscience*, **20**(10), 1892–1902.

Suga, N. (2008). Role of corticofugal feedback in hearing. *Journal of Comparative Physiology A*, **194**(2), 169–183.

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

Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(22), 11807–11814.

Suga, N., Xiao, Z., Ma, X., and Ji, W. (2002). Plasticity and corticofugal modulation for hearing in adult animals. *Neuron*, **36**(1), 9–18.

Suzuki, Y. (2000). Contribution of cochlear nucleus to 80 Hz amplitude-modulation following response. *Nippon Jibiinkoka Gakkai Kaiho*, **103**, 177–187.

Thornton, A. (2007). Instrumentation and recording parameters. In R. Burkard, J. Eggermont, and M. Don, editors, *Auditory Evoked Potentials: Basic Principles and Clinical Application*, pages 42–47. Lippincott Williams & Wilkins, Philadelphia, PA.

Tietze, G. (1979). Stimulation methods for simultaneous derivation of acoustically evoked brainstem and cortical responses. *Scandinavian Audiology. Supplementum*, **11**, 97–104.

Tonnquist-Uhlén, I., Borg, E., Persson, H., and Spens, K. (1996). Topography of auditory evoked cortical potentials in children with severe language impairment: the N1 component. *Electroencephalography and Clinical Neurophysiology*, **100**(3), 250–260.

Tremblay, K., Kraus, N., McGee, T., Ponton, C., Otis, B., and others (2001). Central auditory plasticity: changes in the N1-P2 complex after speech-sound training. *Ear and Hearing*, **22**(2), 79–90.

Tremblay, K. L., Piskosz, M., and Souza, P. (2002). Aging alters the neural representation of speech cues. *Neuroreport*, **13**(15), 1865–1870.

Uther, M., Kujala, A., Huotilainen, M., Shtyrov, Y., and Näätänen, R. (2006). Training in Morse code enhances involuntary attentional switching to acoustic frequency: evidence from ERPs. *Brain Research*, **1073-1074**, 417–424.

Wier, C. C., Jesteadt, W., and Green, D. M. (1977). Frequency discrimination as a function of frequency and sensation level. *The Journal of the Acoustical Society of America*, **61**(1), 178–184.

Winer, J. A. (2005). Decoding the auditory corticofugal systems. *Hearing Research*, **207**(1-2), 1–9.

Winkler, I., Tervaniemi, M., and Näätänen, R. (1997). Two separate codes for missing-fundamental pitch in the human auditory cortex. *The Journal of the Acoustical Society of America*, **102**(2), 1072.

Wong, P. C. M., Skoe, E., Russo, N. M., Dees, T., and Kraus, N. (2007). Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nature Neuroscience*, **10**(4), 420–422.

Woods, D. L., Alain, C., Covarrubias, D., and Zaidel, O. (1993). Frequency-related differences in the speed of human auditory processing. *Hearing Research*, **66**(1), 46–52.

Worden, F. G. and Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology*, **25**(1), 42–52.

Wu, Y. and Yan, J. (2007). Modulation of the receptive fields of midbrain neurons elicited by thalamic electrical stimulation through corticofugal feedback. *The Journal of Neuroscience*, **27**(40), 10651–10658.

Yamaguchi, S. and Knight, R. T. (1991). P300 generation by novel somatosensory stimuli. *Electroencephalography and Clinical Neurophysiology*, **78**(1), 50–55.

Zendel, B. R. and Alain, C. (2014). Enhanced attention-dependent activity in the auditory cortex of older musicians. *Neurobiology of Aging*, **35**(1), 55–63.

Zhou, X. and Jen, P. H.-S. (2007). Corticofugal modulation of multi-parametric auditory selectivity in the midbrain of the big brown bat. *Journal of Neurophysiology*, **98**(5), 2509–2516.

Zurek, P. M. (1992). Detectability of transient and sinusoidal otoacoustic emissions. *Ear and Hearing*, **13**(5), 307–310.

Chapter 4

Concurrent measurement of novelty-detection processes from multiple stages of the auditory hierarchy

4.1 Abstract

Recent work suggests that acoustic change-detection processes operate pre-cortically as well as cortically. However, little is known regarding the relationship between novelty-detection processes at different levels of the auditory pathway. Here we use stimuli designed to simultaneously elicit auditory-evoked potential (AEP) components from multiple stages along the auditory pathway to assess distributed change-detection in 20 normal-hearing adult listeners (15 female; mean age = 19.59, SD =

1.57 years). In addition to eliciting novelty-detection components attributed to generators in secondary auditory cortex (mismatch negativity; MMN) and frontal lobe (P3a), our results show that acoustic deviant stimulation was associated with changes in the morphology of AEP components ascribed to neural activity in subcortical nuclei (frequency-following response; FFR) and thalamocortical circuits (40 Hz auditory steady-state response; ASSR). Further, under certain conditions, we found that novelty-detection processes at subcortical and cortical generators were linked. Relative to the standard-evoked response, the amplitude of the deviant-evoked 700 Hz FFR (but not the 300 Hz FFR) correlated negatively with the latencies of both MMN and P3a components and increased significantly in a window immediately following both components. A follow-up experiment suggested that the frequency specificity of this effect may be related to the large frequency change used in the first experiment.

4.2 Introduction

The mammalian auditory system is adept at detecting novel events in the acoustic environment even in the absence of focused attention. Detecting acoustic change can be vital to the organism as such change may indicate the presence of a threat or of an important communicative signal from a conspecific. In humans, the mechanisms implicated in auditory change-detection involve large populations of cortical neurons such that their activity can be reliably measured at the scalp using electroencephalography (EEG). The mismatch negativity (MMN) and P3a or novelty-P3, both of which have been identified as neural correlates of deviant detection (Escera and Corral, 2007; Näätänen *et al.*, 2007), are two of the largest and most well-characterized late components of the auditory-evoked potential (AEP). Both can be elicited using the oddball

paradigm wherein an infrequent (deviant) sound is presented pseudo-randomly among frequently repeated (standard) sounds. The MMN is characterized as a scalp-frontal negative deflection peaking at 100–250 ms after the onset of a deviant sound and is best visualized in the difference waveform obtained by subtracting the AEP elicited by standard stimulation from that elicited by deviant stimulation. The MMN is pre-attentive, in that it can be elicited without directed listener attention and its topographic distribution across the scalp implicates bilateral generators in secondary auditory cortex (Alho, 1995; Näätänen *et al.*, 2007; Schönwiesner *et al.*, 2007). In contrast to the pre-attentive quality of the MMN, the P3a component—a scalp-frontal positive deflection in the difference waveform which peaks between 185–285 ms after deviant onset—is thought to reflect involuntary reorienting of attention related to alerting processes in frontal lobe triggered by salient novel auditory features (Polich, 2007; Yamaguchi and Knight, 1991).

More recent research has shown that acoustically novel events can modulate the morphology of AEP components that occur at latencies preceding the MMN (Escera and Malmierca, 2014). For example, the middle-latency response (MLR), an early sequence of positive and negative deflections (∼12–50 ms post-stimulus onset) can be modulated by changes in acoustic frequency (Althen *et al.*, 2013; Grimm *et al.*, 2011; Leung *et al.*, 2012; Recasens *et al.*, 2014a,b; Slabu *et al.*, 2010), missing-fundamental tones (Alho *et al.*, 2012), intensity (Althen *et al.*, 2016), temporal regularity (Leung *et al.*, 2013) and location (Grimm *et al.*, 2012; Sonnadara *et al.*, 2006). The sub-components of the MLR have been attributed to input-level activation of primary auditory cortex (AC) by projections from the medial geniculate body (MGB) of the thalamus (McGee *et al.*, 1992).

Evidence of novelty detection has also been found at lower levels than AC. The frequency-following response (FFR) is a steady-state AEP component occurring 5–10 ms post stimulus onset that primarily reflects phasic population-level synaptic activity from nuclei within the midbrain, namely the inferior colliculus (IC). Moreover, this phasic activity entrains to periodicities in the evoking acoustic stimulus (Chandrasekaran and Kraus, 2010; Smith *et al.*, 1975, 1978; Sohmer *et al.*, 1977; Worden and Marsh, 1968). The wide variety of cell types in the IC are together capable of representing complex auditory signals with high temporal precision (Peruzzi *et al.*, 2000), at least up to 1000 Hz (Chandrasekaran and Kraus, 2010; Kraus and Nicol, 2005). Slabu *et al.* (2012) demonstrated that the FFR representation of the second formant of a speech syllable was modulated when evoked by a deviant compared to a standard context. Similarly, Shiga *et al.* (2015) found that amplitude-modulated frequency deviants concurrently modulated FFR and MLR responses in addition to evoking the MMN. Functional magnetic resonance imaging (fMRI) experiments further suggest that activity related to novelty-detection can be localized to the IC and the MGB (Cacciaglia *et al.*, 2015; Gao *et al.*, 2014).

Currently, it is unknown how distributed novelty detection processes along the auditory hierarchy are related. One possibility is that the deviant-related modulation of the FFR reflects stimulus-specific adaptation (SSA) at the IC that is transmitted to higher level auditory structures, such as AC (Fishman and Steinschneider, 2012; Nelken and Ulanovsky, 2007). This transmission might then lead to changes in MLR and elicitation of the MMN and possibly P3a. Alternatively, deviance-detection processes in AC might trigger modulation of activity in subcortical nuclei through direct corticofugal projections to the MGB and IC (Chandrasekaran *et al.*, 2014). Speaking

to the former possibility, single-unit recordings in anesthetized animals show that neurons in the AC (Ulanovsky *et al.*, 2003), as well as in the IC and MGB (Ayala *et al.*, 2015; Ayala and Malmierca, 2013, 2015; Ayala *et al.*, 2013; Lumani and Zhang, 2010; Patel *et al.*, 2012; Ponnath *et al.*, 2013) exhibit reduced firing after a few repetitions of a standard tone (i.e., SSA), but regain robust firing in response to a deviant tone. Further, SSA in AC neurons is similar to that of the IC and MGB in that the difference in neural firing evoked by standard and deviant stimulation is positively correlated with the degree of frequency separation between standards and deviants and negatively correlated with the deviant's probability of occurrence. The rapid rate of adaptation is yet another commonality of SSA in cortical and subcortical neurons (Antunes *et al.*, 2010; Ayala and Malmierca, 2013; Malmierca *et al.*, 2009; Ulanovsky *et al.*, 2003). Speaking to the latter possibility, strong and widespread levels of SSA in subcortical nuclei are largely limited to nonlemniscal divisons of the IC (Ayala *et al.*, 2013; Pérez-González *et al.*, 2005; Zhao *et al.*, 2011) and MGB (Antunes *et al.*, 2010), and it is well known that these regions are the target of dense efferent projections from the AC (Games and Winer, 1988; Malmierca and Ryugo, 2011; Winer and Schreiner, 2005; Winer *et al.*, 2001, 1998). Of course, it is likely that both processes operate concurrently as suggested by Escera and Malmierca (2014), wherein SSA in IC is transmitted through MGB to AC for further processing, but is at the same time modulated by descending cortical control. Such a system would allow for both bottom-up and top-down processing of novel acoustic events (see also, Chandrasekaran *et al.*, 2014; Escera and Corral, 2007).

Previous work has shown that it is possible to record the activity of cortical novelty-detection mechanisms concurrently with subcortical acoustic feature representation in individual listeners (Shiga *et al.*, 2015; Slugocki *et al.*, 2017). But to our knowledge no study has examined whether the timing or magnitude of cortical novelty-detection components is related to the modulation of components from lower level generators, nor has any study examined the timing of FFR modulation relative to the onset of the deviant signal. Here we use stimuli designed for the measurement of the simultaneously-evoked auditory potential (SEAP; Slugocki *et al.*, 2017) to assess novelty-detection processes at different levels of the auditory hierarchy. The SEAP stimuli are designed by amplitude modulating pure tone carriers at the complex sum of a low (∼40 Hz) and high (∼80 Hz) frequency that are not harmonically related. The stimuli are presented in a forward-reverse oddball paradigm wherein AEPs are collected in response to both a frequently occurring standard stimulus and an infrequent deviant stimulus that differs from the standard in both carrier and modulation frequencies. Stimuli are presented over two blocks with standard and deviant stimulus assignment reversed in the second block. This presentation scheme should elicit FFRs at their pure tone carriers, 40 (cortical/thalamocortical; Herdman *et al.*, 2002) and 80 Hz (subcortical; Frisina *et al.*, 1990; Herdman *et al.*, 2002; Suzuki, 2000) auditory steady-state responses (ASSRs) at the amplitude-modulation rates, as well as the obligatory auditory cortical potentials (P1-N1-P2). Additionally, the late novelty-detection components (MMN and P3a) should be evoked by the infrequent change in the pure tone carrier.

Using this method we can make within-subjects comparisons of the neural activity contributing to each aforementioned AEP component as elicited by the same physical

stimulus but in a deviant relative to standard context. Further, we can assess whether deviant-related modulation of subcortical components (i.e., the FFR) precedes or follows later cortical novelty-detection components (i.e., MMN and P3a), and whether the overall strength and timing of these components is related across listeners. If novelty-detection operates exclusively in a bottom-up manner then deviant-related modulation of the FFR should precede or coincide with the latency of MMN and P3a components. On the other hand, if novelty-detection signals are first generated in the AC and feedback to modulate subcortical processing in a top-down manner, then deviant-related modulation of the FFR should follow elicitation of the MMN and P3a. We do not expect deviant stimulation to affect the 80 Hz ASSR because SSA has not been observed below the level of the IC (Ayala *et al.*, 2013), though it may modulate the 40 Hz ASSR as this component likely shares generators with MLRs (Bohórquez and Özdamar, 2008; Galambos and Makeig, 1992; Plourde *et al.*, 1991).

4.3 Experiment 1

4.3.1 Materials and methods

Participants

A total of 20 adult subjects (15 females, mean age = 19.59, SD = 1.57 years) were recruited from the McMaster Undergraduate Psychology subject pool. Written consent was obtained in compliance with a protocol approved by the McMaster Research Ethics Board and participation was remunerated in the form of partial course credit.

Basic demographic information was obtained from each participant via a short questionnaire. All participants self-reported to have normal hearing at the time of experiment. The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* (18 July 1964).

Stimuli

Two amplitude-modulated pure tones were generated following the technique for collecting simultaneously-evoked auditory potentials (SEAP) as described in Slugocki *et al.* (2017) and as detailed in Eq. 4.1:

$$y(t) = \sin(2\pi f_{carrier} t) \times ([\sin(2\pi f_{AM1} t) + \sin(2\pi f_{AM2} t)] + 1) \tag{4.1}$$

The first stimulus, hereafter referred to as *SEAP300*, consisted of a 300 Hz carrier tone that was amplitude-modulated at the sum of 44 and 77 Hz. The second stimulus, hereafter referred to as *SEAP700*, consisted of a 700 Hz carrier tone that was amplitude-modulated at the sum of 37 and 81 Hz. The carriers were modulated at different rates so that they differed in both spectral and temporal dimensions. Both carriers were amplitude modulated to a depth of 100%. Modulation rates were chosen so as to be non-harmonically related but close to the modulation rates evoking maximum amplitude ASSRs (i.e., 40 and 80 Hz) without creating additional modulation of the carrier signal at $f2 - f1$ and allowing for later response extraction and analysis in the frequency domain. Hereafter, we collectively refer to the ASSRs evoked by rates of 37/44 Hz as 40 Hz ASSR and those evoked by 77/81 Hz as 80 Hz ASSR in keeping with the established distinction in the literature. The resultant stimuli contained spectral energy at the carrier frequencies as well as the sum and difference of

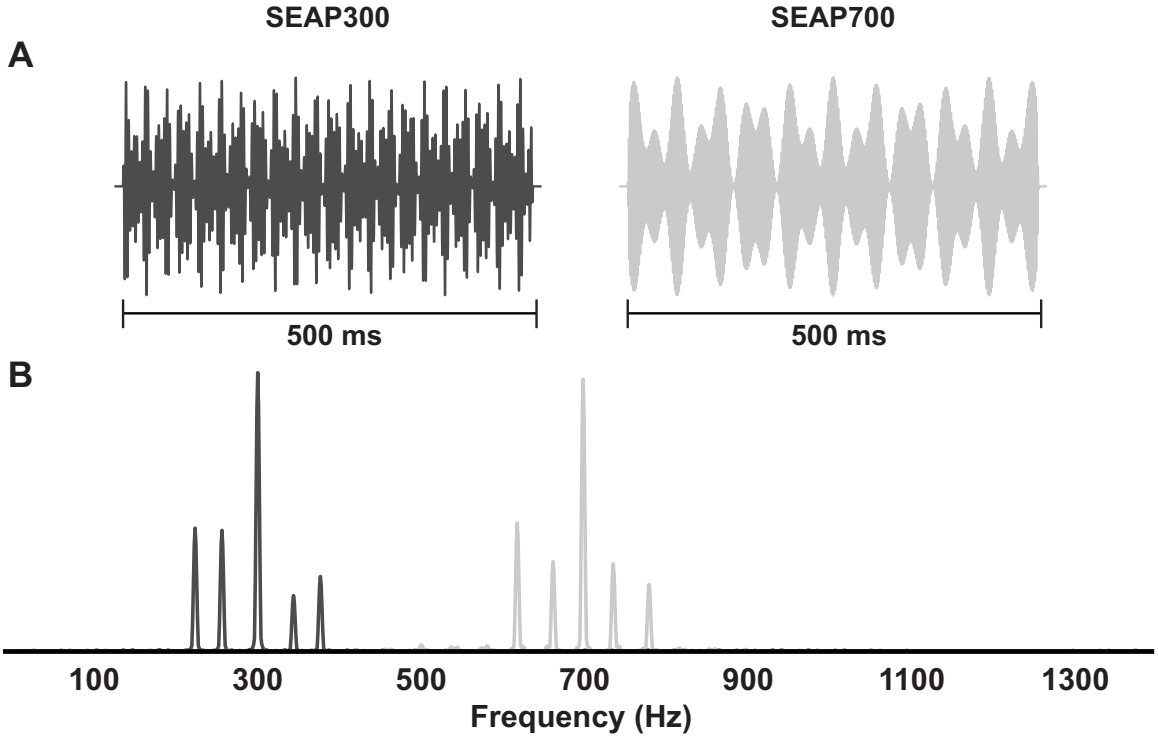

FIGURE 4.1: *SEAP300* and *SEAP700* stimuli used in Experiment 1 as represented in both the time (A) and frequency (B) domains. The 300 Hz (left/black) and 700 Hz (right/grey) carrier signals were amplitude modulated at the sum of 44+77 Hz and 37+81 Hz, respectively. Figures represent the output signal recorded from the foam ear-tip using a Bruël & Kjær Artificial Ear (Type 4152) connected to a Bruël & Kjær 2260 Investigator sound level meter.

the carrier and modulation frequencies but no energy at the modulation frequencies themselves (FIG. 4.1).

