Dual Representations of Temporal Modulations in Human Auditory Cortex

Huizhen Tang

ARC Centre of Excellence in
Cognition and its Disorders

Department of Cognitive Science

Macquarie University, Sydney, Australia

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

July 2015

Table of Contents

Thesis Summary

Our ability to understand speech and other sounds relies crucially on the capacity to detect and perceive temporal amplitude fluctuations in the range of about 1-100 Hz. However, most individual neurons in auditory cortex are capable of precisely aligning their activities only to modulation rates at the lower end of this range. This raises the question of how higher modulation rates might be encoded, and of how the auditory cortex might be organised to accommodate the full range of perceptually relevant amplitude envelope modulations.

Here we show, with noninvasive magnetoencephalography and electroencephalography measurements, that population oscillatory responses of human auditory cortex transition between a mode of strong phase locking to modulation rates below about 40-50 Hz, to a non phase-locked mode of responding at rates higher than about 50 Hz. Such dual response modes are predictable from the behaviours of single neurons in auditory cortex of non-human primates, but only the low rate phase locking mode has been previously observed in the neuronal population responses indexed in human MEG/EEG recordings.

Taken together, the single neuron and MEG and EEG results from current thesis work suggest that two distinct types of neuronal encoding are required to represent the full range of temporal modulation rates that are relevant to everyday perception.

Keywords: temporal processing, rate-based code, temporal code, human auditory cortex, magnetoencephalography, developing brai

Statement

I certify that the work in this thesis entitled "Temporal Representations of Amplitude Modulations in Human Auditory Cortex" has not been previously submitted for a degree nor it has been submitted as a part of requirements for a degree to any other university or institution other than Macquarie University.

I also certify that thesis is an original piece of research and it has been written by me. Any help and assistance that I have received in my research work and the preparation of the thesis itself have been appropriately acknowledged.

In addition, I certify that all information sources and literature used are indicated in the thesis. The research presented in the thesis was approved by Macquarie University Ethics Review Committee, reference number: HE23NOV2007-R05540 on 13th June, 2012.

Signed:

Huizhen Tang (Student Number: 42731232)

31st of July 2015

Acknowledgments

First and foremost, I would like to thank my primary supervisor, Distinguished Professor Stephen Crain, for his guidance, constant support, insightful suggestions, and generous allocation of time throughout my candidature. His passion and devotion in science has driven me to try my very best, especially in difficult times when I thought I have failed in many respects. Without him this thesis would not have been possible.

I am very grateful for the chance to work with A/Professor Blake W. Johnson, who has brought me into the field of neuroimaging and brain oscillations, which has become my wonderland and I would like very much to devote my career into.

Special thanks need to be delivered to Dr Michael Proctor, whose critical thinking and insightful comments have improved this thesis largely. Many thanks to Dr Jon Brock for his enthusiasm and interest in my work as well as his helpful comments.

I also would like to thank Robin Blumfield, Lesley McKnight and other members in the administrative team. They are the people who make the process so much easier for all the PhD students.

Further, many thanks need to deliver to colleagues at Macquarie University for their generosity in sharing their knowledge and expertise. I have benefited so much from the many inspiring discussions I had with them.

I would like to express my sincere appreciation to all the volunteers who tolerated the long testing sessions attentively and with great enthusiasm. I gratefully acknowledge the collaboration of Kanazawa Institute of Technology in establishing the KIT-Macquarie MEG laboratory.

Finally, I need to thank my family and friends for their unconditional love and support.

Chapter 1 - General Introduction

General Introduction

This thesis aims to investigate temporal processing in the auditory cortex in adults, using concurrent magnetoencephalography (MEG) and electroencephalography (EEG), and in preschool-aged children using a paediatric MEG system. Adopting an event-related spectral perturbation (ERSP) analysis, both phase-locked (PL) and non-phase-locked (NPL) brain activities were examined in order to separate the response associated with a stimulus-synchronized temporal representation and the response associated with a non-synchronized rate-based representation. This chapter reviews the current view of cortical processing of temporal modulations and the general findings from previous neurophysiological studies on temporal processing as measured using auditory steady-state responses (ASSRs) and envelope following responses (EFRs). This chapter also provides a general description of the three main methodological components of the thesis: 1) An event-related spectral perturbation (ERSP) analysis, which separately examines non-phase-locked (NPL) temporal processing and phase-locked (PL) temporal processing; 2) Concurrent MEG and EEG recording to validate the findings with different electrophysiological measurement and separate cortical activities from concomitant subcortical activity; 3) The use of a unique paediatric MEG system for measuring brain activity in the developing brain in preschool-aged children.

The main research questions that will be addressed in this thesis are presented at the end of the chapter.

1.1 Cortical processing of temporal modulations

Sound, including the sounds associated with speech, can be characterized in two dimensions: one is the spectral content of the sound, and the other is the temporal modulation of the sound as an acoustic signal (Smith et al., 2002, Joris et al., 2004, Baumann et al., 2015). As compared to the spectral content, the temporal modulation component of sound has been reported to

deliver more information relevant to the intelligibility of speech (Drullman et al., 1994a, Drullman, 1995, Smith et al., 2002), and impairments of temporal processing have been linked to language problems in several studies (Phillips and Farmer, 1990, Souza, 2000, Ben-Yehudah et al., 2004, Boets et al., 2007, Jorgens et al., 2008, Lehongre et al., 2011). Studies of temporal modulations usually use amplitude modulated (AM) stimuli, where a carrier is modulated by an envelope (Purcell et al., 2004, Lehongre et al., 2011, Miyazaki et al., 2013).

1.1.1 Neural encoding of temporal modulations

Unlike other sensory systems where sensory input can be static (e.g. the visual system), one unique property of the auditory system is that its sensory input is a time-varying variable (Rosen, 1992, Wang et al., 2008). The challenge is to discover how the human auditory system encodes the temporal information contained in an acoustic signal, which is constantly changing in time, and which determines how that signal is perceived as sequence of sounds that are perceptually distinct. Without sufficient temporal processing of the acoustic signal, the construction of accurate perceptions of the auditory world or meaningful linguistic representations of the speech signal might be impaired or even impossible (Yin et al., 2011). While there is widespread agreement on how temporal modulations are represented in the auditory periphery in both humans and animals, the evidence pertaining to cortical processing remains relatively elusive.

In healthy humans, an acoustic signal arrives at the ear and leads to vibrations of the cochlea fluid. The hair cells then transform these vibrations into electrical signals, which are subsequently conducted along the auditory nerve through to the brainstem and to other subcortical regions en route to the auditory cortex (Joris et al., 2004, Baumann et al., 2015). The auditory cortex is pivotal in the sense that it connects reciprocally with the auditory periphery, as well as with the frontal and parietal regions, where higher order processes take place. How the time-varying acoustic signal is mapped onto a spike train of cortical neurons is,

therefore, a question of particular interest for our understanding of the neural mechanisms underlying speech processing, as well as for our understanding of the processing of other non-speech sounds, such as music.

Throughout most of the auditory pathway leading to the auditory cortex, it has been consistently found that temporal modulation of an acoustic signal is directly encoded by the temporal patterns of neural activity via a temporal code (see Table 1 for definition) (Joris et al., 2004, Wang et al., 2008). At the level of the cortex, this view of the temporal encoding of the amplitude modulated (AM) stimuli has been updated by recent research in animals (Lu et al., 2001b, Liang et al., 2002, Wang et al., 2003, Joris et al., 2004, Yin et al., 2011). Single unit studies by Liang and colleagues (Liang et al., 2002) found a substantial population of A1 neurons of awake marmoset monkeys that change their firing rate as a function of modulation frequency (MF) without being synchronized to the modulation waveform. More recently, Yin and colleagues (Yin et al., 2011) recorded A1 neurons in awake macaques and found that most (37–78%) neurons were synchronized to at least one MF without exhibiting non-synchronized responses at other frequencies (Yin et al., 2011). This large percentage of neurons exhibiting synchronization to a modulation waveform was, however, accompanied by two other types of neurons: one type were exclusively non-synchronized neurons (7–29%), and the other type were mixed-mode neurons (13–40%) that synchronized to at least one MF and fired non-synchronously to at least one other MF. Taken together, these data suggest that, in addition to the stimulus-synchronized temporal code, a different, non-synchronized rate-based code, is also applied by the auditory cortex to encode temporal modulation (Lu et al., 2001b, Liang et al., 2002, Yin et al., 2011). The following section describes these two different encoding schemes for cortical processing of temporal modulations in more detail.