Stimuli were generated online using a Tucker-Davis Technologies (TDT) RP2.1 Enhanced Real-time Processor controlled by TDT RPvdsEX (v.5.4) software. Signal output was attenuated with a TDT P5A Programmable Attenuator, whose output was then passed through a Hafler Transana P1000 Amplifier which drove a compression driver (Peavey RX-22 8Ω). To prevent electromagnetic artifact in the EEG

data, the compression driver was kept outside of the subject testing chamber. Sound output from the compression driver travelled through a length of polyvinyl chloride tubing (282 cm of tube with 1.27 cm outside diameter) that tapered over 12.5 cm to accommodate another 250 cm of tubing (0.32 cm outside diameter). The end of the narrow tube terminated in a 13 mm disposable ear-insert (ER1-14A). Participants inserted the foam ear tip into their right ear canal for stimulus delivery. The output of the presentation system was measured at the foam ear-insert using a Bruël & Kjær Artificial Ear (Type 4152) connected to a Bruël & Kjær 2260 Investigator sound level meter. A digital room correction algorithm (http://drc-fir.sourceforge.net) was used to correct for the distorted frequency response characteristics of the compression driver assembly. This algorithm produced finite impulse response (FIR) filters to ensure a flat frequency response over a 200 Hz bandwidth centered on each carrier frequency. Sound output was calibrated to ensure an equivalent C-weighted sound pressure level (SPL-C) for both stimuli of 80 dB SPL-C. Analysis of the impulse response generated by the filtered output revealed nearly uniform energy ($\pm$3 dB) in the frequency domain from 200 to 400 Hz and from 600 to 800 Hz.

Recordings of each stimulus were taken at the tip of the ear insert using the B&K apparatus, described above. The AC level output of the sound level meter was passed to a bipolar channel on the Neuroscan SynAmps RT headbox. A stimulus delay of 20.45 ms, attributed to the tube assembly and digital filters, was measured as the time between the transistor-transistor logic pulse in the EEG trace and the time of arrival for sound at the ear insert. All subsequent measures of component latency, as well as markers of stimulus onset, were adjusted to reflect this delay.

Presentation and Recording

The experiment comprised two recording blocks each consisting of 2400 trials. A trial was defined as the presentation of a 500 ms stimulus followed by a 500 ms inter-stimulus interval. In the first block, presentation followed an oddball paradigm, wherein one stimulus (e.g., *SEAP300*) served as the frequent standard (85% of trials) and the other (e.g., *SEAP700*) served as the infrequent deviant (15% of trials). Presentation of deviant trials was pseudo-random with the constraint that successive deviants were separated by at least two standard trials (Picton *et al.*, 2000). In the second block, the assignment of standard and deviant trials was reversed, such that the stimulus serving as the standard in block 1 was the deviant in block 2, and *vice versa*. The order of blocks was counterbalanced across participants. Each block lasted approximately 40 minutes and the total experimental session lasted less than 2 hours. This forward-reverse presentation paradigm allowed for comparison of AEPs as elicited in standard and deviant contexts while controlling for acoustic differences between the stimuli.

During stimulation participants sat on a comfortable chair located inside a sound-attenuating chamber. To maintain participant arousal, a silent subtitled movie was displayed on a 48 cm LCD computer monitor placed at eye-level on a desk positioned 1 m in front of the participant's chair. Subject EEG was recorded using three Ag/AgCl sintered electrodes filled with a conductive gel (Signa Gel). Electrodes were applied in a single-channel vertical montage with the active electrode placed on the vertex of the skull (Cz), the reference electrode placed on the right ear-lobe (ipsilateral to stimulus presentation), and the ground electrode placed on the forehead. Reference and ground electrodes were secured using double-sided tape washers. The vertex electrode was

kept in place with an elastic headband stretched under the participant's chin. Scalp voltage measured from the electrodes was sampled and digitized at 20 kHz using a Neuroscan SymAmps RT amplifier and stored using Compumedics SCAN 4.5 Acquire software. Electrode sites were prepared using NuPrep abrasive gel in an attempt to keep electrode impedances below 5 kΩs, though data were only rejected if electrode impedance was above 30 kΩs (mean = 13.50, SD = 8.13 kΩs).

Data Analysis

Continuous EEG data were processed in MATLAB (v7.10.0) using the EEGLAB (v13; Delorme and Makeig, 2004) toolbox and the ERPLAB (v4.0) (v4.0; Lopez-Calderon and Luck, 2014) plug-in. Recorded data were first re-sampled at 2000 Hz to reduce computation time and memory demands. Data were then zero-phase filtered using a Chebyshev Type II infinite impulse response (IIR) high-pass filter at 0.5 Hz (Table 4.1) to remove slow-wave activity without distorting signal phase. Trials were epoched between −120 and 880 ms relative to stimulus onset. Baseline correction was applied to each epoch by subtracting the average activity from −120 to 0 ms (relative to stimulus onset) from the entire epoch. Artifact trials, identified as those containing activity exceeding ±80 μV, were eliminated from each participant's dataset.

Presence of Relevant Frequency Components

A complete set of artifact-free standard trials was extracted and averaged for each subject to verify significant activity at relevant component frequencies using the F-test method (Zurek, 1992). In this method, discrete Fourier transforms (DFTs) of the standard-evoked sustained AEPs (i.e., low and high ASSRs, and the FFRs) were

Table 4.1: Summary of Chebyshev Type II IIR filters.

	Frequency (Hz)			**Attenuation (dB)**		
AEP	*Low-stop*	*Pass*	*High-stop*	*Low-stop*	*Pass*	*High-stop*
Transients (HP)	0.25	0.5	–	−12	1	–
Transients (LP)	–	20	40	–	1	−24
37 Hz ASSR	13	26–48	96	−24	1	−24
44 Hz ASSR	17	34–54	108	−24	1	−24
77 Hz ASSR	33	66–86	172	−24	1	−24
81 Hz ASSR	35	70–90	180	−24	1	−24
300 Hz ASSR	140	280–320	640	−24	1	−24
700 Hz ASSR	340	680–720	1000	−24	1	−24

calculated over a Hanning-windowed portion of the averaged epoch ranging from 50 to 450 ms post-stimulus onset. A spectral resolution of 1 Hz was interpolated by zero-padding the extracted portion of the response epoch up to 2000 samples. Amplitude values corresponding to the frequency of each component were tested against 40 neighboring non-component frequency bins (20 higher and 20 lower). Significance of the F ratio was evaluated against critical values of F with 2 and 80 degrees of freedom. One subject was excluded from further analyses for failing to pass F-test verification at any component frequency. Therefore, the final sample consisted of 19 subjects (mean age = 19.63, SD = 1.61 years; mean impedance = 12.63, SD = 7.33 kΩ). For all further analyses, standard trial selection was limited to those trials that immediately preceded deviant presentations to minimize any differences in signal-to-noise ratio (SNR) resulting from unequal trial counts.

Transient AEPs

To assess transient AEPs, standard- and deviant-evoked averages from all subjects were forward-reverse low-pass filtered at 20 Hz using a Chebyshev Type II IIR filter (Table 4.1). Filtering was applied in the forward and reverse direction to ensure zero phase distortion. This filter was used to remove higher frequency energy contributed from ASSR components. Difference waves were calculated by subtracting the filtered standard response from the filtered deviant response to the same carrier signal (Picton *et al.*, 2000). The peaks of all transient components were selected using a peak picking algorithm in MATLAB and confirmed with visual inspection (Mathworks File Exchange: pickpeaks.m). Grand averages of the standard, deviant, and difference AEP traces were computed by averaging each response type across all participants. The peaks of the P1, N1, and P2 components were identified in each grand-averaged standard waveform as the largest positive-negative-positive deflection in a window from 10–300 ms post-stimulus onset. Peaks in individual subject standard averages were then identified in a 50 ms window around the grand-averaged peaks. The N1 and P2 components are often considered as a complex thought to reflect synchronous activation of thalmocortical circuits (Näätänen and Picton, 1987; Wolpaw and Penry, 1975; Woods, 1995), hence we included the difference between P2 and N1 amplitudes in our assessment. The grand mean of all averaged difference waves was tested against zero at 1 ms intervals from 0–500 ms post-stimulus onset using a single sample t-test (two-tailed, $\alpha = 0.05$, Bonferroni corrected). The peak of the MMN of each subject was identified as the largest negative peak in the difference wave within a window where the single sample t-test of the grand-averaged difference indicated a significant negative deviation from zero. Similarly, the peak of the P3a component of each

subject was identified as the largest positive peak within a window where the single sample t-test of the grand-average difference indicated a significant positive deviation from zero. Amplitude and latency values of the MMN and P3a peaks from each subject were recorded for later analyses. The equivalence of the 300 and 700 Hz stimuli in eliciting transient components was assessed using paired samples t-tests (α = 0.05, Bonferroni-corrected) under the null hypothesis that the difference between the amplitudes and latencies of each component elicited by *SEAP300* versus *SEAP700* was equal to zero.

Steady-state AEPs: 40 Hz, 80 Hz, and FFR

To assess steady-state AEPs, the DFT of each subject's averaged epoch was calculated over a window from 50 to 450 ms post-stimulus onset following the method described above. Amplitude values corresponding to 3 Hz bins around the frequency of each component were extracted for later comparisons. Use of 3 Hz bins allowed for inclusion of any small deviations in phase locking at the relevant spectral components. For each steady-state component, paired samples t-tests assessed the null hypothesis that the difference between the standard and deviant response amplitude was zero. Unwrapped phase values were also extracted for all steady-state components. Phase data were assessed using CircStat, a circular statistics toolbox for MATLAB (Berens, 2009). First, the Rayleigh test measured the degree to which phase data polarized for each steady-state component in standard- and deviant-evoked data across our sample of listeners (Fisher, 1995). Significant polarization was interpreted as further evidence that oscillatory activity in the AEP was phase locked to the periodicity of the evoking stimulus. Second, the Watson-Williams test assessed the null hypothesis that the

phases of the standard and deviant steady-state components shared the same direction (Stephens, 1969; Watson and Williams, 1956). To better visualize the data in the time domain, standard- and deviant-evoked average responses were band-pass filtered using Chebyshev Type II IIR filters (Table 4.1). Filtering was applied in the forward and reverse direction to ensure zero phase distortion.

Time-frequency Decomposition

The effect of stimulus context (i.e., standard *versus* deviant) on the subcortical stimulus representation, as reflected in the FFR, was further investigated in time-frequency decompositions of each subject's averaged response. First, subject-specific averages were segmented into 880 windows. Each window was 120 ms long and shared 119 ms of overlap with the previous window. This window length could accommodate just over 4 cycles of the lowest oscillatory component (i.e., 37 Hz). Slices were then Hanning-windowed and zero-padded (to equal 2000 samples) before computing the DFT. The resultant DFT coefficients of each slice were attributed to the latency at the center of each window. This process generated unique DFT coefficients for each millisecond of data between −60 and 820 ms relative to stimulus onset.

Second, complex time-frequency representations (TFRs) from each participant were grand averaged separately for each context (standard or deviant) and stimulus (*SEAP300* or *SEAP700*) combination. A T-mean statistic (Martínez-Montes *et al.*, 2008) compared each complex time-frequency point in the grand average of each condition against the average baseline activity (−120 to 0 ms re stimulus onset) in the corresponding frequency bin. The T-mean value was evaluated using an F-distribution with 2 and 2(N-1) degrees of freedom, where N represents the number of

time points belonging to the baseline period. Alpha (0.05) was Bonferroni corrected for the number of comparisons in the time-frequency data according to Eq. 4.2.

$$\alpha/(N_{bins} \times N_{windows} \times N_{subjects}) \tag{4.2}$$

In this way, time-frequency point masks of significant activity were created to limit the number of subsequent comparisons and control the false alarm rate without drastically reducing power. Masks for each stimulus comprised those time-frequency points that were significantly different from baseline in either a standard or deviant context.

Third, the TFRs of steady-state components were examined separately for each carrier using within-subjects nonparametric permutation tests (described in Maris and Oostenveld, 2007). In this procedure, statistical values for each time-frequency point that survived the masking procedure described above are obtained from a paired samples t-test comparing standard- and deviant-evoked activity across participants. A permutation distribution is then generated by measuring paired samples statistics for 1000 permutations of the subject-specific averages (i.e., assignment of standard and deviant). For n subjects, this process draws on a distribution of 2^n possible t-values that are all equally probable under the null hypothesis. The p-value is obtained by locating the observed test statistic under the permutation distribution. P-values were considered significant if they exceeded an alpha of 0.05 (Bonferroni corrected for the number of time-frequency points surviving baseline comparison, as defined above). Clusters of significant p-values were defined by the number of time-frequency points in any contiguous grouping. Only the largest cluster was considered as a possible indicator of a significant context-related difference in steady-state activity.

Correlations between Steady-state and Transient AEPs

To control for the family-wise error rate while maintaining statistical power, only the amplitudes or phases of those steady-state components exhibiting sensitivity to trial type (standard vs. deviant) were compared against the amplitude and latency of the cortical change-detection components (i.e., MMN and P3a) using Pearson's correlation analysis. The actual sample used in each correlation consisted of some subset of the 19 subjects whose data also survived outlier elimination in each morphological category. Outliers were defined as values that fell more than 1.5 times the interquartile range above the third quartile or below the first quartile.

4.3.2 Results

Transient AEPs

The grand-averaged activity evoked by standard presentation of both *SEAP300* and *SEAP700* revealed obligatory cortical P1-N1-P2 components. Similarly, roving window t-tests revealed two significant deflections in the difference waves (*deviant* − *standard*) computed for *SEAP300* and *SEAP700* that coincided with the expected latencies of the MMN and P3a components (Fig. 4.2).

The peak-to-peak amplitude of the standard-evoked N1-P2 complex was significantly greater when elicited by a *SEAP700* relative to *SEAP300* ($t(15) = 3.31, p < 0.01$). Though not significant, the latencies and amplitudes of individual standard-evoked obligatory components (i.e., P1, N1, and P2) trended towards being stronger and occurring earlier when elicited by *SEAP700*. Analysis of the difference waves further revealed stronger ($t(17) = -4.64, p < 0.001$) and earlier ($t(17) = -4.65, p < 0.001$) MMN and significantly earlier P3a ($t(17) = -4.57, p < 0.001$) in response to

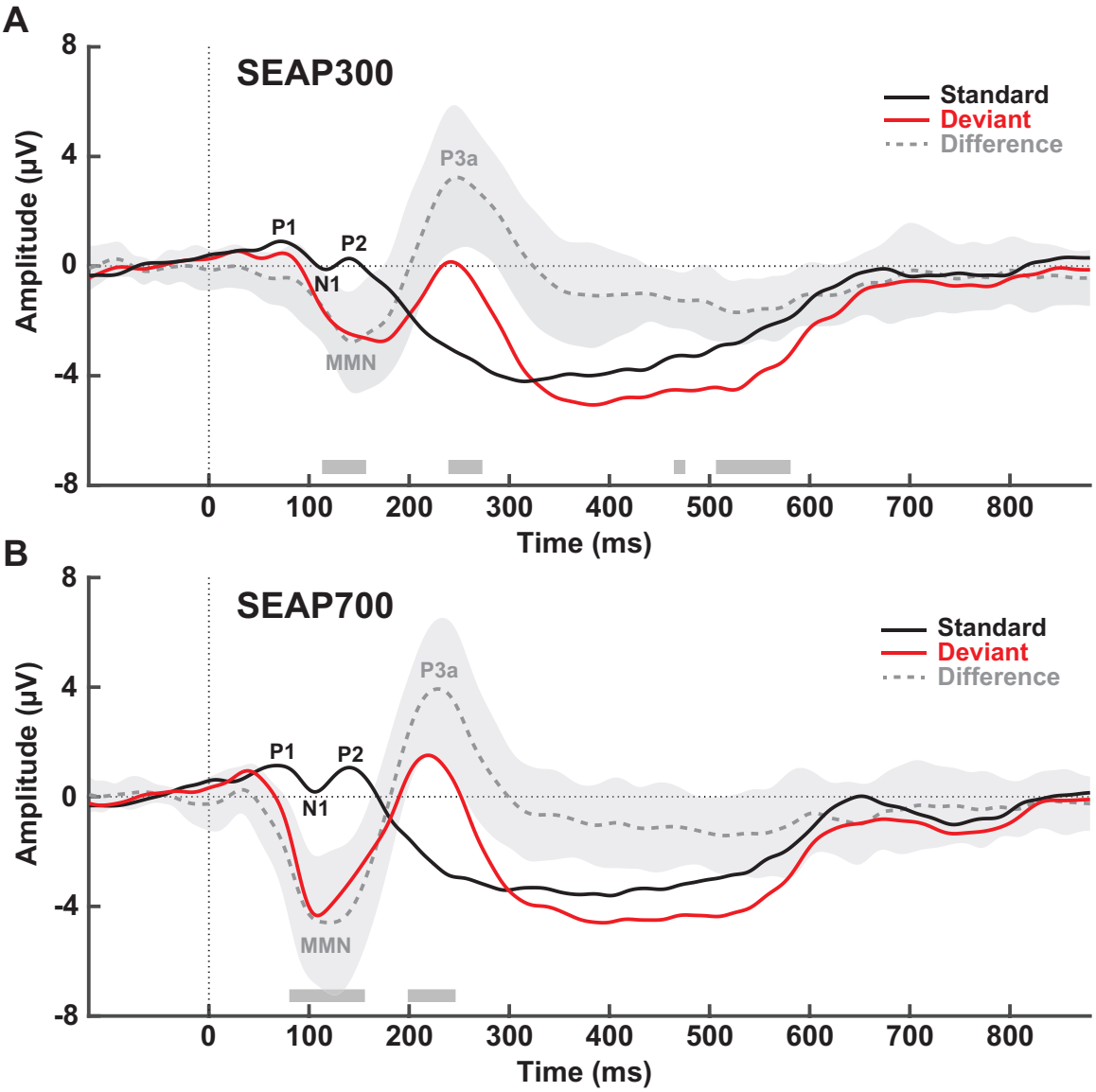

FIGURE 4.2: Grand averaged (n=19) transient AEPs elicited in response to *SEAP300* (A) and *SEAP700* (B) stimuli when presented as standards (black) or deviants (red) in the forward-reverse oddball paradigm. The grand average of the difference wave (standard response subtracted from deviant response) and the standard deviation of the difference wave are represented in the dashed trace and grey area, respectively. Labels identify transient component peaks in the grand averaged waves. Dark grey bars above abscissa indicate significant deflections from zero (single sample t-test at each time point, two-tailed, $\alpha = 0.05$, Bonferroni corrected).