Table 1. Definitions of terms

Term	**Definition**
Neural code	the activity of a population of neurons
Neural representation	the activity of a single neuron
Temporal encoding	the representation of the temporal structure of sound by the modulated spike timing or firing rate of individual neurons or a population of neurons
Temporal representation	the coding of a stimulus by the very precise temporal structure of a spike train; also referred as stimulus-synchronized temporal representation
Temporal code	the encoding of a stimulus by a population of synchronised neurons with a temporal representation of the stimulus.
Rate-based representation	the coding of a stimulus by the firing rate of neuron; also referred as non-synchronised rate representation.
Rate code	the encoding of a stimulus by a population of non-synchronised neurons with a discharge rate-based representation of the stimulus.

1.1.2 Stimulus-synchronised temporal representation

Because time is an essential variable of acoustic inputs, it is not surprising that auditory research has emphasized the mapping of the temporal patterns of neural firing onto the temporal structures of the acoustic signal (Joris et al., 2004, Eggermont, 2014). One of the most prominent neural firing patterns for mapping acoustic temporal information has been found in the auditory system. The purpose of this firing pattern is to synchronize the neuronal firing with the onset of each acoustic event (e.g. each modulation cycle) – so this can be characterized as a stimulus-synchronized temporal representation. The left panel of Figure 1 shows an example of this temporal representation by synchronised neurons.

This stimulus-synchronised temporal coding is usually studied by analysing phase-locked (PL) neural activities. Such PL activity is evident throughout the auditory system and is prevalent at lower levels of the auditory system, with auditory nerve fibres showing precise PL rates up to 4 kHz (Wang et al., 2008, Eggermont, 2014). The upper limit of stimulus-synchronization in single units declines progressively and markedly as one ascends the auditory pathway (Wang et al., 2008, Eggermont, 2014). Auditory nerve fibers will phase lock to envelopes as high as 3-4 kHz, but in the cochlear nucleus, the PL cut offs lie between 750 and 1500 Hz (Rhode and Greenberg, 1994). In the inferior colliculus, the PL limit is reduced to 500 Hz, with a majority of cells losing the ability to follow the envelope above 200 Hz (Rhode and Greenberg, 1994). Studies using amplitude modulated (AM) stimuli in the auditory thalamus suggest a highest sensitivity below 50 Hz, with PL limits on the order of 100 Hz (Preuß and Müller-Preuss, 1990). Using click trains in the medial geniculate body (MGB), PL limits were observed between 100 and 300 Hz with peak sensitivity between 25 and 125 Hz (Bartlett and Wang, 2007). When the system ascends to the auditory cortex, the PL limit is reduced even further. The majority of neurons in the primary auditory cortex in different species showed a temporal best modulation frequency (tBMF) below 20 Hz (Joris et al., 2004).

This reduction of temporal representation along the ascending auditory pathway is considered a consequence of the biophysical properties of neurons and the temporal integration of converging inputs from one station to the next (Wang, 2007). For example, unlike onset responses, the responses to the periodicity of a continuous signal appear to depend more on the minimum spike latency, which determines if the neurons have sufficient time to 'reset' before the next modulation cycle. Therefore, a longer mean spike latency of the cortical neurons has been suggested as being one of the biophysical properties that causes the reduction of cortical capacity to synchronise to modulation waveform at higher rates (Liang et al., 2002).

Regardless of the underlying mechanisms, the sharply limited phase-locked neural capacity of the auditory cortex does not fit well with the perceptual capability of both human and animals, who are extremely competent in perceiving temporal information across a wide range of time scales. For example, both humans and animals are able to discriminate acoustic changes at time scales much shorter than 20 to 30 ms (Wang, 2007). There must be some way for cortical neurons to encode such rapid time-varying acoustic information. One possibility was proposed first by Bieser and Muller-Preuss (Bieser and Muller-Preuss, 1996), and later supported and extended by Lu and colleagues (Lu et al., 2001b, Liang et al., 2002, Wang et al., 2003), who suggested that stimulus-synchronised temporal representation alone is inadequate to represent the entire range of perceptually-relevant temporal information, especially when the system ascends to the thalamo-cortical stations. Besides temporal representation, these researchers argued that auditory cortex must represent temporal modulations at higher rates using an additional encoding strategy – a strategy does not rely on faithfully replicating the temporal structure of the acoustic signal but nonetheless preserves sufficient information for perception.

1.1.3 Non-synchronised rate-based representation

Recordings of A1 neurons in awake marmoset monkeys from Lu and colleagues exhibited stimulus-synchronised and non-synchronised two types of cortical responses to periodic click

trains (Lu et al., 2001b, Liang et al., 2002). In Figure 1, the left panel illustrates the stimulus-synchronised responses where the right panel shows the non-synchronised responses (Lu et al., 2001b). From the left panel of Figure 1, one can see that when the inter-click intervals (ICIs) were longer than about 25 ms (40 click per second), synchronised neurons exhibited significant stimulus-synchronised responses to click trains, but these responses diminished when click trains were presented with shorter ICIs. In contrast, the spiking pattern presented on the right panel of Figure 1 indicated that the discharging rate of this type of neurons increased with acoustic stimulus, but with no synchronisation to the periodicity of the click train.

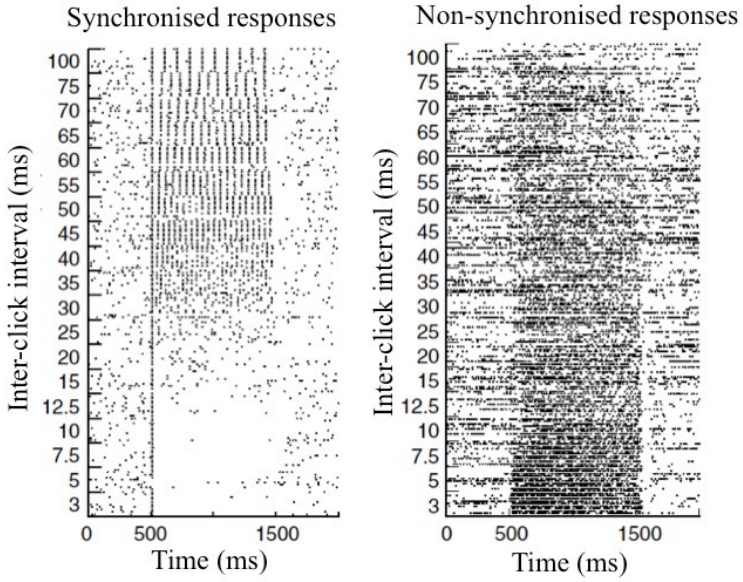

Figure 1. Two distinct types of cortical responses to periodic click trains. The left and right panel shows examples of stimulus-synchronised and non-synchronised responses, respectively, to click trains recorded from primary auditory cortex (A1) of awake marmosets. The stimulation was presented between the time window of 500 ms to 1500 ms in the X-axis. Figure courtesy of Wang et al., (2007).