SEAP700 relative to *SEAP300* deviants (Fig. 4.3). The stimulus-specificity of these effects is explored in Experiment 2.

Steady-state AEPs and Time-frequency decomposition

The grand-averaged time and frequency domain representations of standard- and deviant-evoked steady-state activity can be seen in Fig. 4.4 and Fig. 4.5.

FFR

Paired samples t-tests (Fig. 4.6) revealed that the amplitude of the 700 Hz FFR was enhanced when elicited by a deviant relative to standard trial ($t(18) = -2.12, p < 0.05$). The results of the within-subjects non-parametric permutation test further showed that the greatest period of enhancement in the 700 Hz FFR signal occurred following the transient change detection components (Fig. 4.7), suggesting that this effect was related to a cortical-subcortical feedback process. No such effects were observed for the 300 Hz FFR in either the average windowed FFT (i.e., 50-450 ms post stimulus onset) or in the time-frequency decomposition (Fig. 4.8).

Rayleigh tests of the steady-state components indicated significant polarization of the 300 Hz FFR, but not the 700 Hz FFR (Fig. 4.9). This negative result may have occurred because the period (1.4 ms) of the 700 Hz component is far shorter than the variation typically reported for the onset of the FFR component (Skoe and Kraus, 2010). The Watson-William test revealed that the phases of neither the 300 nor 700 Hz FFRs were affected by trial type (Fig. 4.9).

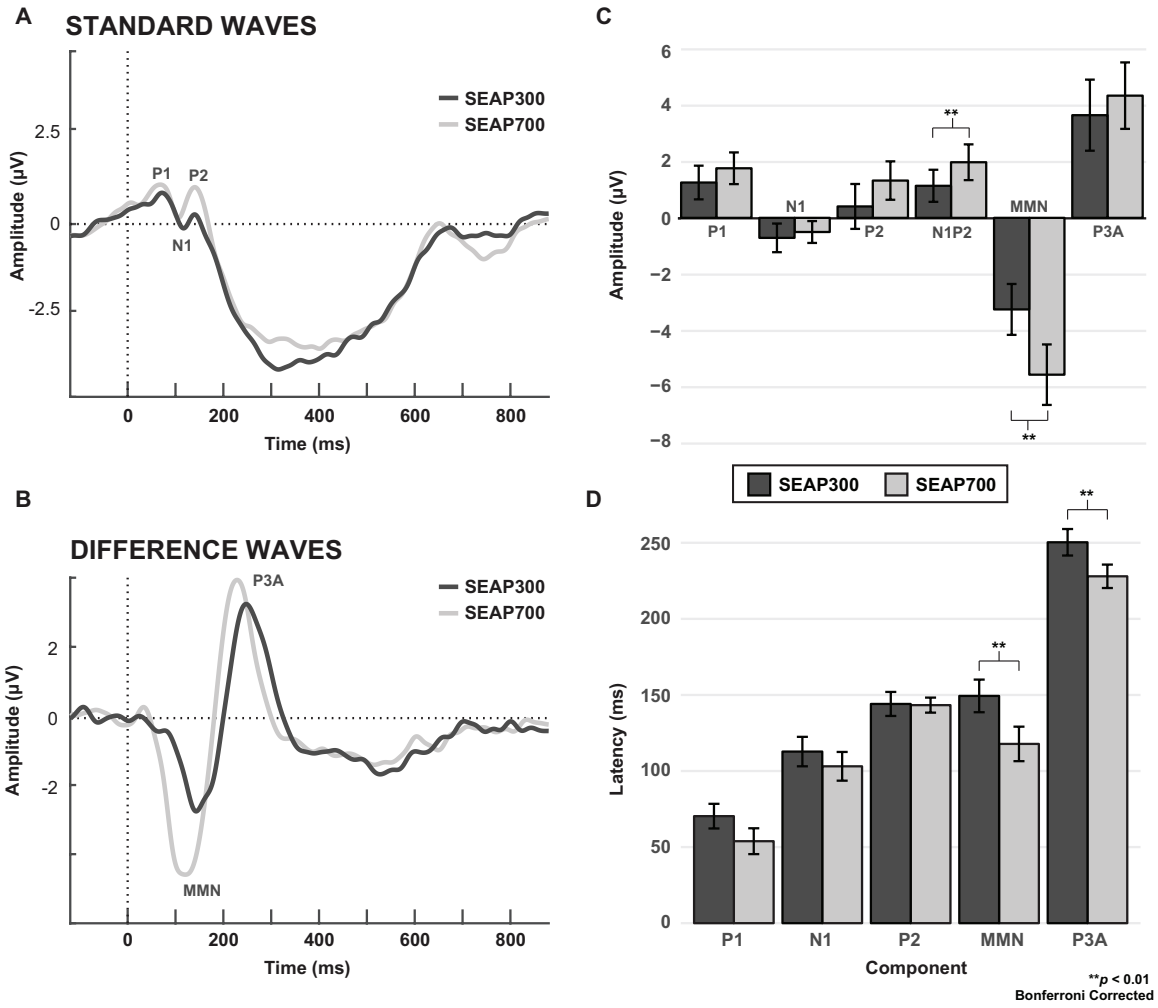

FIGURE 4.3: A time-domain comparison of the transient standard (A) and difference (B) components elicited by *SEAP300* (dark grey) and *SEAP700*(light grey) carrier signals. The amplitudes (C) and latencies (D) of all components labelled in the time-domain figures are plotted in columns with error bars representing the within-subjects 95% confidence interval of the mean. Significant differences between components based on eliciting carrier signal are denoted by asterisks.

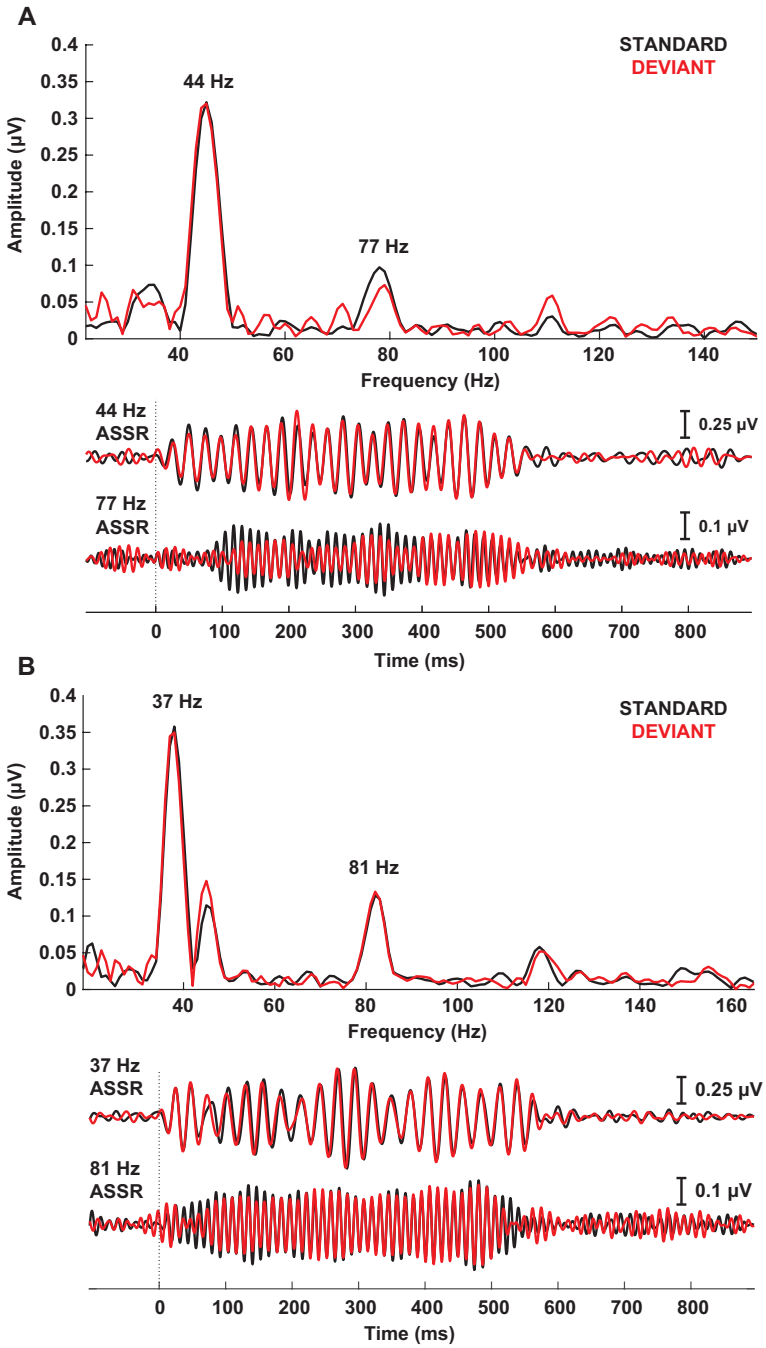

FIGURE 4.4: Grand averaged (n=19) ASSRs as elicited by *SEAP300* (A) and *SEAP700* (B) via standard (black) and deviant (red) trials. Each figure shows the data in the frequency (top) and time (bottom) domains. Labels identify relevant 40 and 80 Hz ASSR components in the frequency domain. To improve visualization, the data were bandpass filtered around each component prior to plotting in the time domain.

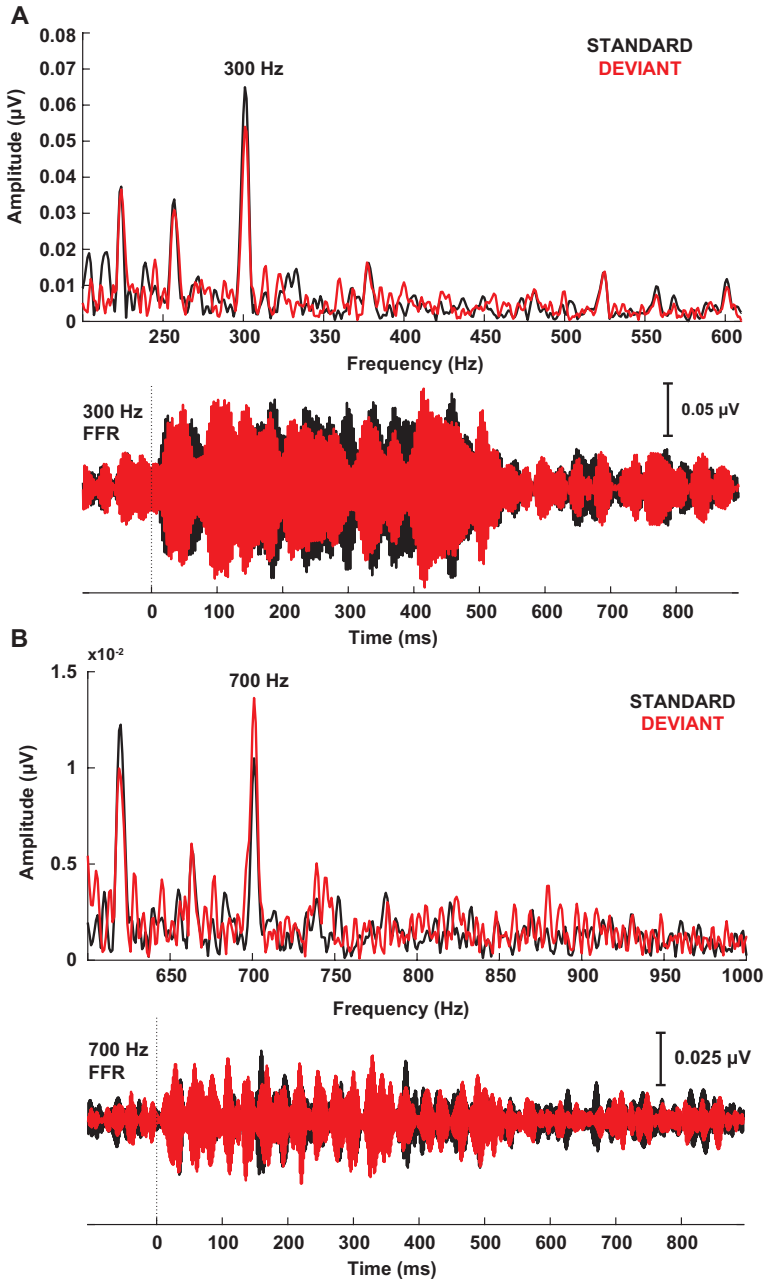

FIGURE 4.5: Grand averaged (n=19) FFRs as elicited by *SEAP300* (A) and *SEAP700* (B) via standard (black) and deviant (red) trials. Each figure shows the data in the frequency (top) and time (bottom) domains. Labels identify relevant FFR components in the frequency domain. To improve visualization, the data were bandpass filtered around each FFR component prior to plotting in the time domain.

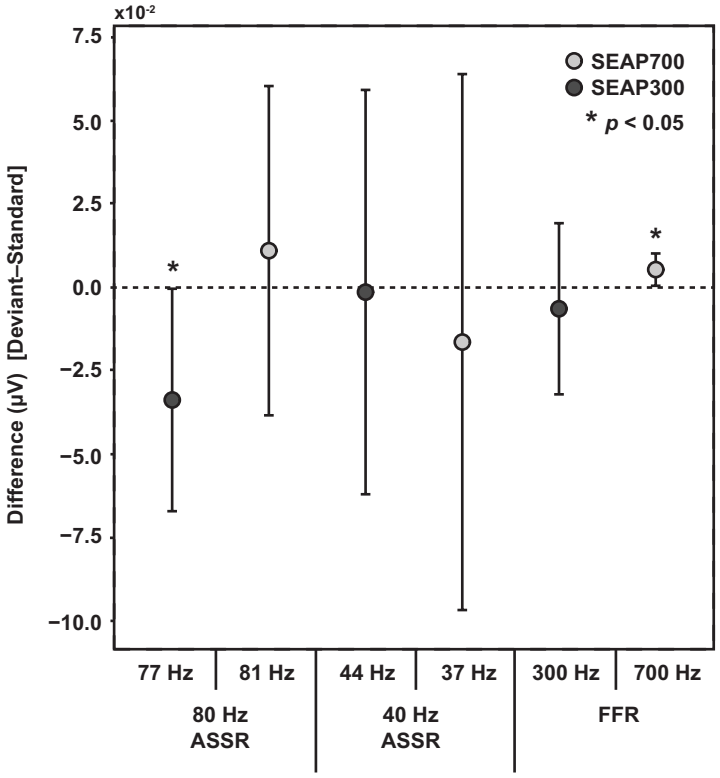

FIGURE 4.6: The average difference in amplitude (standard minus deviant) and the within-subjects 95% confidence interval of the difference are plotted for each steady-state component elicited by *SEAP300* (dark grey) and *SEAP700* (light grey) stimuli. Paired samples t-tests revealed that the 700 Hz FFR signal in the *SEAP700* stimulus and the 80 Hz ASSR (77 AM) signal in the *SEAP300* stimulus were significantly stronger and weaker, respectively, when elicited by deviant relative to standard trials.

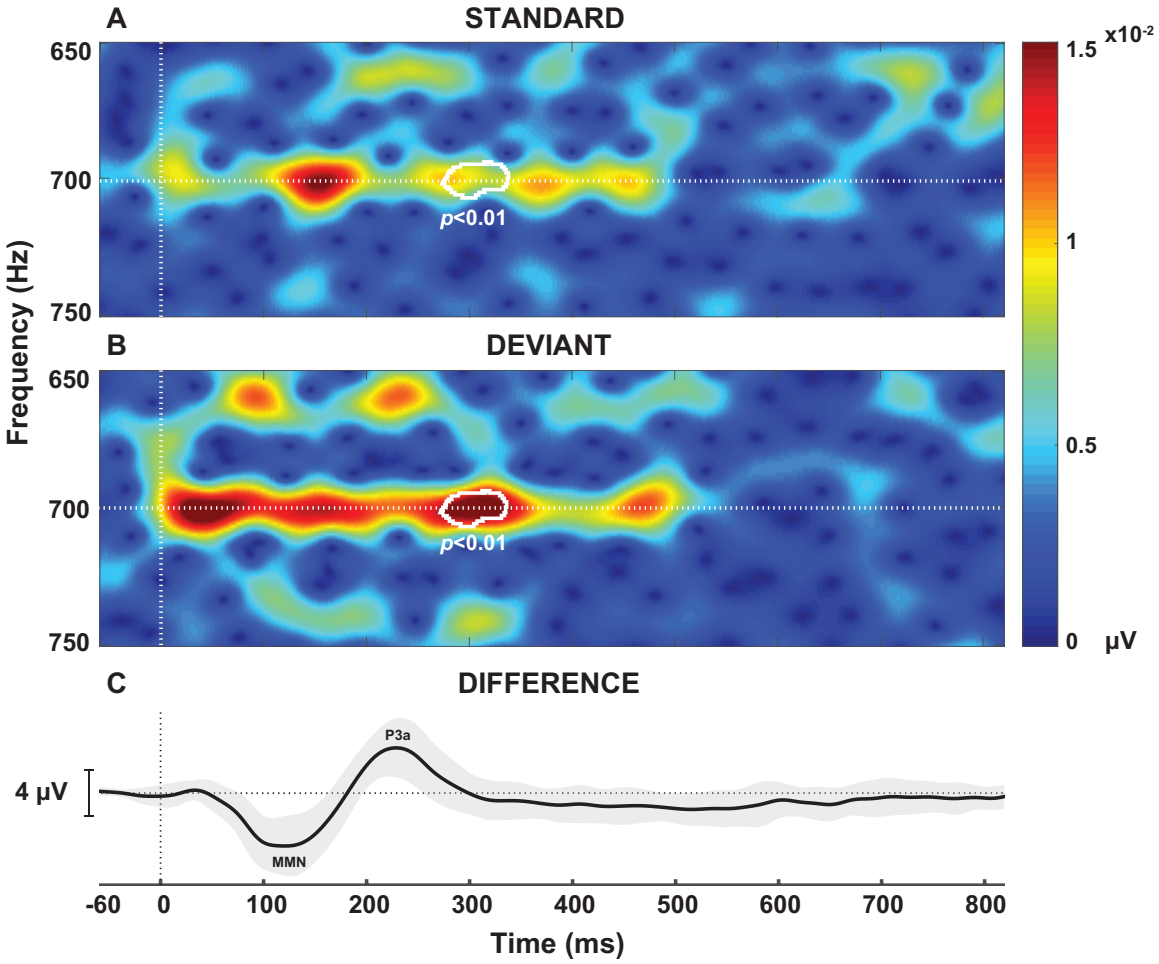

FIGURE 4.7: Time-frequency representations (TFRs) of the grand average standard (A) and deviant (B) responses to *SEAP700* are shown centered upon the 700 Hz component. The grand average of the difference wave, filtered around 0.5–20 Hz to visualize the cortical change detection components, is shown in (C). The vertical dotted lines denote stimulus onset in all three figures. The significant cluster identified by a within-subjects non-parametric permutation test is encircled by a solid white line.

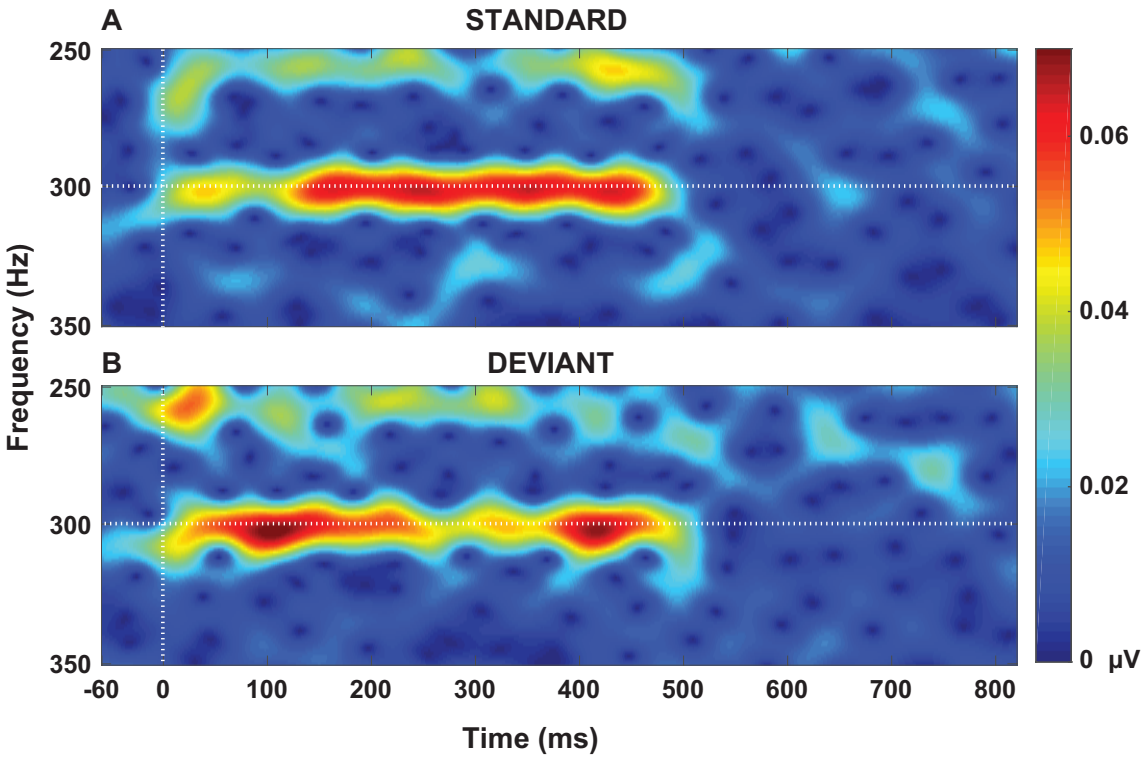

FIGURE 4.8: Time-frequency representations (TFRs) of the grand average standard (A) and deviant (B) responses to *SEAP300* are shown centered upon the 300 Hz component. The vertical dotted white lines denote stimulus onset in both figures. Within-subjects non-parametric permutation tests did not reveal significant differences in the FFR at 300 Hz when elicited by deviant versus standard presentations.

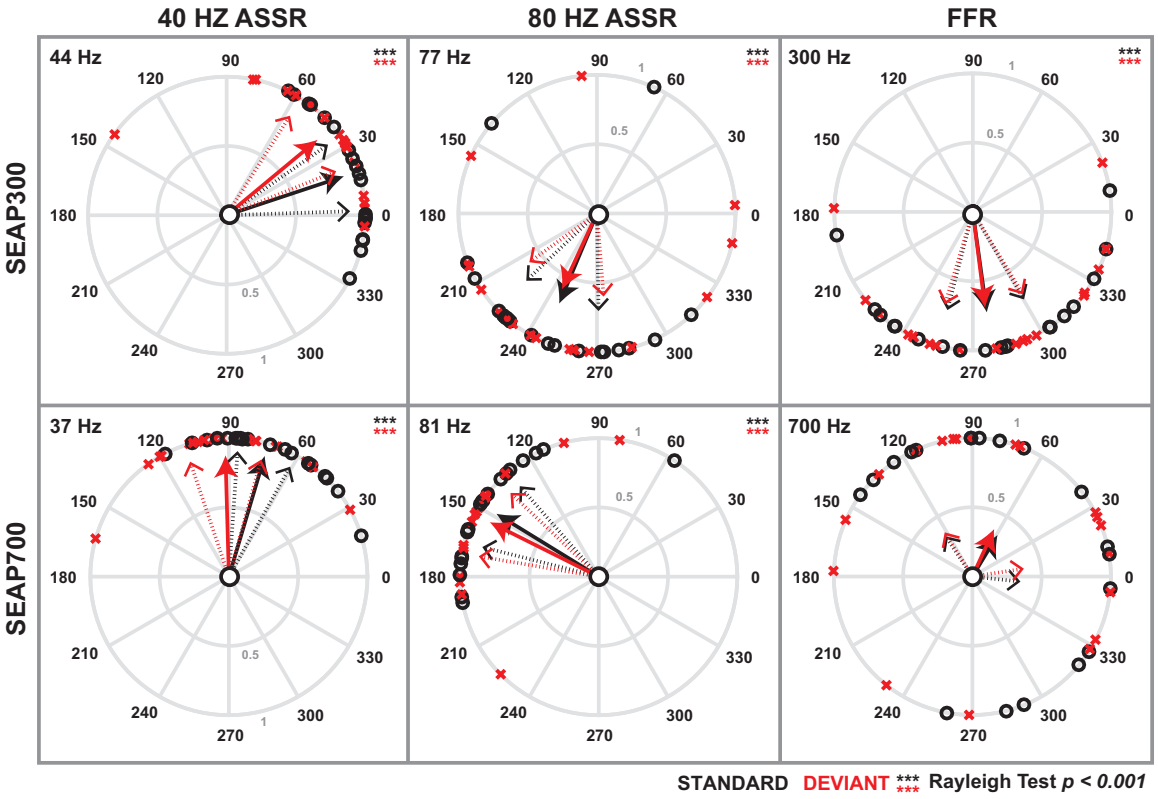

FIGURE 4.9: Polar plots of the 80 Hz ASSR (left), 40 Hz ASSR (middle), and FFR (right) as elicited by *SEAP300* (top) and *SEAP700* (bottom) stimuli. Components evoked by standard and deviant trials are plotted in black and red, respectively. Solid arrows denote the mean resultant vector for each trial type with 95% confidence intervals plotted by broken arrows. The phases of the 40 Hz ASSRs were significantly advanced when elicited by a deviant relative to standard context. Asterisks denote significant polarization as measured by the Rayleigh test.

40 Hz and 80 Hz ASSRs

The amplitude of the 80 Hz ASSR elicited the by 77 Hz AM (*SEAP300*) was significantly attenuated in response to deviant *versus* standard events ($t(18) = -2.33, p < 0.05$). The results of a within-subjects non-parametric permutation test further showed that the greatest difference between standard- and deviant-evoked 80 Hz ASSR occurred near onset of the response (Fig. 4.10). This might suggest that SEAP300 deviants disrupted the onset of phase-locking at neurons contributing to the 80 Hz ASSR. A similar effect was not observed for the 80 Hz ASSR elicited by 81 Hz AM in *SEAP700*.

Rayleigh tests of the steady-state components indicated significant polarization of all 40 Hz and 80 Hz components. Pertaining to the effect of trial type, the Watson-William test revealed that for both *SEAP300* and *SEAP700*, deviant trials significantly advanced the phase of the 40 Hz ASSRs (44 Hz AM, *SEAP300*: $F(1, 36) = -4.61, p < 0.05$; 37 Hz AM, *SEAP700*: $F(1, 36) = 4.01, p = 0.05$). The phase of the 80 Hz ASSRs were not affected by trial type (Fig. 4.9).

Correlations between Transient and Steady-state AEPs

Transient AEPs and FFR

The amplitudes of both the standard- and deviant-evoked 700 Hz FFR components were found to significantly correlate with the latency of the MMN, though the association appeared to be stronger for the deviant-evoked FFR. The amplitude of the deviant-evoked, but not the standard-evoked, FFR was also correlated with the latency of the P3a component (Fig. 4.11A). Moreover, the relationship between MMN latency and deviant FFR amplitude appeared to be particularly well fit by a linear

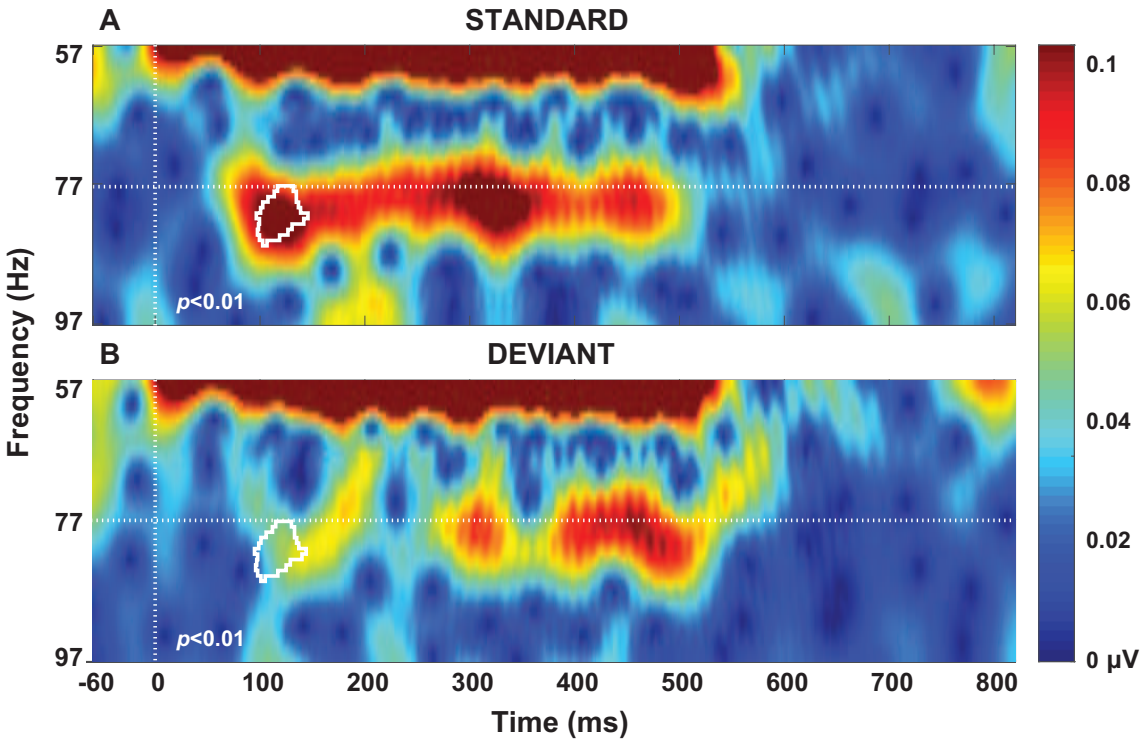

FIGURE 4.10: Time-frequency representations (TFRs) of the grand average standard (A) and deviant (B) responses to *SEAP300* are shown centered on the 77 Hz component. The vertical dotted white lines denote stimulus onset in both figures. The significant cluster identified by a within-subjects non-parametric permutation test is encircled by a solid white line.

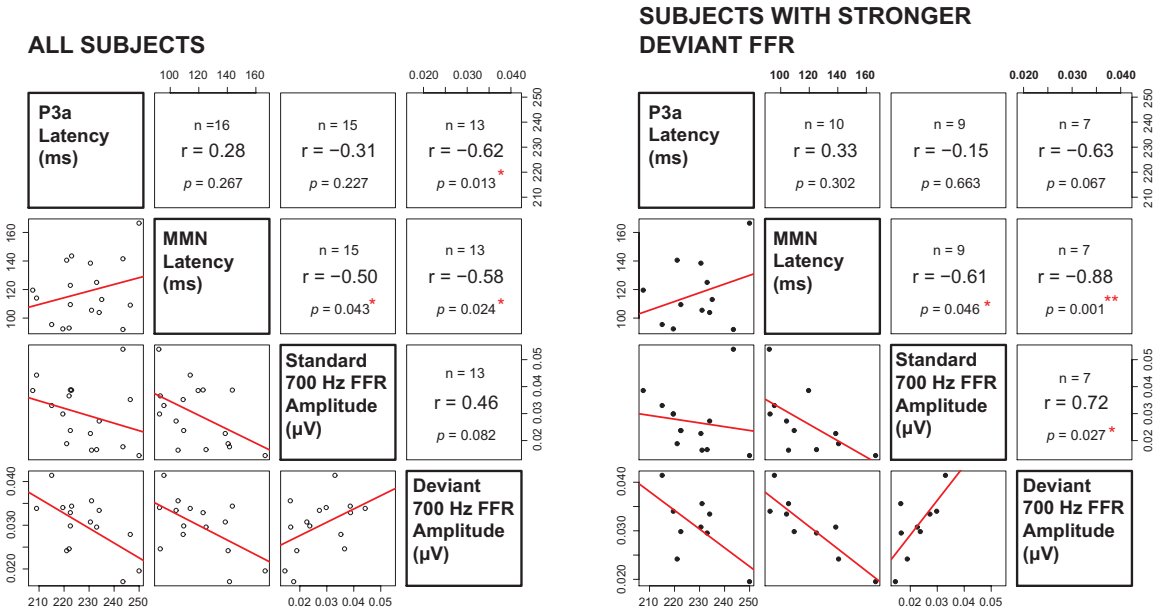

FIGURE 4.11: Scatterplot matrices showing relationships between the amplitudes of the standard- and deviant-evoked 700 Hz FFR components and the latencies of the MMN and P3a change-detection components. Plots are shown for correlations including all subjects surviving outlier elimination (A) and for that subset of subjects who showed a stronger deviant- relative to standard-evoked FFR (B).

regression line in the subset of subjects whose FFR amplitudes increased in response to deviant *versus* standard stimulus presentation (Fig. 4.11B). Similar associations were not observed between the amplitudes of either MMN or P3a components and that of the 700 Hz FFR.

As described in Experiment 1, correlation analyses involving the transient cortical novelty-detection components were limited to include only those steady-state components exhibiting context-dependent modulation. As such, we did not examine possible relationships between the 300 HZ FFR amplitude and the morphology of MMN or P3a.

40 and 80 Hz ASSRs and AEPs

For *SEAP300*, no significant association was found between 80 Hz ASSR (i.e., 77 Hz AM) amplitude and the morphology of change-detection components. The phase of both 40 Hz ASSRs also did not correlate with either the amplitudes of latencies of the MMN and P3a components. No further correlations were examined.

4.4 Experiment 2

4.4.1 Materials and methods

General

Analysis of the transient components revealed significantly stronger and earlier MMN and significantly earlier P3a when elicited by 700 Hz relative to 300 Hz deviant events. Enhancement of MMN amplitude has been previously reported for frequency increments relative to decrements (Peter *et al.*, 2010). In Experiment 2 we examined

the possibility that deviance direction (i.e., increment or decrement) was responsible for the enhancement of novelty-detection components observed in Experiment 1. The experiment comprised 4 different conditions which, in addition to testing for an effect of deviance direction, were also compared to assess the possible role of perceived loudness (*P.Loudness*), amplitude modulation (*PureTone*), and deviance size at high (*SmallD.700*) and low (*SmallD.300*) frequencies.

Participants

A total of 6 adult subjects (3 females; mean age = 26.71, SD = 2.55 years) were recruited from the graduate student community at McMaster University's Department of Psychology, Neuroscience & Behaviour. Written consent was obtained in compliance with a protocol approved by the McMaster Research Ethics Board. Basic demographic information was obtained from each participant via a short questionnaire. All participants self-reported to have normal hearing at the time of experiment. The study conformed to the Code of Ethics of the World Medical Association (Declaration of Helsinki), printed in the *British Medical Journal* (18 July 1964).

Stimuli

Two conditions (*P.Loudness* and *PureTone*) used the same large deviance size as Experiment 1. The *P.Loudness* condition further assessed whether stimulus-specific enhancement favoring the 700 Hz carrier signal was still observed when *SEAP300* and *SEAP700* stimuli were equated for perceived loudness rather than SPL as in Experiment 1. Stimulus parameters were identical to those of Experiment 1 except that stimulus loudness was equated according to the equal-loudness contours described in

Glasberg and Moore (2006). To be presented at equal loudness levels, *SEAP300* and *SEAP700* were presented at intensity levels of 80 and 76 dB SPL-C, respectively, or an equated loudness of 62.5 phons. The *PureTone* condition used pure tones of 300 and 700 Hz presented at 80 dB SPL-C to control for the possibility that the complex amplitude-modulation envelopes used in Experiment 1 interacted with the carrier signals in a way that facilitated pre-attentive detection of 700 Hz deviant events.