Further examination showed that the responses of these two types of neuron populations change as a function of the inter-click interval (ICI) (Lu et al., 2001b, Wang, 2007). Specifically, neurons in the synchronized population were responsive and showed stimulus-synchronised discharges at long (ICIs, but few responses at short ICIs, whereas the non-synchronised sustain firing population appeared to behave in an opposite way, in which more this type of neurons responded to short ICIs and less to long ICIs (see Figure 2). These results

suggest that the stimulus-synchronised neuron populations represent slow temporal modulations explicitly using a temporal code; and the non-synchronised population represent fast modulations implicitly using discharging rate-based code. The synchronised representation was termed 'isomorphic', which provides a faithful replica of the temporal modulation, whereas the non-synchronised representation is 'non-isomorphic', which means the temporal modulation has been converted into an internal representation that is no longer a faithful replica of the temporal modulation. The combination of these two types of cortical neuron populations, therefore, provides a dual encoding mechanism that allows the auditory cortex to represent a wide range of temporal modulations that are perceptually relevant (Lu et al., 2001).

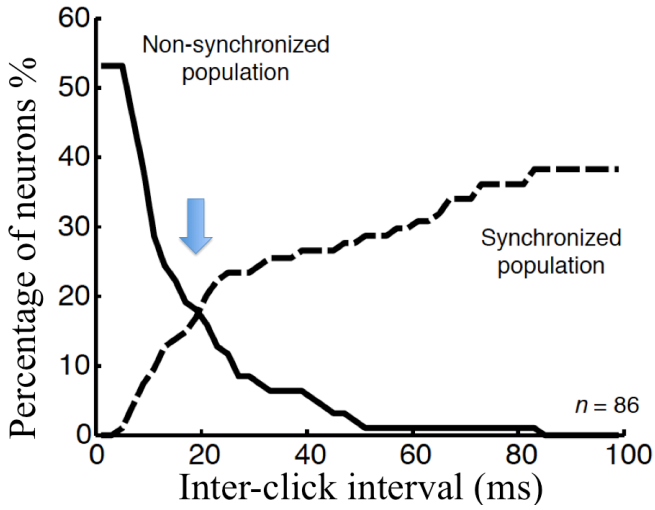

Figure 2. Dual encoding mechanisms for encoding slow and fast repetition rates or temporal modulations in the auditory cortex. These dual mechanisms are provided by a combination of stimulus-synchronised neurons and non-synchronised neurons, which are capable of encoding acoustic periodicity either using an isomorphic faithful representation or a non-isomorphic internal representation, respectively. The blue arrow indicates the point where the neuronal activity is dominated by synchronised neurons transitioning to one dominated by non-synchronised neurons as the ICIs become shorter.

1.1.4 Temporal-to-rate transform and its functional role

The results described above indicate that neurons in marmoset auditory cortex apply a dual encoding scheme to AM sounds. This dual encoding scheme is achieved with two distinct populations of neurons using either a synchronised temporal code or a non-synchronised rate-based code to represent acoustic periodicity (Lu et al., 2001b, Liang et al., 2002, Wang et al.,

2003). Moreover, the synchronisation-based neurons seem to be responsive to slow rates, whereas the non-synchronised neurons seem to be responsive at faster modulation rates. A comparison of the percentages of these two types of neurons reveals that temporal encoding makes a transition from being dominated by synchronisation-based temporal codes to one that is dominated by non-synchronised rate-based codes around 45 – 50 Hz (corresponding to ICI = ~ 20 – 25 ms, see Figure 2). In other words, when the repetition rate is below 50 Hz, the temporal modulation is mostly encoded with a temporal code, but, as the repetition rate increases, the temporal modulation becomes mostly encoded using a rate-based code.

Lu and colleagues (2001a) demonstrated a potential function of rate coding in discriminating acoustic transients using two different types of temporally modulated signals, which they called ramped and damped sinusoids. These researchers found that some neurons responded almost exclusively to one type (ramped or damped) of modulated signals with a segment length of 25 ms. However, this selectivity of responsiveness was mostly observed in neurons showing a non-synchrony rate-based representation, but not in the synchronised neurons (Lu et al., 2001a). This led Lu and colleagues to suggest that a discharge rate-based representation is necessary for discriminating the acoustic changes in rapidly changing signals (with a length of 25 ms). Wang and colleagues (2007) compared the asymmetric responsiveness to damped and ramped modulation of A1 neurons in marmoset monkeys with humans' psychophysical performance in discriminating these two types of modulations (Patterson, 1994). They found that the percentage of A1 neurons in marmosets was closely associated with the psychophysical performance in humans. These findings led them to conclude that discharge rate-based representation in the auditory cortex serve to discriminate rapid acoustic transients without involving stimulus-synchronised temporal representations of those transients.

One may question why the auditory system would go to the trouble of representing precise timing information of acoustic signals in the auditory periphery, especially the auditory nerve, but only utilise that information at higher levels with low-fidelity rate-based codes.

Wang and colleagues (2007) suggested that this transformation of encoding is necessary to achieve multi-sensory integration, in view of the fact that other sensory signals (e.g., visual) do not usually vary as rapidly as do auditory signals, resulting in much slower peripheral representations of sensory signals than that of auditory signals. At the level of cerebral cortex, however, the integration of neural representations of different sensory signals across multiple modalities needs to take place within the same time frame (Mauk and Buonomano, 2004). To this end, the cerebral cortex must slow down the much faster auditory signals coming from the periphery, by converting them into non-synchronised responses (Wang, 2007). This seems to be in agreement with a fundamental principle of cortical processing (Goldstein et al., 1959, Whitfield and Evans, 1965). Specifically, this principle relies on the transformation of stimulus features into internal neuronal representations that are no longer faithful replicas of the temporal structures of the acoustics (Goldstein et al., 1959, Whitfield and Evans, 1965, Whitfield, 1980, Wang et al., 2008).

1.1.5 Temporal processing in human auditory cortex

As mentioned earlier, the evident limitation of phase locked (PL) brain activity in the auditory cortex stands in contradiction with the remarkable ability of both humans and animals to detect temporal changes in rapid acoustic signals (Wang, 2007). Single unit studies have shown that the marmoset auditory cortex is likely to apply a dual encoding scheme to achieve this. This dual encoding scheme is provided by two distinct populations of neurons, one using a synchronised temporal code and the other using a non-synchronised rate-based code to represent acoustic periodicity (Lu et al., 2001b, Liang et al., 2002, Wang et al., 2003). Despite the converging evidence in signal unit studies in animals, the manifestation of such dual temporal representations in human auditory cortex is unclear. This is because studies of temporal encoding in humans has relied mainly on measurements of evoked responses that are tightly phase-locked to temporal modulation in sounds.

Neurophysiological research in humans can be conducted either using non-invasive techniques such as EEG/MEG for investigating auditory cortical processing with a high temporal resolution (Linden et al., 1987, Lins and Picton, 1995, John and Picton, 2000, Roß et al., 2000, Roß et al., 2002, Picton et al., 2003, Purcell et al., 2004, Lehongre et al., 2011, Tlumak et al., 2011). A commonly used EEG/MEG measure of auditory temporal processing is the auditory steady-state response (ASSR), a brain response that is evoked by periodic acoustic stimuli (Regan, 1989). The frequency component of the ASSR is stable in both amplitude and phase and sustained over a long temporal window (Regan, 1989). Brain responses can also elicited with acoustic stimuli that "sweep" over a range of modulation rates. Since these brain responses are not "steady" they are referred to by the more general moniker of "envelope following response (EFR; Regan, 1989, Purcell et al., 2004). EEG/MEG recorded ASSRs/EFRs provide a non-invasive measurement of how the periodicity of an acoustic signal, such the temporal modulation at a fixed rate or sweeping over a range of rates, is mapped onto the dynamics of brain oscillations (Regan, 1989, Levi et al., 1995, Ross et al., 2000, Picton et al., 2003, Nodarse et al., 2011, Picton, 2013).