Two additional conditions (*SmallD.700* and *SmallD.300*) used smaller deviance sizes than in Experiment 1. This allowed for assessment of whether the stimulus-specific enhancement favoring the higher frequency stimulus (i.e., increment) found in Experiment 1 was related to the relatively large size of the frequency difference between standard and deviant carriers. Both *SmallD.700* and *SmallD.300* stimuli used a difference of a sixth of an octave: the *SmallD.700* condition used carrier frequencies of 700 and 817 Hz and the *SmallD.300* condition used carrier frequencies of 300 and 350 Hz. All stimuli in the *SmallD.300* and *SmallD.700* conditions were presented at 80 dB SPL-C.

Presentation and Recording

Stimulus presentation and recording of EEG data largely followed the same design outlined in Experiment 1. Each experiment session comprised 8 recording blocks with each block defined by the presentation of 600 trials (90 deviant) from one condition in one direction. The forward and reverse directions of each condition were always presented consecutively. The order of conditions and the order of the directions within each condition were pseudo-randomly chosen for each subject. The electrode impedance of all subjects was below 15 kΩs (mean = 7.33, SD = 3.01 kΩ).

Data Analysis

Standard and deviant data were epoched and averaged separately for each recording block following the methods for analysis of transient components outlined in Experiment 1. Peak amplitudes and latencies of the MMN and P3a components were extracted from the difference wave computed by subtracting the average response elicited by each stimulus when presented as a standard trial from that elicited when presented as a deviant trial.

The possible effect of deviance direction (i.e., increment or decrement) was examined separately for the MMN and P3a components, and for conditions using small and large deviance sizes. Thus, a total of four repeated-measures MANOVAs were conducted, the first two on conditions with large deviance sizes, and the second two on conditions with small deviance sizes. The first MANOVA assessed the within-subject factors of deviance direction and stimulus (i.e., *P.Loudness* or *PureTone*) on the dependent measures: MMN amplitude and MMN latency. The second MANOVA assessed the within-subject factors of deviance direction and stimulus (i.e., *P.Loudness* or *PureTone*) on the dependent measures: P3a amplitude and P3a latency. The final two MANOVAs were structured similarly to the first two except that amplitude and latency data, and hence the within-subjects factor of stimulus, corresponded to the MMN and P3a components evoked by *SmallD.700* and *SmallD.300* stimuli.

4.4.2 Results

Deviance direction had a significant effect on MMN and P3a morphologies when deviance size was large but not when deviance size was small (Table 4.2). With respect

to stimulus type, the repeated measures MANOVAs found no significant effect of stimulus (i.e., *P.Loudness* or *PureTone*) for conditions with large deviant size. There was a significant effect of stimulus for conditions with small deviant size but, importantly, for both large and small deviant sizes, there was no interaction between stimulus and deviance direction. This indicates that potential loudness differences and potential effects of amplitude modulation cannot explain the effects seen in Experiment 1.

TABLE 4.2: Summary of repeated-measures MANOVAs for MMN and P3a components as elicited by stimuli using large (i.e., *P.Loudness* and *PureTone*) and small (i.e., *SmallD.700* and *SmallD.300*) deviance sizes. For each MANOVA, component amplitudes and latencies were the dependent measures and stimulus and deviance direction (i.e., increment or decrement) were the within-subjects factors.

Dev. size	**AEP**	**Effect**	**Pillai's Trace**	***F***	***df***	**Error *df***	***p***	
		STIMULUS	0.68	4.24	2	4	0.1	
L	MMN	DEV.DIRECTION	0.79	7.6	2	4	0.04	*
A		STIM:DIR	0.12	0.28	2	4	0.77	
R								
G		STIMULUS	0.77	6.64	2	4	0.05	
E	P3A	DEV.DIRECTION	0.88	14.96	2	4	0.01	*
		STIM:DIR	0.66	3.91	2	4	0.11	
		STIMULUS	0.9	19.52	2	4	0.009	**
S	MMN	DEV.DIRECTION	0.28	0.76	2	4	0.53	
M		STIM:DIR	0.64	3.6	2	4	0.13	
A								
L		STIMULUS	0.91	3.12	2	4	0.15	
L	P3A	DEV.DIRECTION	0.01	0.02	2	4	0.98	
		STIM:DIR	0.43	1.5	2	4	0.33	

To increase power of subsequent tests the data from *P.Loudness* and *PureTone* were combined in the large deviant case. As the effect of stimulus was significant

for small deviance size conditions these data were not combined. Post-hoc paired samples t-tests revealed that stimuli with 700 Hz carriers elicited earlier MMN ($t(5) = -3.82, p < 0.05$) and P3a ($t(5) = -4.18, p < 0.01$) and stronger MMN ($t(5) = -2.71, p < 0.05$) than stimuli with 300 Hz carriers for conditions using a large deviance size (FIGURE 4.12.A and FIGURE 4.12.B). These effects were not significant in the case of small deviance sizes.

In sum, the results suggest that large frequency increments enhance the MMN and P3a component morphology relative to large decrements. This is not the case for small frequency changes. Furthermore, these enhancements were not related to differences in perceived loudness or amplitude envelopes between the low and high frequency stimuli.

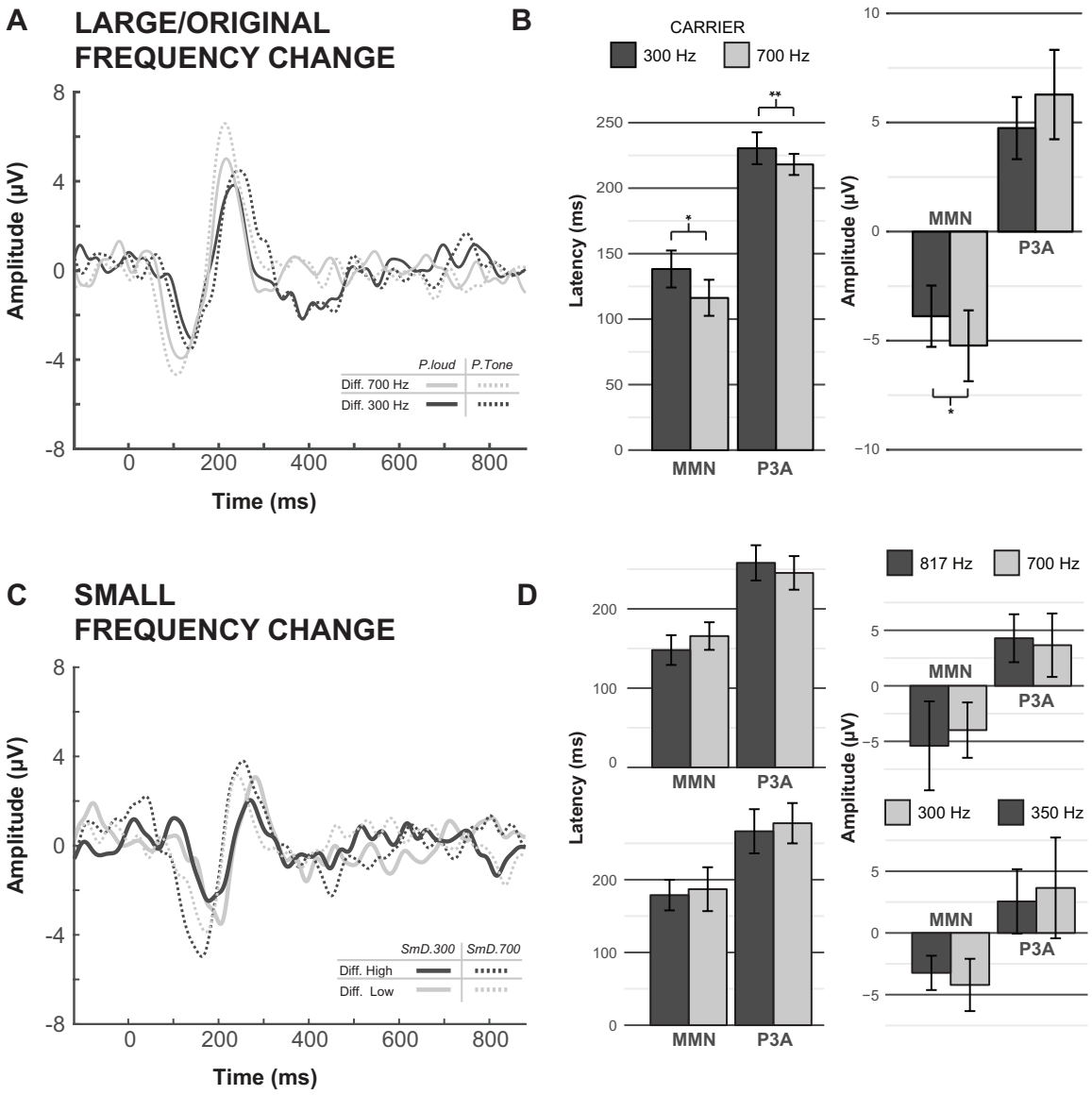

FIGURE 4.12: A comparison of the difference waves elicited by the 4 stimulus sets used in Experiment 2. (A) Difference waves as evoked by 300 (dark grey) and 700 Hz (light grey) carrier signals from *P.Loudness* (solid line) and *PureTone* (dotted line) stimulus sets. (B) Amplitudes and latencies of the change detection components from the combined *P.Loudness* and *PureTone* datasets. (C) Difference waves as evoked by lower (dark grey) and higher (light grey) frequency carrier signals from *SmallD.300* (solid line) and *SmallD.700* (dotted line) stimulus sets. (D) Column plots for (bottom) *SmallD.300* and (top) *SmallD.700* datasets. Error bars represent the within-subjects 95% confidence interval of the mean. Significant differences between the components based on the evoking carrier signal are denoted by asterisks.

4.5 Discussion

To our knowledge, this is the first study to demonstrate an association between acoustic feature encoding in subcortex and the onset of cortical change-detection components. Our data are also in agreement with recent studies reporting novelty-related modulation of activity generated at pre-cortical auditory structures (Althen *et al.*, 2013; Escera and Malmierca, 2014; Grimm *et al.*, 2011; Leung *et al.*, 2012; Recasens *et al.*, 2014b,a; Shiga *et al.*, 2015; Slabu *et al.*, 2010, 2012). This is supported by the observed attenuation of the subcortical FFR evoked by a 700 Hz carrier signal when presented as a standard relative to deviant event, and the relationship between the 700 Hz FFR amplitude and the latencies of the MMN and P3a components.

The contrast between standard- and deviant-evoked FFRs suggests more than local SSA at subcortical generators. According to work in animal models, midbrain neurons exhibiting SSA are largely found in dorsal regions of the IC and are primarily "onset" types whose robust responses are confined to the first 40 ms of the stimulus (Pérez-González *et al.*, 2005; Malmierca *et al.*, 2009; Pérez-González and Malmierca, 2012; Lumani and Zhang, 2010; Duque *et al.*, 2012). Consistent with the interpretation that the FFR primarily reflects activity of neurons in the IC, we observed the standard 300 and 700 Hz FFRs were attenuated at stimulus onset, and the standard 700 Hz FFR remained largely attenuated thereafter. Indeed, limiting our analysis to those standards which immediately preceded deviant events probably maximized the likelihood of observing adaptation-like attenuation of the standard FFR response (Skoe *et al.*, 2013). However, our data also revealed a late enhancement of the deviant-evoked FFR in a window immediately following the cortical P3a

component. Though we cannot directly infer the extent to which this late enhancement reflects the influence of corticofugal projections, the strength of the correlation between 700 Hz FFR amplitude and the latency of the novelty detection components (i.e., MMN and P3a), especially in those subjects that exhibited FFR enhancement to the deviant event, warrants considering this possibility.

The discovery of SSA in the midbrain (Duque *et al.*, 2012; Malmierca *et al.*, 2009; Pérez-González *et al.*, 2005) has given way to the suggestion that change-detection may be a fundamental attribute of auditory processing. Though the exact nature of the relationship between cortical and subcortical SSA remains unclear, it is possible that at least part of SSA at midbrain may be inherited from cortical structures (Nelken and Ulanovsky, 2007). For example, Bäuerle *et al.* (2011) found a substantial reduction in SSA at the thalamus of gerbils following pharmacological lesioning of AC. Similarly, reversible deactivation of rat AC with cooling reduced but did not abolish SSA in the IC (Anderson *et al.*, 2013), suggesting that corticocollicular projections are not essential for generating subcortical SSA but may still have a role in its modulation. In such an integrated system, novel acoustic events release subcortical neurons from local SSA which feeds forward to trigger change-detection mechanisms at higher levels of the central auditory system. This higher level activity can then modulate the encoding of the ascending auditory signal at lower levels to enhance potentially biologically-relevant acoustic features. By adapting to repetitive stimulation in this way, the auditory system would be primed to respond to any significant immediate change in the acoustic environment (Chandrasekaran *et al.*, 2014). The action

of such a system might also explain why in our data enhancement of the deviant-evoked FFR, relative to the standard response, was observable at stimulus onset but strongest following cortical change-detection components.

We must acknowledge that the stimulus-specific asymmetry of these effects remains a significant caveat to any interpretation. Stimulus salience may partly explain why FFR modulation was only observed for the 700 Hz carrier signal. Slabu *et al.* (2012) found that deviant presentation of a /ba/ syllable among frequently presented /wa/ syllables modulated the strength of the second harmonic in the FFR compared to presentations of /ba/ in either a standard context or in a block where stimulus occurrence was equiprobable. Such modulation was not observed in reverse for any harmonics of the /wa/ syllable. One possibility offered by the authors is that deviance detection at the brainstem level is more sensitive to the greater saliency of certain stimulus features (e.g., faster rising formant transition of /ba/) in the context of less salient features (slower rising formant transition of /wa/). In the context of the present study, the higher pitched 700 Hz carrier signal might have been more salient than the lower pitched 300 Hz signal. Difference limens for frequency discriminations are smaller when testing with increments *versus* decrements (Kishon-Rabin *et al.*, 2004). The amplitude of the MMN has also been reported to be larger for frequency increments relative to decrements (Peter *et al.*, 2010). Consistent with this possibility, the 700 Hz carrier signal elicited stronger N1-P2 and MMN as well as earlier MMN and P3a components relative to the 300 Hz carrier. The results of Experiment 2 further suggest that the enhanced transient AEPs evoked by 700 Hz carriers did not result from differences in perceived loudness, unique amplitude-modulation patterns, or deviance direction alone. Rather, the magnitude of the increment (i.e., more than

an octave *versus* 1/6 of an octave) was critical to reproducing enhanced transients. Future studies might look at smaller differences at high and low frequencies, similar to those used in Experiment 2, to assess whether modulation of the FFR found in Experiment 1 is still observed when the contrast between standard and deviant presentations is less salient.

Previous work has demonstrated modulation of 40 Hz ASSRs to changing stimulus parameters. For example, Ross (2008) reported phase advancement of the 40 Hz ASSR in response to infrequent changes in interaural phase difference resulting in a perceptual change from a focal to a spacious sound. By comparing ASSRs evoked by the same stimuli in different contexts, our results further show that such modulation is not related to physical differences in the evoking stimuli. Curiously, phase advances for deviant-evoked 40 Hz ASSRs were observed for both carrier frequencies but were not associated with either change-related modulation of the FFR or the morphology of cortical change-detection components. There is some suggestion that deviant-related modulation of the 40 Hz ASSR might result from fresh afferent stimulation at the earliest stages of auditory processing (i.e., the auditory nerve and cochlear nucleus). Chakalov *et al.* (2014) demonstrated this possibility in an experiment using binaural beats in a forward-reverse oddball paradigm, similar to that employed in the present study. The binaural beat is an acoustic phenomenon that can be experienced by dichotic presentation of two different tones via headphones (Oster, 1973). The resultant percept is that of a rough or "beating" sound at the difference frequency of the two tones, so long as that difference is less than 30 Hz. This beating is considered to result from neural coupling in the brainstem structures that receive input from both ears (i.e., superior olivary complex and beyond). Using this phenomenon,

Chakalov *et al.* (2014) dichotically presented two tones (250 Hz and 276 Hz) to elicit a cortical ASSR at the beat frequency of 26 Hz. Infrequently, they would switch the ear to which each tone was presented, thereby stimulating fresh afferents from each cochlea without altering the beating precept that presumably results from integration of the tones at higher auditory structures. Comparison of the activity evoked by each dichotic assignment in standard and deviant contexts revealed significant deviant-related enhancement of the 26 Hz ASSR, strongly suggesting the involvement of SSA at auditory nerve and cochlear nucleus in deviant-detection processes at the cortical generators of the ASSR. It is possible that the faster synchronization of thalamocortical activity observed in the present experiment as phase advancement of the 40 Hz ASSR reflects a similar contribution of enhanced activity from fresh afferents early in the auditory pathway. In a previous experiment, Slugocki *et al.* (2017) found the phase of the 80 Hz ASSR, with purported generators in cochlear nucleus (Suzuki, 2000), to be correlated with that of the 40 Hz ASSR but not with the onset of an FFR at 500 Hz. Given the complexity of ascending and descending projections in the auditory system, it is not unreasonable to speculate that separate change-detection sub-circuits might exist and cooperate in a way that cannot be delineated using field recordings. It remains to be seen whether novelty-related modulation at 40 Hz ASSR generators behaves differently to that observed at generators of the FFR and MLRs and how such differences might inform us as to the hierarchical arrangement of change-detection processes along the auditory neuraxis.

As far as we are aware, no study has reported context-dependent modulation of the 80 Hz ASSR. Here, we observed stronger standard- relative to deviant-evoked 80 Hz ASSR (77 Hz; *SEAP300*), indicating that this effect does not reflect SSA. This

is consistent with purported generators of 80 Hz ASSR residing predominantly in cochlear nucleus (Suzuki, 2000), whereas SSA is not a widespread phenomenon in nuclei below the IC (Ayala *et al.*, 2013). It is possible that deviant presentation of *SEAP300* acted to disrupt the contribution of cortical sources to the 80 Hz ASSR (Herdman *et al.*, 2002). It is unknown why similar modulation was not observed for the 80 Hz ASSR evoked by *SEAP700*. Perhaps the cortical contribution to the 80 Hz ASSR is particularly tuned for encoding repetitive stimulus features in the amplitude envelopes of lower frequency sounds. Temporal envelope information is indeed an important cue to speech intelligibility (Drullman, 1995). Further, the long-term average speech spectrum is highly concentrated between 250 and 500 Hz and drops off thereafter by 6 dB per octave until 4 kHz (Byrne *et al.*, 1994). The periodicity-coding capabilities of the FFR show a similar maximum at 500 Hz but stop at 1.5 kHz. Future work should examine whether speech-shaped noise modulated at 80 Hz shows similar context-dependent modulation.

In this discussion we adopted a largely neuroanatomical and neurophsysiological-based interpretation of our data in the context of recent integrative models of distributed change-detection and plastic processes along the mammalian auditory hierarchy (Chandrasekaran *et al.*, 2014; Escera and Malmierca, 2014). In so doing, we chose to accept the most widely purported generators of each AEP under consideration. However, it must be acknowledged that our data are limited to field recordings of a single scalp channel. Field recordings are at best indirect correlates of SSA as such phenomena are typically characterized in single neurons (Ayala and Malmierca, 2013; Malmierca *et al.*, 2009; Pérez-González and Malmierca, 2012; Pérez-González *et al.*, 2005; Zhao *et al.*, 2011).

As we did not include a control block with equiprobable presentation, our data also cannot speak to whether attenuation of the standard 700 Hz FFR is a prerequisite correlate of true novelty-detection at the subcortical level or simply an artifact of neural refractoriness. However, in previous reports, modulation of the FFR has survived contrasts of deviant-evoked activity and that evoked by the same stimuli presented in equiprobable blocks when such a control is used (Shiga *et al.*, 2015; Slabu *et al.*, 2012). Further, all inferences regarding the relationship between subcortical and cortical activity, as well as those concerning the underlying mechanisms of novelty-detection and plasticity at either level, are based on a single waveform and so must be considered speculative. Recent work has shown the FFR to include a cortical component (Coffey *et al.*, 2016, 2017). Though this alone does not rule out co-modulation of the subcortical contribution, we must consider the possibility that the context-dependent modulation of FFR observed here is occurring at the cortical level. Despite these caveats, we feel such discussion is constructive in considering how recently-developed methodologies for concurrent measurement of cortical and subcortical AEPs might be used to test predictions of distributed deviance detection models in human listeners. In turn, we hope this may guide future use of multi-channel concurrent recordings in verifying the effects described herein with the added confidence of better estimating their respective underlying neural sources.

4.6 Conclusion

This study adds to a growing body of literature demonstrating that change-detection processes are not limited to the auditory cortex. We show that modulation of subcortical generators follows the onset of the cortical MMN and P3a change-detection

components, suggesting that it may reflect corticofugal information flow. Future research should make use of multi-channel concurrent recordings of subcortical and cortical activity to better delineate how change-detection at these different levels of the auditory hierarchy might contribute to the formation of perceptual auditory objects (Nelken, 2004; Winkler *et al.*, 2009) and attention-orienting mechanisms (Escera *et al.*, 1998, 2000).

4.7 References

Alho, K. (1995). Cerebral generators of mismatch negativity (MMN) and its magnetic counterpart (MMNm) elicited by sound changes. *Ear and Hearing*, **16**(1), 38–51.

Alho, K., Grimm, S., Mateo-León, S., Costa-Faidella, J., and Escera, C. (2012). Early processing of pitch in the human auditory system: early processing of pitch. *European Journal of Neuroscience*, **36**(7), 2972–2978.

Althen, H., Grimm, S., and Escera, C. (2013). Simple and complex acoustic regularities are encoded at different levels of the auditory hierarchy. *European Journal of Neuroscience*, **38**(10), 3448–3455.

Althen, H., Huotilainen, M., Grimm, S., and Escera, C. (2016). Middle latency response correlates of single and double deviant stimuli in a multi-feature paradigm. *Clinical Neurophysiology*, **127**(1), 388–396.

Anderson, S., White-Schwoch, T., Parbery-Clark, A., and Kraus, N. (2013). Reversal of age-related neural timing delays with training. *Proceedings of the National Academy of Sciences of the United States of America*, **110**(11), 4357–4362.

Antunes, F. M., Nelken, I., Covey, E., and Malmierca, M. S. (2010). Stimulus-specific adaptation in the auditory thalamus of the anesthetized rat. *Public Library of Science One*, **5**(11), e14071.

Ayala, Y. A. and Malmierca, M. S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Frontiers in Neural Circuits*, **6**, 1–19.

Ayala, Y. A. and Malmierca, M. S. (2015). Cholinergic modulation of stimulus-specific adaptation in the inferior colliculus. *The Journal of Neuroscience*, **35**(35), 12261–12272.

Ayala, Y. A., Pérez-González, D., Duque, D., Nelken, I., and Malmierca, M. S. (2013). Frequency discrimination and stimulus deviance in the inferior colliculus and cochlear nucleus. *Frontiers in Neural Circuits*, **6**, 1–19.

Ayala, Y. A., Udeh, A., Dutta, K., Bishop, D., Malmierca, M. S., and Oliver, D. L. (2015). Differences in the strength of cortical and brainstem inputs to SSA and non-SSA neurons in the inferior colliculus. *Scientific Reports*, **5**, 10383.

Bäuerle, P., Behrens, W. v. d., Kössl, M., and Gaese, B. H. (2011). Stimulus-specific adaptation in the gerbil primary auditory thalamus is the result of a fast frequency-specific habituation and is regulated by the corticofugal system. *The Journal of Neuroscience*, **31**(26), 9708–9722.

Berens, P. (2009). CircStat: a MATLAB toolbox for circular statistics. *Journal of Statistical Software*, **31**(10), 1–21.

Bohórquez, J. and Özdamar, Ö. (2008). Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution. *Clinical Neurophysiology*, **119**(11), 2598–2607.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C., Kiessling, J., Kotby, M. N., Nasser, N. H., El Kholy, W. A., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Fronlenkov, G. I., Westerman, S., and Ludvigsen, C. (1994).

An international comparison of longterm average speech spectra. *The Journal of the Acoustical Society of America*, **96**(4), 2108–2120.

Cacciaglia, R., Escera, C., Slabu, L., Grimm, S., Sanjuán, A., Ventura-Campos, N., and Ávila, C. (2015). Involvement of the human midbrain and thalamus in auditory deviance detection. *Neuropsychologia*, **68**, 51–58.

Chakalov, I., Paraskevopoulos, E., Wollbrink, A., and Pantev, C. (2014). Mismatch negativity to acoustical illusion of beat: how and where the change detection takes place? *NeuroImage*, **100**, 337–346.

Chandrasekaran, B. and Kraus, N. (2010). The scalp-recorded brainstem response to speech: neural origins and plasticity. *Psychophysiology*, **47**(2), 236–246.

Chandrasekaran, B., Skoe, E., and Kraus, N. (2014). An integrative model of subcortical auditory plasticity. *Brain Topography*, **27**(4), 539–552.

Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, **7**, 11070.

Coffey, E. B. J., Musacchia, G., and Zatorre, R. J. (2017). Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *The Journal of Neuroscience*, **37**(4), 830–838.

Delorme, A. and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, **134**(1), 9–21.

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *The Journal of the Acoustical Society of America*, **97**(1), 585–592.

Duque, D., Pérez-González, D., Ayala, Y. A., Palmer, A. R., and Malmierca, M. S. (2012). Topographic distribution, frequency, and intensity dependence of stimulus-specific adaptation in the inferior colliculus of the rat. *The Journal of Neuroscience*, **32**(49), 17762–17774.

Escera, C. and Corral, M. (2007). Role of mismatch negativity and novelty-P3 in involuntary auditory attention. *Journal of Psychophysiology*, **21**(3-4), 251–264.

Escera, C. and Malmierca, M. S. (2014). The auditory novelty system: an attempt to integrate human and animal research. *Psychophysiology*, **51**(2), 111–123.

Escera, C., Alho, K., Winkler, I., and Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, **10**(5), 590–604.

Escera, C., Alho, K., Schröger, E., and Winkler, I. (2000). Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neurotology*, **5**(3-4), 151–166.

Fisher, N. (1995). *Statistical Analysis of Circular Data*. Cambridge University Press.

Fishman, Y. I. and Steinschneider, M. (2012). Searching for the mismatch negativity in primary auditory cortex of the awake monkey: deviance detection or stimulus specific adaptation? *The Journal of Neuroscience*, **32**(45), 15747–15758.

Frisina, R. D., Smith, R. L., and Chamberlain, S. C. (1990). Encoding of amplitude modulation in the gerbil cochlear nucleus: I. A hierarchy of enhancement. *Hearing Research*, **44**(23), 99–122.

Galambos, R. and Makeig, S. (1992). Physiological studies of central masking in man. I: The effects of noise on the 40Hz steadystate response. *The Journal of the Acoustical Society of America*, **92**(5), 2683–2690.

Games, K. D. and Winer, J. A. (1988). Layer V in rat auditory cortex: projections to the inferior colliculus and contralateral cortex. *Hearing Research*, **34**(1), 1–25.

Gao, P. P., Zhang, J. W., Cheng, J. S., Zhou, I. Y., and Wu, E. X. (2014). The inferior colliculus is involved in deviant sound detection as revealed by BOLD fMRI. *NeuroImage*, **91**, 220–227.

Glasberg, B. R. and Moore, B. C. J. (2006). Prediction of absolute thresholds and equal-loudness contours using a modified loudness model. *The Journal of the Acoustical Society of America*, **120**(2), 585–588.

Grimm, S., Escera, C., Slabu, L., and Costa-Faidella, J. (2011). Electrophysiological evidence for the hierarchical organization of auditory change detection in the human brain: hierarchical organization of auditory change detection. *Psychophysiology*, **48**(3), 377–384.

Grimm, S., Recasens, M., Althen, H., and Escera, C. (2012). Ultrafast tracking of sound location changes as revealed by human auditory evoked potentials. *Biological Psychology*, **89**(1), 232–239.

Herdman, A. T., Lins, O., Roon, P. V., Stapells, D. R., Scherg, M., and Picton, T. W. (2002). Intracerebral sources of human auditory steady-state responses. *Brain Topography*, **15**(2), 69–86.

Kishon-Rabin, L., Roth, D. A.-E., Dijk, B. V., Yinon, T., and Amir, O. (2004). Frequency discrimination thresholds: the effect of increment versus decrement detection of frequency. *Journal of Basic and Clinical Physiology and Pharmacology*, **15**(1-2), 29–40.

Kraus, N. and Nicol, T. (2005). Brainstem origins for cortical what and where pathways in the auditory system. *Trends in Neurosciences*, **28**(4), 176–181.

Leung, S., Cornella, M., Grimm, S., and Escera, C. (2012). Is fast auditory change detection feature specific? An electrophysiological study in humans. *ResearchGate*, **49**(7), 933–42.

Leung, S., Recasens, M., Grimm, S., and Escera, C. (2013). Electrophysiological index of acoustic temporal regularity violation in the middle latency range. *Clinical Neurophysiology*, **124**(12), 2397–2405.

Lopez-Calderon, J. and Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, **8**, 1–14.

Lumani, A. and Zhang, H. (2010). Responses of neurons in the rat's dorsal cortex of the inferior colliculus to monaural tone bursts. *Brain Research*, **1351**, 115–129.

Malmierca, M. S. and Ryugo, D. K. (2011). Descending connections of auditory cortex to the midbrain and brain stem. In J. A. Winer and C. E. Schreiner, editors, *The Auditory Cortex*, pages 189–208. Springer US, Boston, MA.

Malmierca, M. S., Cristaudo, S., Pérez-González, D., and Covey, E. (2009). Stimulus-specific adaptation in the inferior colliculus of the anesthetized rat. *The Journal of Neuroscience*, **29**(17), 5483–5493.

Maris, E. and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, **164**(1), 177–190.

Martínez-Montes, E., Cuspineda-Bravo, E. R., El-Deredy, W., Sánchez-Bornot, J. M., Lage-Castellanos, A., and Valdés-Sosa, P. A. (2008). Exploring event-related brain dynamics with tests on complex valued timefrequency representations. *Statistics in Medicine*, **27**(15), 2922–2947.

McGee, T., Kraus, N., Littman, T., and Nicol, T. (1992). Contributions of medial geniculate body subdivisions to the middle latency response. *Hearing Research*, **61**(1-2), 147–154.

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**(4), 375–425.

Näätänen, R., Paavilainen, P., Rinne, T., and Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: a review. *Clinical Neurophysiology*, **118**(12), 2544–2590.

Nelken, I. (2004). Processing of complex stimuli and natural scenes in the auditory cortex. *Current Opinion in Neurobiology*, **14**(4), 474–480.

Nelken, I. and Ulanovsky, N. (2007). Mismatch negativity and stimulus-specific adaptation in animal models. *Journal of Psychophysiology*, **21**(3-4), 214–223.

Oster, G. (1973). Auditory beats in the brain. *Scientific American*, **229**(4), 94–102.

Patel, C. R., Redhead, C., Cervi, A. L., and Zhang, H. (2012). Neural sensitivity to novel sounds in the rat's dorsal cortex of the inferior colliculus as revealed by evoked local field potentials. *Hearing Research*, **286**(12), 41–54.

Pérez-González, D. and Malmierca, M. S. (2012). Variability of the time course of stimulus-specific adaptation in the inferior colliculus. *Frontiers in Neural Circuits*, **6**, 1–12.

Pérez-González, D., Malmierca, M. S., and Covey, E. (2005). Novelty detector neurons in the mammalian auditory midbrain. *European Journal of Neuroscience*, **22**(11), 2879–2885.

Peruzzi, D., Sivaramakrishnan, S., and Oliver, D. L. (2000). Identification of cell types in brain slices of the inferior colliculus. *Neuroscience*, **101**(2), 403–416.

Peter, V., McArthur, G., and Thompson, W. F. (2010). Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neuroscience Letters*, **482**(1), 71–75.

Picton, T. W., Alain, C., Otten, L., Ritter, W., and Achim, A. (2000). Mismatch negativity: different water in the same river. *Audiology and Neurotology*, **5**(3-4), 111–139.

Plourde, G., Stapells, D. R., and Picton, T. W. (1991). The human auditory steady-state evoked potentials. *Acta Oto-Laryngologica*, **111**(sup491), 153–160.

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology*, **118**(10), 2128–2148.

Ponnath, A., Hoke, K. L., and Farris, H. E. (2013). Stimulus change detection in phasic auditory units in the frog midbrain: frequency and ear specific adaptation. *Journal of Comparative Physiology A*, **199**(4), 295–313.

Recasens, M., Grimm, S., Wollbrink, A., Pantev, C., and Escera, C. (2014a). Encoding of nested levels of acoustic regularity in hierarchically organized areas of the human auditory cortex. *Human Brain Mapping*, **35**(11), 5701–5716.

Recasens, M., Grimm, S., Capilla, A., Nowak, R., and Escera, C. (2014b). Two sequential processes of change detection in hierarchically ordered areas of the human auditory cortex. *Cerebral Cortex*, **24**(1), 143–153.

Ross, B. (2008). A novel type of auditory responses: temporal dynamics of 40-Hz steady-state responses induced by changes in sound localization. *Journal of Neurophysiology*, **100**(3), 1265–1277.

Schönwiesner, M., Krumbholz, K., Rübsamen, R., Fink, G. R., and von Cramon, D. Y. (2007). Hemispheric asymmetry for auditory processing in the human auditory brain stem, thalamus, and cortex. *Cerebral Cortex*, **17**(2), 492–499.

Shiga, T., Althen, H., Cornella, M., Zarnowiec, K., Yabe, H., and Escera, C. (2015). Deviance-related responses along the auditory hierarchy: combined FFR, MLR and MMN evidence. *Public Library of Science One*, **10**(9), e0136794.

Skoe, E. and Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and Hearing*, **31**(3), 302–324.

Skoe, E., Krizman, J., Spitzer, E., and Kraus, N. (2013). The auditory brainstem is a barometer of rapid auditory learning. *Neuroscience*, **243**, 104–114.

Slabu, L., Escera, C., Grimm, S., and Costa-Faidella, J. (2010). Early change detection in humans as revealed by auditory brainstem and middle-latency evoked potentials: early auditory change detection in humans. *European Journal of Neuroscience*, **32**(5), 859–865.

Slabu, L., Grimm, S., and Escera, C. (2012). Novelty detection in the human auditory brainstem. *The Journal of Neuroscience*, **32**(4), 1447–1452.

Slugocki, C., Bosnyak, D., and Trainor, L. J. (2017). Simultaneously-evoked auditory potentials (SEAP): a new method for concurrent measurement of cortical and subcortical auditory-evoked activity. *Hearing Research*, **345**, 30–42.

Smith, J. C., Marsh, J. T., and Brown, W. S. (1975). Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. *Electroencephalography and Clinical Neurophysiology*, **39**(5), 465–472.

Smith, J. C., Marsh, J. T., Greenberg, S., and Brown, W. S. (1978). Human auditory frequency-following responses to a missing fundamental. *Science*, **201**(4356), 639–641.

Sohmer, H., Pratt, H., and Kinarti, R. (1977). Sources of frequency following responses (FFR) in man. *Electroencephalography and Clinical Neurophysiology*, **42**(5), 656–664.

Sonnadara, R. R., Alain, C., and Trainor, L. J. (2006). Effects of spatial separation and stimulus probability on the event-related potentials elicited by occasional changes in sound location. *Brain Research*, **1071**(1), 175–185.

Stephens, M. A. (1969). Tests for randomness of directions against two circular alternatives. *Journal of the American Statistical Association*, **64**(325), 280–289.

Suzuki, Y. (2000). Contribution of cochlear nucleus to 80 Hz amplitude-modulation following response. *Nippon Jibiinkoka Gakkai Kaiho*, **103**, 177–187.

Ulanovsky, N., Las, L., and Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, **6**(4), 391–398.

Watson, G. S. and Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika*, **43**(3/4), 344–352.

Winer, J. A. and Schreiner, C., editors (2005). *The Inferior Colliculus*. Springer, New York, NY.

Winer, J. A., Larue, D. T., Diehl, J. J., and Hefti, B. J. (1998). Auditory cortical projections to the cat inferior colliculus. *The Journal of Comparative Neurology*, **400**(2), 147–174.

Winer, J. A., Diehl, J. J., and Larue, D. T. (2001). Projections of auditory cortex to the medial geniculate body of the cat. *Journal of Comparative Neurology*, **430**(1), 27–55.