Temporal Modulation transfer functions (TMTFs) are used to characterise the properties of ASSRs/EFRs. TMTFs depict a system's sensitivity to amplitude modulations of an acoustic signal as a function of the modulation rate or periodic frequency (Viemeister, 1979). Also, as the term TMTF implies, the temporal modulations are represented with a synchrony-based temporal code. In single unit studies or electrophysiological studies in animals, TMTFs are often related to another term, the rMTF, which stands for rate modulation transfer functions. rTMFs quantify the response to AMs as a function of MF when the AMs are represented with a rate-based code (Joris et al., 2004). In auditory studies in humans, the use of rMTF is rare.

TMTFs of ASSRs/EFRs in human adults consistently exhibit a low-pass/band-pass filter profile, respectively. This filter profile is characterised by a best modulation frequency (BMF) occurring at about 40 Hz. The strength of the response begins to decline steadily at frequencies

higher than about 50 Hz (Picton et al., 2003, Purcell et al., 2004, Poulsen et al., 2007, Miyazaki et al., 2013). Figure 3 shows the TMTF of ASSRs summarised from multiple studies (Picton et al., 2003) and TMTF of EFRs of a single study (Purcell et al., 2004).

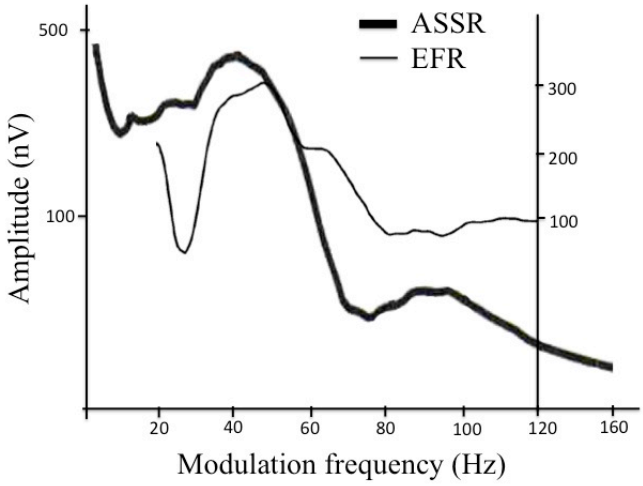

Figure 3. TMTFs of ASSRs and EFRs. The curve shown with thick line is re-produced from Picton et al., (2003). This curve was generated from multiple TMTFs of ASSRs from a list of various data sets. For the exact data sets and the method of generating this smooth curve, please see Picton et al., (2003). The thin curve is re-produced from Purcell et al., (2004) showing the grand average of EFRs to 100% modulated white noise from 10 waking individuals. The modulation frequency range covered in the EFRs is from 20 to 100 Hz.

The TMTF profiles of ASSRs/EFRs in the auditory cortex are consistent with a hierarchical filter bank organisation in the auditory system (Giraud et al., 2000). Giraud and colleagues (2000) conducted an fMRI study to investigate the representation of temporal modulations in the auditory system using sinusoid amplitude-modulated white noise at fixed frequencies between 4 to 256 Hz. The main structures in the central auditory system all showed responses to the AMs, including the brainstem, inferior colliculus, the medial geniculate body, Heschl's gyrus (HG), the superior temporal gyrus, the superior temporal sulcus, and the inferior parietal region. The preferential AM rates at each level of structure demonstrated a hierarchical filter bank organisation, in which lower levels yielded faster preferential AM rates. For instance, in the brainstem, the preferential AM was 256 Hz; in the primary auditory cortex, the preferential AM was 8 Hz. Interestingly, Giraud and colleagues found that temporal

modulations were processed by distinct (hierarchically organised series of filters) structures along the auditory pathway, as well as by common neural substrates in the same cortical region, where they found different responses to high and low AM frequencies. The finding of the same cortical region responding to high and low AMs differently might be related to the possibility of a dual encoding mechanism in the auditory cortex in humans. However, due to the poor temporal resolution of fMRI, few conclusions can be drawn in this regard.

Recently, ECoG studies conducted on epilepsy patients have shown in addition to the typical evoked ASSRs and EFRs usually reported in EEG/MEG auditory studies, induced power at high gamma band (70-110 Hz) were associated with processing temporal information (Steinschneider et al., 2008, Brugge et al., 2009, Nourski et al., 2013). For example, Brugge and colleagues (Brugge et al., 2009) used click trains presented at different rates between 4 and 200 Hz to examine the capacity of primary auditory cortex to encode repetitive transients in epilepsy patients using ECoG. In addition to the conventional measure of averaging method based auditory evoked potential (AEP), they also examined the event-related band power (ERBP), which is a measure based on the power change relative to the defined pre-stimulus baseline occurring on a single trial, including both phase-locked and non-phase-locked power within the frequency range of interest. They found both robust auditory evoked potentials (AEPs) to each individual click along with a frequency following response (FFR) in the posteromedial Heschl's gyrus. Moreover, event-related band power (ERBP) showed stimulus-related increases in gamma band frequencies as high as 250 Hz. Further analysis of induced power revealed a long-latency response in the belt region responding to the higher click rates between 100 Hz and 200 Hz (see Figure 4). AEP to 25 Hz click trains showed distinct onset responses to each click which were not unidentifiable as the rate increased. Similarly, the FFR was clearly observed at click rate about 100 Hz or below, and prominent at rates below about 50 Hz, but strongly attenuated at higher rates. The time course of AEP possessed an isomorphic relationship with temporal structure of the click trains and, therefore, can be considered to

explicitly represent the inter-click intervals. In contrast, the long-latency non-phase-locked gamma band activity did not show such an isomorphic relationship. Brugge and colleagues (2009) suggested that the gamma band non-phase-locked responses found in both the core and belt regions indicates a transformation from an isomorphic to a non-isomorphic representation of the temporal structure of click trains comparable to the transformation from synchrony-based temporal representation to rate-based representation in marmoset monkeys' primary auditory cortex (Lu et al., 2001b). Furthermore, the non-synchronised neuron populations, which undergo rate-based coding, exhibited sustained responses to click trains with firing rates that were monotonic functions of the click rate (Lu et al., 2001b). In other words, although the rate-based representation is non-isomorphic, the firing rates of the neurons increased as the click rate increased. Consistent with this, the non-phase locked (NPL) gamma band power showed a monotonic increase as the click rate increased and this was in company with the decrease in power of the FFR. These results point to the existence of a non-synchronised rate-based representation at higher click rates. However, the NPL gamma responses in this study occurred in a broad band of frequencies from about 20 Hz up to about 150 Hz with no clear upper cut-off. The nature and origin of broadband gamma is not clear, but a proposed interpretation considers that it is the summation of a large number of band-limited oscillatory responses (Crone et al., 2011).

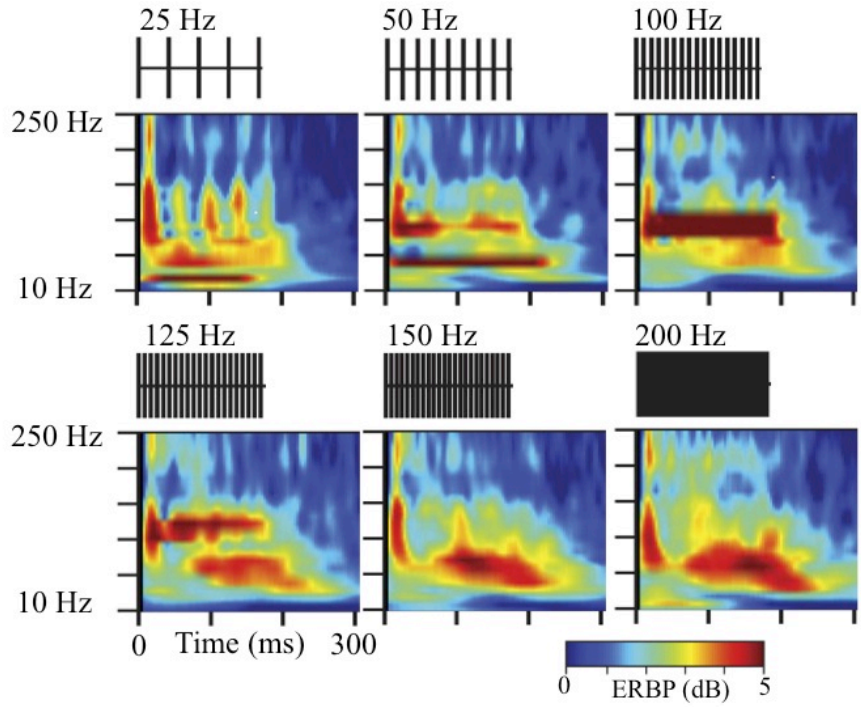

Figure 4. Time-frequency data recorded from a representative ECoG depth electrodes placed in the HG region when click trains were presented at different rates between 25 Hz and 200 Hz. The time-frequency representations were computed using a wavelet transform based on complex Morlet wavelets. ERBP was computed relatively to a 300 ms pre-stimulus baseline. Figure is reproduced from Brugge et al. (2009).

1.2 Development of temporal processing in the auditory cortex

A large body of behavioural evidence suggests that auditory processing in typically developing children may be quite different than that in adults. However, little research has been devoted to the cortical processing of temporal modulations in the developing brain. Moreover, most of the electrophysiological research on children has been conducted with either infants or school-aged children. Little is known to date about the maturation of cortical temporal processing in preschoolers, i.e., children aged from 3 to 5 years. Nevertheless, this is a crucial period of dramatic development in speech and language.

1.2.1 Behavioural evidence

Psychophysical studies investigating children's temporal processing usually use paradigms such as gap detection, backward masking, and modulation detection. In general, such behavioural evidence indicates a protracted maturational time course for temporal processing (Irwin et al.,

1985, Wightman et al., 1989, Hall and Grose, 1994, Trehub et al., 1995, Bishop et al., 1999, Hill et al., 2004, Dawes and Bishop, 2008, Banai et al., 2011, Moore et al., 2011, Fox et al., 2012, Buss et al., 2014). For example, using a gap detection task, Trehub and colleagues (Trehub et al., 1995) studied the sensitivity of infants (6.5 months and 12 months) and young children (5 years) to discriminate a pair of tones with a gap inserted between them. Different gap durations from 8 ms to 40 ms were examined. Infants and children performed much poorly than adults on gap durations of 12 and 16 ms. Likewise, using psychophysical temporal modulation transfer functions to study the detection of amplitude modulations (AM) in children from 4- to 10-years-old, Hall and Grose (1994) demonstrated that the sensitivity of 4- to 7-year-old children were lower than adult controls in response to modulations at rates between 5 Hz to 200 Hz. Significant differences were also found between the group of 4-5 year-olds and the group of 9-10 year-olds. These findings were interpreted as evidence that auditory processing of temporal modulations does not reach mature levels until mid- to late-childhood (Hall and Grose, 1994). This conclusion was further confirmed in a study conducted by Moore and colleagues (Moore et al., 2011), who evaluated the auditory processing of a large number of 6- to 11-year-old children using multiple paradigms, including simultaneous masking, backward masking, amplitude and frequency modulation detection.

In short, children consistently demonstrate poorer auditory temporal resolution than adults in psychophysical studies (Irwin et al., 1985, Wightman et al., 1989, Hall and Grose, 1994, Trehub et al., 1995, Bishop et al., 1999, Hill et al., 2004, Dawes and Bishop, 2008, Banai et al., 2011, Moore et al., 2011, Fox et al., 2012, Buss et al., 2014). Although much about the cortical processing in the brain can be learnt from behavioural performance, psychophysical measures in children are much less reliable than they are in adults due to children's difficulty in complying with experimental requirements (Bishop, 2007). Consequently, it is important to obtain direct and objective neurophysiological data on temporal processing in the developing brain, in addition to obtaining reliable behavioural measures. The understanding of temporal

processing in the developing brains relies on the converging of behavioural evidence and neurophysiological findings.

1.2.2 ASSRs and EFRs in children

Compared with the massive body of research devoted to the study of auditory processing in adults, as measured using ASSRs or EFRs, the number of similar studies in children is relatively modest (Rickards et al., 1994, Levi et al., 1995, Picton et al., 2003, John et al., 2004, Nodarse et al., 2012, Tlumak et al., 2012, Muhler et al., 2014). Most of these studies were conducted in infants or very young children when they were either in sleep or sedated, which could supress the cortical response even in adults (Pethe et al., 2001). It is now well established that the ASSR at 40 Hz consistently found in adults is almost entirely absent in infants. Specifically, the amplitude of the infant ASSR has been found to be monotonically decreasing as the rate of modulation or repetition increases, without showing the typical peak around 40 Hz evident in adults (Picton et al., 2003). Interestingly, although the amplitude of ASSRs to slower rates is attenuated in infants, their responses to higher rates around 80 Hz are much closer to that of adults (John et al., 2004, Tlumak et al., 2012). For example, Levi and colleagues (1995) recorded EFRs to amplitude-modulated tones in 1-month old infants with EEG. They found that the best modulation frequency (BMF) was 80 Hz for infants and 40 Hz for adults. Although there were some age differences in the EFRs, the 80 Hz ASSRs were largely adult-like in children who were 1 month of age. Considering the abundant psychophysical evidence that children have lower temporal resolution, it seems unlikely that the infant auditory cortex possesses such a higher temporal resolution.

In fact, scalp recorded EEG responses to slow rates reside primarily in the auditory cortex in children, whereas adult responses to higher rates primarily involve subcortical (brainstem) responses (Herdman et al., 2002, Picton et al., 2003, Purcell et al., 2004). This suggests that the difference between children's best modulation frequency (BMF) at a higher frequency, as compared to the lower BMF for adults, may simply reflect the earlier maturation

of brainstem responses in children, as compared to their cortical responses (Moore, 2002, Joris et al., 2004, Moore and Linthicum Jr, 2007).

As mentioned above, the contribution of cortical responses in the infant studies may have been further reduced in previous studies because measurements were made when infants were either asleep or sedated (Goldstein et al., 1959, Cohen et al., 1991, Lu et al., 2001a, Pethe et al., 2001). This conclusion is supported by a recent EEG study that examined the ASSR in children aged 6-9 years who were asked to maintain their attention to trains of tone-burst at a range of repetition rates during the recording (Tlumak et al., 2012). Although the ASSRs at 20 and 40 Hz in children were smaller than those of adults, the magnitudes of responses at low repetition rates (i.e. 0.75 Hz, 1.25 Hz, 2.5 Hz, and 5 Hz)in school-aged children were larger than those obtained in adults (Tlumak et al., 2012). These findings might be an indication of different developmental trajectories of temporal processing in the cortex and in peripheral regions (Moore, 2002).

1.3 Perception of temporal modulation

The perception of the temporal modulation in AM sounds is highly dependent on the MF as well as on the types of carriers. The focus of much current work is on the perceptual effect of MF (Joris et al., 2004). The effect of carrier is beyond the scope of this thesis. Therefore, we will only consider the perception of AM sounds when the carrier is noise, which provides no obvious spectral information.

1.3.1 Perceived temporal modulation

The MF determines what perceptual sensation we would experience when we are presented with an AM noise. Generally, when the modulation occurs slowly, with a rate below about 8-10 Hz, each modulation is perceived as an individual event with clear boundaries between the modulations (Joris et al., 2004, Nourski and Brugge, 2011). As the MF increases to above about ~10, however, the AM noise starts to evokes a sensation of “flutter”, which is the result of

losing the boundaries between individual events (Miyazaki et al., 2013). Additionally, another sensation, a type of "roughness", is also reported within this ~10 to ~50 Hz range (Fishman et al., 2000, Eggermont, 2001). This sensation of flutter or roughness continues until the MF goes above ~ 50 Hz. Pitch is perceive at modulation rates of ~50 – 500 Hz (Joris et al., 2004, Nourski and Brugge, 2011). Figure 5 illustrates the correspondence between the MF and the perception of AMs in both non-speech sounds and natural speech. We will have more to say on speech perception in the next section. It is important to keep in mind, at this point, that the boundaries between different perceptual classes described here are variable between subjects and between studies, largely due to the subjective nature of auditory perception.