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, **13**(12), 532–540.

Wolpaw, J. R. and Penry, J. K. (1975). A temporal component of the auditory evoked response. *Electroencephalography and Clinical Neurophysiology*, **39**(6), 609–620.

Woods, D. L. (1995). The component structure of the N1 wave of the human auditory evoked potential. *Perspectives of Event-Related Potentials Research*, **44**, 102–209.

Worden, F. G. and Marsh, J. T. (1968). Frequency-following (microphonic-like) neural responses evoked by sound. *Electroencephalography and Clinical Neurophysiology*, **25**(1), 42–52.

Yamaguchi, S. and Knight, R. T. (1991). P300 generation by novel somatosensory stimuli. *Electroencephalography and Clinical Neurophysiology*, **78**(1), 50–55.

Zhao, L., Liu, Y., Shen, L., Feng, L., and Hong, B. (2011). Stimulus-specific adaptation and its dynamics in the inferior colliculus of rat. *Neuroscience*, **181**, 163–174.

Zurek, P. M. (1992). Detectability of transient and sinusoidal otoacoustic emissions. *Ear and Hearing*, **13**(5), 307–310.

Chapter 5

Discussion and conclusion

5.1 Summary and unique contributions

This thesis contributes to the human AEP literature by demonstrating that concurrent measurement of subcortical and cortical AEP components is a viable approach for studying distributed change-detection processes in infancy and adulthood. Using this approach, Chapter 2 shows for the first time that 2-month-old infants respond to infrequent changes in sound source location with neural activity implicating both subcortically- and cortically-driven change-detection mechanisms. The results of Chapter 2 directly question a leading hypothesis which posits that behavioural insensitivity to changes in sound source location reflect suppression of subcortically-mediated sound localization mechanisms between 1 and 3 months-of-age (e.g., Muir and Hains, 2004). In Chapter 3, I demonstrate the viability of a new stimulation protocol for the concurrent recording of subcortical and cortical AEPs in normal hearing adults. Normative data, collected from adult listeners, further show that the morphologies of several well-known AEP components attributed to either subcortical

or cortical generators are related. The results of Chapter 3 are among the first (see also: Musacchia *et al.*, 2008) to suggest a link between acoustic feature encoding at subcortical nuclei, as indexed by the stimulus-to-response fidelity of the FFR, and higher level acoustic feature integration as indexed by the latency of the N1. Finally, Chapter 4 builds upon the methods developed in Chapter 3 by presenting the SEAP stimuli in a forward-reverse oddball. This approach allowed for a direct comparison of subcortical and cortical AEPs as elicited by the same physical stimuli, but in standard relative to deviant contexts. The results of Chapter 4 suggest that regularity encoding operates in both a bottom-up and top-down manner when normal-hearing adults are presented with novel acoustic events. Together, these studies emphasize the potential for EEG techniques to be applied in studying the interaction of subcortical and cortical processing in human listeners. Moreover, the results of Chapters 2 through 4 lend support to models wherein change-detection is considered a distributed and perhaps fundamental attribute of the auditory hierarchy, and one which might ultimately underlie the formation of auditory objects (e.g., Winkler *et al.*, 2009).

Chapters 2 and 4 further a growing body of evidence that implicates thalamic neurons and thalamocortical circuitry as important early generative stages of sensory prediction. As previously discussed, the presence of a slow-positive MMR in early infancy is thought to reflect subcortically-driven stimulation of layer IV of A1 via thalamocortical afferents (Eggermont and Moore, 2012). The difference between standard- and deviant-evoked electrophysiological activities in the infant brain might possibly reflect the stimulation of fresh (i.e., un-adapted) neurons at lower level auditory nuclei. Indeed, neurons in the ICc and MGBv of other mammals are known to be sensitive to the kind of spatial information used to evoke the MMR in Chapter

2 (Chase and Young, 2008). These neurons are also known to exhibit SSA (e.g., Anderson *et al.*, 2009; Ayala and Malmierca, 2013), thereby providing a putative mechanism for simple stimulus-feature based regularity encoding without the involvement of higher level change-detection processes that rely on mature cortical circuitry. Such a mechanism might also explain why the MMR elicited by spatial deviants was observed to persist longer into development in Chapter 2 compared to what is typically reported for other acoustic deviants (e.g., pitch; He *et al.*, 2007). Whereas He *et al.* (2007) used relatively small pitch changes (1/12 of an octave), Chapter 2 used maximally large ($\pm 90°$) spatial deviants as well as broadband (i.e., white) noise which can broadly stimulate many frequency-specific neurons. Such stimulation may have activated a population of fresh afferents large enough to support a strong layer IV current sink and thus overcome the newly developing inputs to layers I and/or II and their associated current sinks (Eggermont and Moore, 2012). The difference in the developmental trajectory of the MMR when elicited by pitch *versus* spatial deviants also suggests that the relative contribution of subcortical and cortical change-detection mechanisms may not strictly be determined by neuroanatomical maturation of auditory cortex.

Some researchers have posited that modulation of the 40 Hz ASSR by infrequent acoustic events might also result from fresh afferent stimulation at early stages of central processing (Chakalov *et al.*, 2014). In Chapter 4, the phase of deviant-evoked ASSRs around 40 Hz advanced relative to standard-evoked ASSRs of the same frequency when the acoustic change involved large (i.e., $>$ octave) frequency differences. Phase advancement was also observed following short-term auditory training and passive exposure to particular stimuli (Bosnyak *et al.*, 2004; Gander *et al.*, 2010),

wherein training/exposure may have presumably re-tuned the response sensitivities of closely-matched neurons towards the spectrotemporal characteristics of the target stimuli (Suga *et al.*, 2000). Should the effects observed in Chapters 2 and 4 indeed reflect SSA at thalamocortical neurons, then they are in agreement with previous studies showing deviant-related modulation of AEPs associated with activity in auditory midbrain, namely the MLR (Grimm *et al.*, 2016).

The results of Chapter 4 further implicate corticofugal (i.e., top-down) modulation of acoustic feature encoding at pre-cortical sources as part of a distributed change-detection system. Specifically, the infrequent presentation of an amplitude modulated pure tone (700 Hz carrier) significantly enhanced the power of subcortical phase-locked activity (i.e., FFR). Lau *et al.* (2017) also reported modulation of the FFR in an experiment designed to disentangle bottom-up regularity encoding processes from the engagement of corticofugal feedback loops. The study compared the subcortical encoding of a falling lexical pitch pattern (Cantonese, a tonal language) in native Cantonese speakers when the pitch pattern was presented in three contexts. In the repetitive context, the pitch pattern was presented on all trials. In the variable context, the pitch pattern occurred randomly and at equal probability with two other tones. In the patterned context, the pitch pattern was presented at equal probability with two other tones but in a repetitive, predictable pattern. The fidelity of the FFR (i.e., the stimulus-to-response cross-correlation) for the falling pitch pattern was stronger when it was presented in the repetitive relative to the variable context, and strongest when presented in the patterned and predictable context. The authors posited that the repetitive context engaged top-down predictive coding mechanisms

to a greater degree than the variable context such that the effect of predictive coding overrode the local effects of reduced SSA at subcortical generators of the FFR. Further, the authors suggested that the stimulus-to-response fidelity of the FFR was strongest in the patterned context because it reflected the combined effects of a top-down predictive coding mechanism, in that the tone sequence was 100% predictable, and an attenuation in local SSA in subcortex in that the falling pitch pattern never occurred on consecutive trials.

If similar mechanisms were engaged by the auditory oddball stimulation (85% transitional probability) used in Chapter 4, then the influence of top-down predictive coding mechanisms on subcortical FFR should have been smaller than those reported by Lau *et al.* (2017), wherein the repetitive stimulus occurred with 100% probability. Differences in the strength of the FFR evoked by a stimulus in the standard *versus* deviant context might thereby be expected to reflect the different degrees to which each stimulus context induced SSA in subcortical generators. Attenuation of standard relative to deviant activity at stimulus onset, though non-significant, was observed for both 300 and 700 Hz FFRs (Fig. 4.7 and Fig. 4.8). Brainstem neurons exhibiting SSA are largely found in dorsal regions of the IC and are primarily "onset" types whose response is confined to the first 40 ms of the stimulus (Pérez-González *et al.*, 2005; Malmierca *et al.*, 2009; Pérez-González and Malmierca, 2012; Lumani and Zhang, 2010; Duque *et al.*, 2012). This observation is at least consistent with SSA at subcortical FFR generators. However, enhancement of the deviant- relative to standard-evoked 700 Hz FFR activity was greatest in a window that immediately followed cortical change-detection components (i.e., MMN and P3a). Moreover, the latencies of the MMN and P3a components were significantly correlated with the

strength of the deviant-evoked FFR energy in this window. In terms of predictive coding theory, the modulation of ascending auditory information following a novel acoustic event could indicate processes related to model updating, wherein the response properties of lower level neurons are adjusted to reflect features of the new auditory input. The degree to which different elements of a predictive coding system are dependent on the transitional probability implied by presentation context is a topic that warrants further investigation.

It is curious that modulation of the deviant-evoked FFR was not observed for 300 Hz carrier stimulus in Chapter 4. A possible explanation is that modulation of the deviant-evoked 700 Hz FFR reflected the engagement of attention triggered by the salience of a rare higher frequency sound event in a stream of otherwise repetitive lower frequency (i.e., 300 Hz) sounds. Consistent with this idea, the difference limens for frequency discriminations were reported to be smaller when testing with increments *versus* decrements (Kishon-Rabin *et al.*, 2004). The amplitude of the MMN response was also reported to be larger for frequency increments relative to decrements (Peter *et al.*, 2010). Indeed, larger MMN and earlier MMN and P3a components were observed for 700 relative to 300 Hz deviants in both Experiments 1 and 2 of Chapter 4. The reason for such enhanced sensitivity to frequency increments in all of these studies remains unclear.

One possibility is related to asymmetries in the shape of the auditory filter and pattern of excitation on the basilar membrane, wherein the slope of the low-frequency end is steeper than that of the high-frequency end (Moore, 1973, 1997). According to Zwicker's (1956) place model, pitch discrimination relates to the detection of changes in basilar membrane excitation at the point of the steepest slope (i.e., low-frequency

edge). However, the two carriers used in Chapter 4 were separated by more than the estimated equivalent rectangular bandwidth (∼100 Hz) of an auditory filter centered on 700 Hz (Moore and Glasberg, 1996). Another possibility is that the auditory system is more likely to treat an ongoing low-frequency sound as background and a higher-frequency infrequent sound as a possible communications signal. Speaking to this, Fujioka *et al.* (2005) showed that both musicians and non-musicians exhibit a larger MMN in response to tonal deviants in a "high voice" melody (C5–G5; 523.25–789.99 Hz) compared to a "low voice" melody (C4–G4; 261.63–392.00 Hz). Marie and Trainor (2013) further demonstrated that a similar high-voice superiority can be observed in 7-month-olds, suggesting a possibly innate basis, although it appears that extensive training with a low-voiced instrument (e.g., cello or bass) can close the amplitude divide between MMN components elicited by high-voiced *versus* low-voiced deviants (Marie *et al.*, 2012). Perhaps endowing the 300 Hz deviant stimulus with behavioural-relevance would have engaged similar mechanisms and also resulted in modulation of the FFR. Unfortunately, to my knowledge no auditory oddball experiments to date have recorded the FFR while manipulating stimulus relevance in this way.

Interpreting the difference between the 300 and 700 Hz deviant-evoked FFRs is further complicated by recent work suggesting that the FFR elicited at frequencies as high as 200-250 Hz also reflects contributions from cortical sources (Coffey *et al.*, 2016, 2017; Giraud *et al.*, 2000). Thus, cortical sources may have contributed to responses to the 300 Hz but not the 700 Hz stimulus. Unfortunately, the extent to which asymmetric cortical contributions might have resulted in the asymmetry of effects for 300 *versus* 700 Hz FFRs in Chapter 4 cannot be determined from the single-channel

recordings. Smaller frequency deviants were used in Chapter 3 (500/600 Hz) and in Experiment 2 of Chapter 4 (817/700 Hz and 300/350 Hz), but Chapter 3 did not use a forward-reverse oddball, and too few trials were collected in Experiment 2 of Chapter 4 for reliable analysis of the FFR. Whether similar modulation, or lack thereof, of the FFR would be observed for small frequency differences remains an open question for future investigations.

5.2 Challenges and limitations

As briefly discussed in Chapter 3, the design of viable methods for simultaneous measurement of subcortical (i.e., the FFR) and cortical AEP components involves several practical considerations. For example, subcortical and cortical AEP components adapt to repeated acoustic stimulation at different rates. Whereas brainstem and midbrain sources are relatively unaffected by reductions in ISI down to about 10 ms (Ballachanda *et al.*, 1992), stark refractory effects in cortical neurons necessitate that cortical AEP components be evoked by relatively slow presentation rates (i.e., 0.5–2 s; Davis *et al.*, 1966; Picton *et al.*, 1977). Unlike cortical generators which are often proximal to the scalp surface, subcortical potentials must also propagate through cortical tissue. Hence, the amplitude of subcortical AEP components is orders of magnitude smaller than that of cortical components. For this reason, capturing subcortical AEP components at an acceptably high SNR requires averaging over a large number ($\sim$2000) of acoustic presentations. Together, these factors pose a trade-off between the optimal ISI for recording cortical activity and the time required for simultaneous collection of enough trials to permit detection and analysis of subcortical components (Skoe and Kraus, 2010).

Chapters 3 and 4 opted to use an ISI of 500 ms because pilot work suggested it provided an acceptable compromise between the amplitude of cortical change-detection components and the SNR of the FFR. In my experiments the amplitude of the obligatory cortical components (i.e., P1-N1-P2) was smaller than that typically reported in studies where stimuli are presented at ISIs greater than 1 s (e.g., Pereira *et al.*, 2014). This might, in part, explain why the correlation between N1 latency and FFR fidelity was observed for standard-evoked responses measured in Chapter 3 but not those measured in Chapter 4. Alternatively, this discrepancy might reflect differences in neural responsiveness to the spectrotemporal features of stimuli used in Chapter 3 *versus* Chapter 4 (i.e., carrier frequencies, modulation rates, or some interaction between these). The standard-evoked responses analyzed in Chapters 3 and 4 further differed in which trials were included in the averaging process. Whereas Chapter 3 included all standard trials not immediately following deviants, Chapter 4 only included standards immediately preceding deviants to ensure comparable trial counts for both types of presentation. It is not known if cortical refractoriness affected the relationship between P1-N1-P2 morphology and that of other AEP components in this thesis.

Recording the FFR itself is also rife with technical challenges. The Nyquist theorem dictates that accurate measurement of the FFR, up to its proposed upper frequency limit ($\sim$1.5 kHz), requires equipment capable of sampling scalp voltage at a minimum of twice the signal frequency. This requirement is comparatively relaxed for recording of cortical AEPs, as the energy of cortical components is typically contained below 100 Hz. Consequently, the use of dense electrode arrays, as prevalent in the cortical AEP literature, faces barriers related to data storage and management

when sampling rates are high enough to accurately capture the FFR. Indeed, many studies involving the FFR, including those described in Chapters 3 and 4, opt to record from a single electrode channel arranged in a vertical montage. Though dense-array recordings have shown that the FFR is strongest near vertex and largely fits to dipoles in midbrain sources (Bidelman, 2015), the scalp topography of cortical AEP components often carries meaningful information about possible generators which can vary between experimental conditions. Whether such differences existed between the components collected in Chapter 3 and 4 is unknown.

As is true of all steady-state components, recording the FFR also greatly benefits from minimal trial-to-trial variation (i.e., "jitter") between the arrival of stimulus energy at the tympanic membrane and the bit code (i.e., "trigger") denoting stimulus onset in the EEG trace. Ensuring minimal stimulus-to-trigger jitter often involves the use of calibrated ear-insert headphones and hardware capable of generating and/or tracking stimulus presentation with sub-millisecond precision. The use of commercially-available ear-insert headphones can be problematic because the proximity of the transducers to the recording electrodes introduces electromagnetic interference in the recording that overlaps with the spectrotemporal features of the FFR. Though it is possible to control stimulus artifact by separately averaging and summating the response evoked by stimuli of opposite polarities (i.e., phase shifted by 180°), doing so distorts the neural representation of temporal fine structure which would otherwise be captured by the FFR (Aiken and Picton, 2008). Such an approach is thereby impractical whenever the neural encoding of temporal fine structure is of interest to the experiment, as was the case for Chapters 3 and 4. Shielding the transducers in a grounded ferromagnetic metal casing (i.e., tinned-copper braided

tape) is another approach for controlling stimulus artifact. Though effective, care must be taken to ensure the shielding is complete and remains so throughout each recording session. Distancing the ear-insert transducers from the recording electrodes via an acoustic delay line (e.g., polyvinyl chloride tubing) is perhaps the most effective method of ensuring artifact-free recordings, and this is the approach I adopted although it requires correcting for the altered frequency-response introduced by the tubing with digital filters. Moreover, the bandwidth over which such filters can ensure a relatively flat frequency response becomes increasingly limited with the length of the delay line. Extensive piloting was required at the outset of my thesis to determine the the acoustic delay line that offered the best approach for collecting artifact-free FFRs without distorting the amplitude-modulated pure tone stimuli central to the methods of Chapters 3 and 4.

Some of the considerations outlined above can also preclude the use of identical approaches for collecting subcortical and cortical responses in infancy and adulthood. For example, wakeful infants tend to move abruptly and often and this introduces large amplitude myogenic artifacts into EEG recordings. I have observed this to be especially true in auditory experiments wherein the infant's primary caregiver is asked to remain silent for the duration of the recording session. Therefore, design of auditory EEG experiments involving wakeful infants should not assume more than 20 uninterrupted minutes of recording time before the subject becomes too restless or fussy to continue (DeBoer *et al.*, 2007). Presenting stimuli at sufficiently slow rates to evoke cortical AEPs might then yield too few artifact-free trials for visualizing brainstem components. Applying the SEAP method outlined in Chapter 3 early in development is further complicated by the relatively small amplitude of steady-state

components in infancy. The amplitude of the 40 Hz ASSR, in particular, is known to be much smaller and less reliable in infants compared to adults (Maurizi *et al.*, 1990; Suzuki and Kobayashi, 1984). Even the 80 Hz ASSR, which is more reliable in infancy compared to its 40 Hz counterpart, is small in amplitude and best recorded while the infant is asleep (Aoyagi *et al.*, 1993). The use of ear-insert headphones can also be problematic as infants respond negatively to the sensation of foreign objects in the ear canal and are prone to dislodge the ear-tips through abrupt movements of the head. In my pilot studies, I overcame this challenge with the use of a custom cap where the acoustic delay line was physically attached and directed into the infants ear by way of a strategically placed grommet. Still, over 40 attempts at collecting concurrent FFR and MMR/MMN recordings in 5-month-old ($n = 30$) and then 9-month-old ($n = 14$) wakeful infants, with variations of the methods developed in Chapter 3, ultimately proved unreliable and so the project was abandoned. Nevertheless, viable approaches to the concurrent measurement of subcortical and cortical AEPs in infants are of great value because they have the potential to better inform our understanding of the processes underlying auditory object formation and auditory learning.