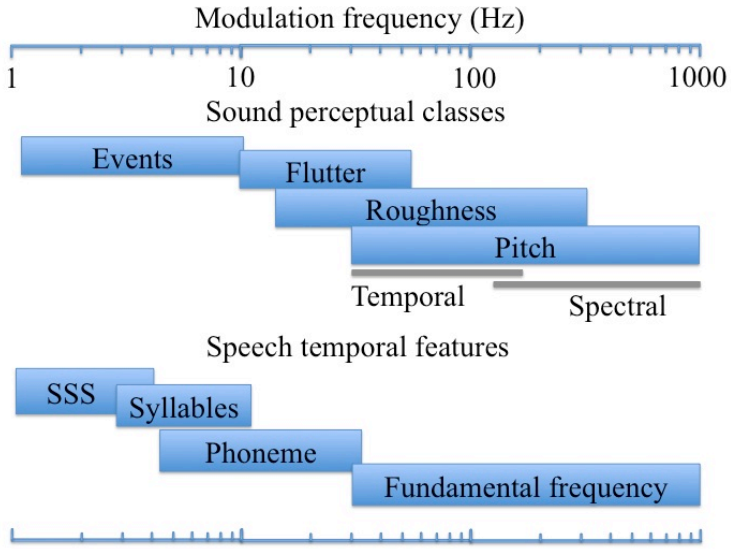

Figure 5. Scale of perceived temporal modulation in non-speech and speech. The top half illustrates the perceptual classes of perceived temporal modulations in non-speech; the bottom half illustrates the potential correspondence between temporal modulations and linguistic units. This figure is reproduced from Nourski & Brugge (2011). SSS stands for "supra-syllabic sequences".

Among these different types of percepts, of interest are intermediate percepts, flutters and roughness (see Figure 5), which characterise the transition from the perceptually distinguishable individual events to the perception of well-defined pitch. This perceptual transition frequency range (~25 to ~ 70 Hz) has been linked to the transition of temporal encoding from a synchrony-based temporal coding scheme to a rate-based coding scheme that

emerges in the thalamus-cortical regions of the auditory system (Wang et al., 2003; Poeppel et al., 2008; Nourski and Brugge, 2011). The sensation of 'roughness' is therefore thought to result from the co-existence of stimulus-synchronised and non-synchronised representations (Eggermont, 2001, Poeppel et al., 2008).

1.3.2 Relation to speech perception

Speech contains a large variety of complex temporally varying features, including both periodic and aperiodic components, and both frequency and amplitude modulations. Rosen (1992) provided a framework for partitioning the temporal structure of speech based on its dominant temporal modulations in three different dimensions – envelope, periodicity and fine-structure – in relation to their acoustic, auditory and perceptual and linguistic correlates (Rosen, 1992). Rosen defined envelope is the overall amplitude fluctuation at frequency between 2 and 50 Hz, while periodicity was used to distinguish the speech signal as periodic versus aperiodic. As mentioned by Rosen (1992), his use of "envelope" differs from the notion of "envelope" that is derived from the analytic signal, which typically contains both "envelope" and "periodicity" information (Smith et al., 2002). Since many studies used "envelope" as derived from analytic signal processing terms (e.g., Drullman, 1995, Smith et al., 2002, Aiken and Picton, 2008, Ghitza et al., 2013, Kubanek J et al., 2013), therefore, from here forward, we will use the term "amplitude envelope" to refer to Rosen's concept of "envelope" and we will use the simple term "envelope" to refer to the analytic signal term which consists of both amplitude envelope and periodicity.

Potential roles of various temporal features in linguistic contrasts were proposed and a list of correspondences between temporal features and linguistic information was provided (see Table 2 for details) in Rosen's framework (1992). Table 2 reveals that the prosodic cues, including intonation, stress, syllabicity and rhythm, are either delivered through amplitude envelope or through the periodicity of speech. Fine structure contributes little (if anything) to

prosodic perceptions. For segmental cues, envelope is important in the perception of manner and voicing, whereas fine structure contains information crucial for place and voice quality. Taken together, as compared to fine structure, envelope plays a more important role for speech intelligibility. This has been demonstrated in several studies using different techniques to manipulate these two fundamental dimensions of speech signals (Drullman et al., 1994b, a, Drullman, 1995, Smith et al., 2002).

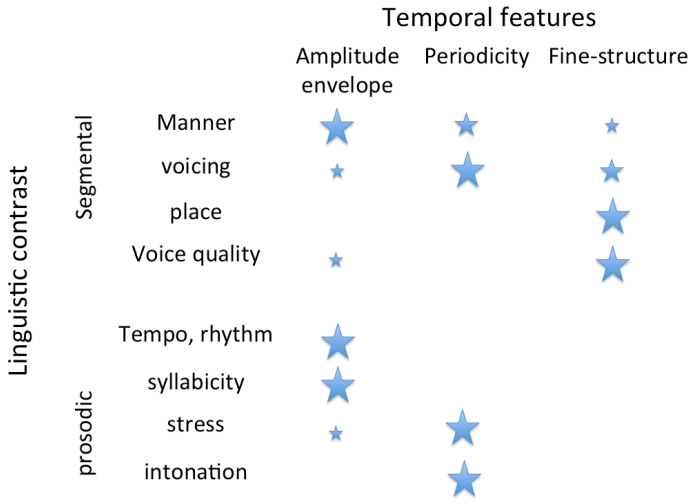

Table 2. List of correspondences between temporal features and linguistic contrasts. This table is reproduced from Rosen (1992). The size of the stars indicates the extent to which a particular feature operates in a particular linguistic contrast and a blank space indicating very weak or non-existent cues.

Rosen's framework suggests the complexity of the temporal features of envelope and their corresponding roles in linguistic contrasts. In line with this, several recent hypotheses for speech perception have highlighted the role of multiple temporal structures within the speech envelope (Rosen, 1992, David and Shamma, 2013, Goswami and Leong, 2013). For example, Goswami and Leong (2013) identified three temporal tiers as most dominant in speech: a stress AMs (<2.5 Hz); a syllable AM (2.5 – 12 Hz); and a phoneme AM (12 – 40 Hz). Moreover, it has been suggested that no single level of modulation can extract all the relevant information from speech (Zeng et al., 2005). Similarly, the 'asymmetric sampling in time' (AST) model proposed by Poeppel (2003) suggests there are two relevant time scales: a medium-duration time window (about 150-300 ms) which is relevant for encoding syllables; and a short time

window (~20 – 40 ms) which relevant for encoding formant transitions in stop consonants. The corresponding frequency range for the short time window is roughly 25 to 50 Hz, which falls into the MF range where AM is perceived with a sensation of flutters or roughness. From Table 1, we can also see that the envelope in this frequency range in speech sound conveys temporal information that is critical for manner of articulation, voicing, vowel identity, and prosody (Rosen, 1992, Nourski and Brugge, 2011). Despite the direct associations in these models between speech envelope and the perception of linguistic information, the underlying neural mechanisms linking these temporal features in speech signal to the perception of corresponding linguistic units are mostly unknown (Zeng et al., 2005, Ding and Simon, 2012, Pasley et al., 2012, Gaucher et al., 2013, Steinschneider et al., 2013).

1.4 Methodology

The current thesis recorded brain activities during stimulation using both MEG and EEG. Data analyses were performed using event-related spectral perturbation (ERSP) analysis adopted from Makeig (1993). In this methodology section, we first describe the general analytic method that was used. Then we introduce the EEG and MEG systems that were used to record brain activity. The use of MEG and EEG was motivated by three considerations: (1) the use of non-invasive measurements of brain activities in healthy humans; (2) a high temporal resolution (in this case, 1 ms); (3) the measurement of complementary electrophysiological signals generated from overlapping but not identical brain regions.

1.4.1 Analytic methods for measuring oscillatory responses

1.4.1.1 Brain oscillations

In the past decades, converging evidence has shown that brain oscillations play a crucial role in brain functions (Başar et al., 1997, Schurmann and Basar, 2001, Sauseng et al., 2007, Mazaheri and Jensen, 2008, Dugue et al., 2011, Thut et al., 2011, Giraud and Poeppel, 2012, Kraus, 2012,

Peña and Melloni, 2012, Thut, 2014). It has been shown that the prominent alpha rhythms (~ 10 Hz) originally discovered by Hans Berger encompass three main functionally and spatially distinct oscillations: the classic posterior alpha rhythm, the rolandic mu rhythm, and a temporal rhythm – "tau rhythm" (Jansen and Brandt, 1991, Başar et al., 1997, Niedermeyer, 1997, Obleser and Weisz, 2012). In addition to speech processing, these different subtypes of alpha rhythms are associated with memory mechanisms, sensory responses, and motor processes (Başar et al., 1997, Obleser and Weisz, 2012, Strauß et al., 2015). For example, recent work by Strauß and colleagues (2015) showed that the pre-stimulus phase distribution of alpha oscillations differentiated between correct and incorrect lexical decision in a speech-in-noise experimental setting, which provides a strong link between phase modulation of alpha oscillations and sensory processing.

Another class of brain rhythms that has received intensive research interest is the gamma band rhythms (> ~ 30 Hz). In the auditory system, a 40-Hz transient oscillatory response to acoustic stimuli has been recorded in both MEG and EEG in the humans (Galambos et al., 1981, Makela and Hari, 1987, Pantev et al., 1991, Galambos, 1992, Ross et al., 2005, Ross, 2008). This 40-Hz gamma band response is considered to be associated with auditory perceptual processing (Pantev et al., 1991). In addition, non-phase-locked gamma band oscillatory responses were also observed during activation of auditory cortex (Pantev, 1995) and visual cortex (Tallon-Baudry et al., 1996). High broadband gamma oscillations (~ 80 – 100 Hz) recorded with ECoG showed non-phase-locked responses to auditory stimuli including pure tones and phonemes (Crone et al., 2001). Compared with tones, phonemes induced a larger power increase in the high gamma band suggesting the potential role of these oscillations in sensory and speech perception (Crone et al., 2001). Compellingly, Helfrich and colleagues (2014) successfully manipulated visual perception by driving subjects' brain oscillations in the gamma band (40 Hz) into synchrony using oscillatory current generated with transcranial alternating current stimulation (tACS).

In the field of speech perception, much attention has been paid to the correspondence between linguistic contrasts at different time scales in speech signals and the associated neuronal oscillations that are generated at different frequency bands (Giraud and Poeppel, 2012, Peelle and Davis, 2012, Goswami and Leong, 2013). Based on the close correspondences between speech contrasts and the patterns of brain oscillations, Giraud and Poeppel (2012) proposed an oscillation-based model for speech perception. According to this model the quasi-rhythmic dynamics of speech at multiple timescales are encoded through stimulus-brain alignment of neuronal oscillations at delta (1-3 Hz), theta (4-8 Hz), and gamma (25-30 Hz) bands, respectively. Such oscillation-dependent operations highlight the crucial role in human cognition played by neuronal oscillations. Given the functional significance of brain oscillations, understanding the neural mechanisms that underlie cortical processing promises to benefit greatly from the application of appropriate analytic methods to reveal the dynamics of brain oscillations. The development and application of suitable analytic methods as well as the development of an appropriate paradigm is therefore critical for advancing our understanding of human cognition (Basar et al., 2001, Gross, 2014).

1.4.1.2 Time-frequency analysis

Time-frequency analysis is a method for transforming the temporally varying brain signals recorded with EEG/MEG/ECoG into both time and frequency domains, in order to unfold the frequency components of these brain signals as well as the temporal evolution of their frequency components (Pfurtscheller and Silva, 1999). In other words, time-varying brain signals are transformed into a time-frequency plane. This time-frequency plane not only reveals the frequency components of brain signals, but also how these frequency components change over time (Durka et al., 2001, Gross, 2014). The resulting spectrogram data is therefore called a time-frequency representation. Analytical techniques for the transformation of time series into the frequency domain include non-parametric methods such as Fourier transform, wavelet transform, and Hilbert transform, and parametric methods such as autoregressive modelling

(Gross, 2014). In cognitive science, time-frequency analysis can be performed either on the brain signals averaged cross trials or on single trial brain signals. Performing these analyses on different kinds of brain signals reveals different types of oscillatory responses and thus reflects the different dynamics of the oscillations (Pfurtscheller and Silva, 1999, Durka et al., 2001).

1.4.1.3 PL activity and averaging-based amplitude measures

The amplitude of most electrophysiological responses in the brain is quite small. This makes them difficult to be observed in single trials. By averaging across trials, the background brain activity, which is considered as randomly distributed, cancels each other out through phase cancellation, leaving the time- and phase-locked activity in the averaged waveform. Traditionally, this phase-locked activity is termed event-related potential (ERP) when it is recorded using EEG/ ECoG, and event-related field (ERF) when it is recorded using MEG.

Auditory research investigating temporal processing usually uses analytic methods such as averaging across trials to extract the auditory evoked potentials/fields (AEP/Fs). The underlying assumption at work is that the AEP/F reflects the time- and phase-locked responses that are elicited by an acoustic signal presented to the subject. Averaging in the time domain is carried out to enhance the signal to noise ratio of phase-locked responses and to cancel non phase-locked electrophysiological noise. This procedure implicitly assumes that non-phase locked brain activity is noise and is of no importance. Additionally, in previous ASSR/EFR studies, continuous acoustic stimulation has usually been presented continuously (without pauses in the stream of stimulation or between sweeps), so that it is not possible to compare stimulus-related brain activity with non stimulus-related baseline. This is a very efficient and useful paradigm for studying temporal processing, but only if one assumes that non phase-locked brain responses are of no importance. As indicated above however, it is becoming increasingly clear from invasive studies in animals and human that non-phase locked responses are a prominent feature of temporal information processing in auditory cortex, particularly at higher temporal modulation rates.

1.4.1.4 PL activity and inter-trial phase coherence measures

In addition to the averaging-based amplitude/power measure, a phase synchronisation measure – the inter-trial phase coherence (ITPC) – has also been commonly applied to extract ERPs/Fs. ITPC quantifies phase consistency across trials for a specific frequency and at a specific time point (Ding and Simon, 2013). For example, Edwards and colleagues used the ITPC to study the phase-consistent time-frequency responses that contribute to ERPs (Edwards et al., 2009). Similarly, inter-trial phase-locking factor was computed by Miyazaki and colleagues (Miyazaki et al., 2013) in order to exhibit the auditory following responses to a two-beat complex sweep between 3 Hz to 60 Hz.

Compared to the averaging-based amplitude/power measure, ITPC is thought to be a more robust and sensitive measure of phase-locked activity (Ding and Simon, 2013, Miyazaki et al., 2013). When the phase-locking factor of the following responses was compared to the amplitude of the following responses, the former showed a much stronger tracking pattern than did the latter, especially in the higher frequency range (Miyazaki et al., 2013). Analytically and numerically, Ding and Simon demonstrated that phase synchronization is much more sensitive to the existence of stimulus-synchronized/phase-locked activity than is response power (Ding and Simon, 2013). The observation that phase is more sensitive than amplitude/power in response to a stimulus, however, does not necessary imply a difference in the underlying physiological mechanisms. According to Ding and Simon (Ding and Simon, 2013), this differential sensitivity might be caused by intrinsic statistical properties of the two measures. Therefore, ITPC as a sensitive measure to extract and characterize stimulus-driven neural activity has been used in many studies (Luo and Poeppel, 2007, Lakatos et al., 2008, Kayser, 2009).