5.3 Future directions

Previous developmental research has shown that subcortical function appears to be well-established at birth (Levi *et al.*, 1995). Conversely, AC is anatomically and functionally immature in early infancy (Eggermont and Moore, 2012; Kushnerenko *et al.*, 2002; Moore and Guan, 2001) and exhibits a protracted developmental trajectory that lasts into teenage years (Ponton *et al.*, 2000; Shahin *et al.*, 2004). As discussed in Chapter 3, the morphologies of many cortical AEP components recorded

in normal-hearing adults, including the MMN, N1, P2, and 40 Hz ASSR, and even subcortical FFR, are sensitive to long-term auditory experience and/or short-term training. This suggests that the underlying neural generators of these components stay neuroplastic into adulthood. Similar neuroplastic effects were observed to follow training in infancy. For example, Trainor *et al.* (2011) conducted a study in which caregivers were asked to play prerecorded melodies (either in harp or guitar timbre) to their 4-month-olds infants at home every day for one week (∼160 min of exposure). The experimenters found that the slow-positive MMR was enhanced when elicited by deviants in the trained timbre compared to the untrained timbre. In another study, Butler and Trainor (2013) found that in typically-developing 4- and 8-month-olds, MMN is not evoked by difficult discriminations involving iterated-rippled noise stimuli—sounds which elicit a pitch percept in the absence of resolvable spectral cues. However, the MMN can be evoked by these stimuli, in the same age groups, following only 8 minutes of passive exposure (i.e., training) when the stimuli are "primed" for a pitch percept by the addition of a sine wave component. Given the relative immaturity of the AC in the first few months after birth, it is reasonable to speculate that such auditory learning relies on stimulus-driven (i.e., bottom-up) processes originating in subcortical structures. Only after sufficient maturation of the AC would neuroplastic effects be expected to manifest in the enhancement of cortically-driven change-detection components. Thereafter, higher-level structures might begin to modulate feature encoding at subcortex via corticofugal (i.e., top-down) projections and produce the kinds of effects observed in Chapter 4. If the challenges reviewed in the preceding section can be overcome, methods such as SEAP offer promise as a means of studying how the relative contribution of these mechanisms and their interaction

change throughout early development. Developing a deeper appreciation of processes underlying early experience-dependent learning should be considered paramount to a stronger understanding of different auditory disorders, as auditory experience is critical to sculpting later auditory function (Sanes and Bao, 2009).

The SEAP method can also be useful for studying later stages of human development. Chapter 3 observed that the latency of the N1 component in normal-hearing adults was related to the fidelity of acoustic feature encoding by presumably subcortical generators of the FFR. The monkey analog of the human N1 is attributed primarily to a current source (hyperpolarization) in lamina III elements in A1 and partly from current sinks (depolarizations) in laminae I/II and V (Steinschneider *et al.*, 1994). Relative to deeper thalamocortical input layers, maturation of these superficial layers is prolonged in human development. Axonal neurofilaments in upper layers III and II are not observed until 5 years of age and these layers do not reach functional maturity until age 12 (Moore and Guan, 2001). This pattern of structural development parallels the emergence and asymptotic maturation of the N1 component, marked by a progressive decrease in N1 latency (Johnstone *et al.*, 1996; Ponton and Eggermont, 2001; Ponton *et al.*, 2000, 1999; Sharma *et al.*, 1997; Tonnquist-Uhlén *et al.*, 1996), as well as behavioural improvements in the perception of degraded speech (Eisenberg *et al.*, 2000) and speech in noise (Elliott, 1979) from 5 to 12 years of age. Further, maturation of superficial cortical layers and, ergo, development of the N1 component appears to depend on sufficiently patterned auditory stimulation during a critical period between 3 and 6 years of age (Eggermont and Ponton, 2003). Given that auditory information must first be represented at subcortical nuclei before reaching cortex, it is reasonable to speculate that the quality of subcortical representation also impacts

N1 morphology. To my knowledge no systematic study has measured the development of N1 concurrent with the FFR to determine how these components interact from 5 to 12 years of age and whether such interaction predicts speech intelligibility in adverse listening conditions. Understanding such a relationship could help with assessing the efficacy of clinical interventions (i.e., cochlear implants or hearing aids) in restoring spectrotemporal representations of stimulus features at subcortex in ways that facilitate acoustic feature integration and ultimately perception.

Finally, it is still unknown what role attention might have in coordinating change-detection processes at subcortical and cortical structures and how such coordination might operate to facilitate auditory learning in the short-term (i.e., based on task demands) or long-term (i.e., through training or extensive experience). In Chapter 4, significant deviant-related modulation of the FFR was only observed for a stimulus which also elicited a significantly earlier P3a component, suggesting this stimulus was associated with faster involuntary capture of attention (Escera *et al.*, 1998). Moreover, the latencies of P3a as well as MMN components elicited by this stimulus were also correlated with the strength of the deviant-evoked FFRs. However, in a pilot study I conducted (normal-hearing adults; $n = 10$), manipulating listener attention toward targets embedded in either an auditory or visual stream did not result in significant modulation of the FFR. This is in contrast to previous studies showing that the FFR is enhanced for spectral features of an attended vowel sound in a dichotic listening task (Galbraith *et al.*, 1998). Future work should manipulate listener attention between two oddball streams presented dichotically to assess whether modulation of FFR is observed for deviant features in the attended stream and whether this modulation

is associated with the overall morphology of change-detection responses generated downstream.

5.4 Concluding remarks

Ever increasingly, research in humans and other mammals demonstrates that auditory change-detection processes are not limited to cortical structures. This thesis demonstrates that concurrent recording of subcortically- and cortically-generated AEPs is a viable means of studying distributed change-detection in human listeners. However, there is still much to understand regarding how simple change-detection operates at subcortical nuclei and feeds forward to influence predictions of acoustic regularity in higher auditory structures. How these higher level structures might then encode and use abstract stimulus regularities to alter the response properties of lower level neurons is also largely unknown. Understanding the interaction of subcortical and cortical structures during the encoding of acoustic regularity not only offers great promise for advancing our models of auditory processing and perception, but might be critical to identifying and developing interventions for auditory processing disorders where such interaction is perturbed.

5.5 References

Aiken, S. J. and Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research*, **245**(1-2), 35–47.

Anderson, L. A., Christianson, G. B., and Linden, J. F. (2009). Stimulus-specific adaptation occurs in the auditory thalamus. *The Journal of Neuroscience*, **29**(22), 7359–7363.

Aoyagi, M., Kiren, T., Kim, Y., Suzuki, Y., Fuse, T., and Koike, Y. (1993). Optimal modulation frequency for amplitude-modulation following response in young children during sleep. *Hearing Research*, **65**(12), 253–261.

Ayala, Y. A. and Malmierca, M. S. (2013). Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Frontiers in Neural Circuits*, **6**, 1–19.

Ballachanda, B. B., Moushegian, G., and Stillman, R. D. (1992). Adaptation of the auditory brainstem response: effects of click intensity, polarity, and position. *Journal of the American Academy of Audiology*, **3**(4), 275–82.

Bidelman, G. M. (2015). Multichannel recordings of the human brainstem frequency-following response: scalp topography, source generators, and distinctions from the transient ABR. *Hearing Research*, **323**, 68–80.

Bosnyak, D. J., Eaton, R. A., and Roberts, L. E. (2004). Distributed auditory cortical representations are modified when non-musicians are trained at pitch discrimination with 40 Hz amplitude modulated tones. *Cerebral Cortex*, **14**(10), 1088–1099.

Butler, B. E. and Trainor, L. J. (2013). Brief pitch-priming facilitates infants discrimination of pitch-evoking noise: evidence from event-related potentials. *Brain and Cognition*, **83**(3), 271–278.

Chakalov, I., Paraskevopoulos, E., Wollbrink, A., and Pantev, C. (2014). Mismatch negativity to acoustical illusion of beat: how and where the change detection takes place? *NeuroImage*, **100**, 337–346.

Chase, S. M. and Young, E. D. (2008). Cues for sound localization are encoded in multiple aspects of spike trains in the inferior colliculus. *Journal of Neurophysiology*, **99**(4), 1672–1682.

Coffey, E. B. J., Herholz, S. C., Chepesiuk, A. M. P., Baillet, S., and Zatorre, R. J. (2016). Cortical contributions to the auditory frequency-following response revealed by MEG. *Nature Communications*, **7**, 11070.

Coffey, E. B. J., Musacchia, G., and Zatorre, R. J. (2017). Cortical correlates of the auditory frequency-following and onset responses: EEG and fMRI evidence. *The Journal of Neuroscience*, **37**(4), 830–838.

Davis, H., Mast, T., Yoshie, N., and Zerlin, S. (1966). The slow response of the human cortex to auditory stimuli: recovery process. *Electroencephalography and Clinical Neurophysiology*, **21**(2), 105–113.

DeBoer, T., Scott, S., L., and Nelson, A., C. (2007). Methods for acquiring and analyzing infant event-related potentials. In M. de Haan, editor, *Infant EEG and Event-Related Potentials*, pages 5–37. Psychology Press, New York, NY.

Eggermont, J. and Moore, J. K. (2012). Morphological and functional development of the auditory nervous system. In L. Werner, R. Fay, and A. Popper, editors, *Springer Handbook of Auditory Research: Human Auditory Development*, pages 61–105. Springer, New York, NY.

Eggermont, J. and Ponton, C. W. (2003). Auditory-evoked potential studies of cortical maturation in normal hearing and implanted children: correlations with changes in structure and speech perception. *Acta Oto-Laryngologica*, **123**(2), 249–252.

Eisenberg, L. S., Shannon, R. V., Martinez, A. S., Wygonski, J., and Boothroyd, A. (2000). Speech recognition with reduced spectral cues as a function of age. *The Journal of the Acoustical Society of America*, **107**(5), 2704–2710.

Elliott, L. L. (1979). Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability. *The Journal of the Acoustical Society of America*, **66**(3), 651–653.

Escera, C., Alho, K., Winkler, I., and Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, **10**(5), 590–604.

Fujioka, T., Trainor, L. J., Ross, B., Kakigi, R., and Pantev, C. (2005). Automatic encoding of polyphonic melodies in musicians and nonmusicians. *Journal of Cognitive Neuroscience*, **17**(10), 1578–1592.

Galbraith, C., G., Bhuta, M., S., Choate, K., A., Kitahara, M., J., and Mulen, T. A. J. (1998). Brain stem frequency-following response to dichotic vowels during attention. *NeuroReport*, **9**(8), 1889–1893.

Gander, P. E., Bosnyak, D. J., and Roberts, L. E. (2010). Acoustic experience but not attention modifies neural population phase expressed in human primary auditory cortex. *Hearing Research*, **269**(12), 81–94.

Giraud, A.-L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., and Kleinschmidt, A. (2000). Representation of the temporal envelope of sounds in the human brain. *Journal of Neurophysiology*, **84**(3), 1588–1598.

Grimm, S., Escera, C., and Nelken, I. (2016). Early indices of deviance detection in humans and animal models. *Biological Psychology*, **116**, 23–27.

He, C., Hotson, L., and Trainor, L. J. (2007). Mismatch responses to pitch changes in early infancy. *Journal of Cognitive Neuroscience*, **19**(5), 878–892.

Johnstone, S. J., Barry, R. J., Anderson, J. W., and Coyle, S. F. (1996). Age-related changes in child and adolescent event-related potential component morphology, amplitude and latency to standard and target stimuli in an auditory oddball task. *International Journal of Psychophysiology*, **24**(3), 223–238.

Kishon-Rabin, L., Roth, D. A.-E., Dijk, B. V., Yinon, T., and Amir, O. (2004). Frequency discrimination thresholds: the effect of increment versus decrement detection of frequency. *Journal of Basic and Clinical Physiology and Pharmacology*, **15**(1-2), 29–40.

Kushnerenko, E., Ceponiene, R., Balan, P., Fellman, V., and Näätänen, R. (2002). Maturation of the auditory change detection response in infants: a longitudinal ERP study. *Neuroreport*, **13**(15), 1843–1848.

Lau, J. C., Wong, P. C., and Chandrasekaran, B. (2017). Context-dependent plasticity in the subcortical encoding of linguistic pitch patterns. *Journal of Neurophysiology*, **117**, 594–603.

Levi, E. C., Folsom, R. C., and Dobie, R. A. (1995). Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hearing Research*, **89**(1-2), 21–27.

Marie, C. and Trainor, L. J. (2013). Development of simultaneous pitch encoding: infants show a high voice superiority effect. *Cerebral Cortex*, **23**(3), 660–669.

Marie, C., Fujioka, T., Herrington, L., and Trainor, L. J. (2012). The high-voice superiority effect in polyphonic music is influenced by experience: a comparison of musicians who play soprano-range compared with bass-range instruments. *Psychomusicology: Music, Mind, and Brain*, **22**(2), 97–104.

Maurizi, M., Almadori, G., Paludetti, G., Ottaviani, F., Rosignoli, M., and Lucianob, R. (1990). 40-Hz steady-state responses in newborns and in children. *International Journal of Audiology*, **29**(6), 322–328.

Moore, B. (1973). Frequency difference limens for shortduration tones. *The Journal of the Acoustical Society of America*, **54**(3), 610–619.

Moore, B. (1997). *Introduction to the Psychology of Hearing*. Academic Press, London, UK, 4th edition.

Moore, B. C. J. and Glasberg, B. R. (1996). A revision of Zwicker's loudness model. *Acta Acustica united with Acustica*, **82**(2), 335–345.

Moore, J. K. and Guan, Y.-L. (2001). Cytoarchitectural and axonal maturation in human auditory cortex. *JARO - Journal of the Association for Research in Otolaryngology*, **2**(4), 297–311.

Muir, D. and Hains, S. (2004). The U-shaped developmental function for auditory localization. *Journal of Cognition and Development*, **5**(1), 123–130.

Musacchia, G., Strait, D., and Kraus, N. (2008). Relationships between behavior, brainstem and cortical encoding of seen and heard speech in musicians and non-musicians. *Hearing Research*, **241**(12), 34–42.

Pereira, D. R., Cardoso, S., Ferreira-Santos, F., Fernandes, C., Cunha-Reis, C., Paiva, T. O., Almeida, P. R., Silveira, C., Barbosa, F., and Marques-Teixeira, J. (2014). Effects of inter-stimulus interval (ISI) duration on the N1 and P2 components of the auditory event-related potential. *International Journal of Psychophysiology*, **94**(3), 311–318.

Peter, V., McArthur, G., and Thompson, W. F. (2010). Effect of deviance direction and calculation method on duration and frequency mismatch negativity (MMN). *Neuroscience Letters*, **482**(1), 71–75.

Picton, T. W., Woods, D., Baribaeu-Braun, J., and Healy, T. (1977). Evoked potential audiometry. *Journal of Otolaryngology*, **6**(2), 90–119.

Ponton, C. W. and Eggermont, J. (2001). Of kittens and kids: altered cortical maturation following profound deafness and cochlear implant use. *Audiology and Neurotology*, **6**(6), 363–380.

Ponton, C. W., Moore, J. K., and Eggermont, J. (1999). Prolonged deafness limits auditory system developmental plasticity: evidence from an evoked potentials study in children with cochlear implants. *Scandinavian Audiology. Supplementum*, **51**, 13–22.

Ponton, C. W., Eggermont, J., Kwong, B., and Don, M. (2000). Maturation of human central auditory system activity: evidence from multi-channel evoked potentials. *Clinical Neurophysiology*, **111**(2), 220–236.

Sanes, D. H. and Bao, S. (2009). Tuning up the developing auditory CNS. *Current Opinion in Neurobiology*, **19**(2), 188–199.

Shahin, A., Roberts, L. E., and Trainor, L. J. (2004). Enhancement of auditory cortical development by musical experience in children. *Neuroreport*, **15**(12), 1917–1921.

Sharma, A., Kraus, N., J. McGee, T., and Nicol, T. G. (1997). Developmental changes in P1 and N1 central auditory responses elicited by consonant-vowel syllables. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section*, **104**(6), 540–545.

Skoe, E. and Kraus, N. (2010). Auditory brainstem response to complex sounds: a tutorial. *Ear and Hearing*, **31**(3), 302–324.

Steinschneider, M., Schroeder, C. E., Arezzo, J. C., and Vaughan Jr., H. G. (1994). Speech-evoked activity in primary auditory cortex: effects of voice onset time. *Electroencephalography and Clinical Neurophysiology*, **92**(1), 30–43.

Suga, N., Gao, E., Zhang, Y., Ma, X., and Olsen, J. F. (2000). The corticofugal system for hearing: recent progress. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(22), 11807–11814.

Suzuki, T. and Kobayashi, K. (1984). An evaluation of 40-Hz event-related potentials in young children. *International Journal of Audiology*, **23**(6), 599–604.

Tonnquist-Uhlén, I., Borg, E., Persson, H., and Spens, K. (1996). Topography of auditory evoked cortical potentials in children with severe language impairment: the N1 component. *Electroencephalography and Clinical Neurophysiology*, **100**(3), 250–260.

Trainor, L. J., Lee, K., and Bosnyak, D. J. (2011). Cortical plasticity in 4-month-old infants: specific effects of experience with musical timbres. *Brain Topography*, **24**(3-4), 192–203.

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, **13**(12), 532–540.

Zwicker, E. (1956). Die elementaren grundlagen zur bestimmung der informationskapazität des gehörs. *Acta Acustica united with Acustica*, **6**(4), 365–381.