ITPC quantifies phase consistency cross trials for each frequency and time point, and is usually computed from the single trial time-frequency representation using the following formula.

$$ITPC(t,f) = \frac{1}{N}\sum_{n=1}^{N} e^{i\varphi_n(t,f)}$$

where N = total number of trials; and $\varphi_n(f,t)$ = the phase in trial n.

Two models have been offered concerning the neural mechanisms of ERP/Fs. One is an additive model, which proposes that ERP/Fs are generated by new phase-locked activity superimposed onto intrinsic brain oscillations. More specifically, the additive model assumes that a constant evoked response is added onto the ongoing brain oscillations on each trial, and the ongoing brain oscillations are considered to be randomly distributed noise that is completely unrelated to the evoked response. Therefore, averaging single trial waveforms across a large number of trials increases the signal-to-noise ratio of the evoked response, by cancelling out the randomly distributed noise (Sauseng et al., 2007, Howard and Poeppel, 2012). The alternative is a phase reset model. This model proposes that the phase resetting of the intrinsic brain oscillations generates ERP/Fs. The ongoing brain oscillations may be realigned by sensory input and, therefore, yield a high phase concentration across trials. Although the amplitude of the oscillation does not increase, the high phase concentration results in an evoked response when the signals are averaged across trials (Sauseng et al., 2007, Howard and Poeppel, 2012). Howard and Poeppel (2012) developed an analytical method that examines the correlation between amplitude change and phase coherence. This method allowed them to determine whether the ERFs recorded during spoken sentences were generated by additive activity or by a phase-resetting of the ongoing theta band oscillations. A co-modulation of amplitude and phase was found and the authors concluded that the theta band phase-locked response to attended speech, as recorded with MEG, was dominated by additive activity that phase-locked to the onset of each sentence (Howard and Poeppel, 2012). Consistent with Howard and Poeppel (2012), Ding and Simon (2013) showed that the phase of neural response is synchronized across trials even without a phase resetting of the ongoing background oscillatory activity.

1.4.1.5 NPL oscillatory activities and ERSP measure

Neurophysiological and electrophysiological evidence has shown that non phase-locked activity is a prominent feature of many aspects of cognitive processing, especially at higher frequencies (Crone et al., 2001, Brugge et al., 2009, Crone et al., 2011). For example, Steinschneider and colleagues (2008) examined EEG responses to tones in the primary auditory cortex of awake monkeys. Post-stimulus broadband power increases were found between 4 Hz and 290 Hz. These power increases were mostly PL activity at frequencies below about 70 Hz, and primarily NPL at higher frequencies. Interestingly, compared with PL power increase, the NPL power increases at higher frequencies were better correlated with the A1 tonotopic organization, which suggests that the high frequency NPL activity is a valuable way of studying the functional organisation of auditory cortex and important for the inclusion of NPL responses in EEG/MEG data analysis (Steinschneider et al., 2008)

As indicated in the previous sections, both averaging-based amplitude/power measure and ITPC measure allow researchers to investigate event-related PL activity. However, if temporal information is processed in rate-place codes without any synchrony between spike-timing and the onset of stimulation, then this averaging-based measure is inadequate for a comprehensive understanding of temporal processing in the auditory cortex (Makeig, 1993, Pfurtscheller and Silva, 1999, Crone et al., 2001, Brugge et al., 2009, Nourski and Brugge, 2011). Methods for examining the NPL activity (or induced responses) and the inclusion of both PL and NPL in analysis of the electrophysiological data is thus important for a comprehensive understanding of temporal processing in the auditory cortex.

To measure NPL activity at a given frequency band (e.g., alpha suppression), event-related (de)synchronization (ERS/D) analysis can be applied (Pfurtscheller and Silva, 1999). ERS/D quantifies the degree to which the amplitude of certain prominent ongoing frequency band EEG was “synchronized”/”desynchronized” by the stimulation events. To compute the ERS/D, a specific frequency range needs to be chosen a *priori*. Usually, this is fulfilled with a

trial-and-error procedure (Makeig, 1993, Makeig et al., 2004). As such, the selection of the frequency range could be subjective and unsuitable for different participants given the potential for substantial individual variation of each frequency band (Doppelmayr et al., 1998).

ERSPs, introduced by Makeig and others (Makeig, 1993, Makeig et al., 2004), can be considered as a generalization of ERS/D. ERSP quantifies the increase or decrease in amplitude/power of the broad-band frequency spectrum as a function of time with reference to the ongoing background brain activity recorded in the defined baseline (Makeig, 1993, Pfurtscheller and Silva, 1999, Brugge et al., 2009). Compared to ERS/D, full-spectrum ERSP analysis provides richer information of the brain signal being analyzed (Makeig, 1993).

More importantly, since this ERSP measure quantifies the total amplitude/power change with reference to the baseline, it contains both the PL activity (phase-locked related change) and NPL activity (non-phase-locked by event-related change) (Makeig, 1993, Tallon-Baudry and Bertrand, 1999, Brugge et al., 2009). The separation of PL and NPL can be accomplished by simple subtraction and therefore provide a window for us to investigate event-related activity originates from different cortical processes that are temporally and spatially overlapping (Nourski and Brugge, 2011). Although other researchers invoke different terms, such as event-related band power (ERBP) (Brugge et al., 2009, Nourski et al., 2013), the method is similar. The reason we used ERSP instead of ERBP is to avoid any pre-defined specific frequency band. A wide range of frequencies remain of interest and, thus, are included. Furthermore, the concept of "perturbation" fits well with our event-related paradigm for perturbing the ongoing brain oscillations with an AM sound whose modulation rate sweeps from 1 Hz to 80 Hz. This sweep stimulus creates a brief perturbation of the brain signals at each frequency within the 1-80 Hz range.

More specifically, ERSPs are computed using the complex demodulation method (Hoechstetter et al., 2004), with a frequency step of 1 Hz and a time step of 50 ms. The ERSP (Makeig, 1993, Pfurtscheller and Silva, 1999, Brugge et al., 2009) quantifies the increase or

decrease in amplitude/power at a given frequency or frequency band with reference to the ongoing background brain activity recorded in a defined baseline. Amplitude is normalized independently in each frequency or frequency band. The ERSP index is calculated using the following formula.

$$ERSP(\mathrm{t,f}) = \frac{\sum_{\mathrm{n=1}}^{\mathrm{N}} \left(\frac{\mathrm{A_n(t,f)} - \overline{\mathrm{A_{baseline,\ n}(f)}}}{\overline{\mathrm{A_{baseline,\ n}(f)}}} \times 100\% \right)}{\mathrm{N}}$$

where $A_n(t,f)$ is the absolute amplitude at time *t* and frequency *f* in trial *n*; and $\overline{A_{baseline,\ n}(f)}$ is the mean absolute amplitude at frequency *f* over the defined baseline in trial *n*.

The phase-locked component of ERSP (*PL)* is computed by applying spectral analysis on averaged waveform, which is obtained from averaging the MEG signals across trials, using the following formula:

$$PL(t,f) = \frac{\bar{A}(t,f) - \overline{\overline{A_{baseline}}(f)}}{\overline{\overline{A_{baseline}}(f)}} \times 100\%$$

where $\bar{A}(t,f)$ is the absolute amplitude at time *t* and frequency *f* of the averaged signal; and $\overline{\overline{A_{baseline}}(f)}$ is the mean absolute amplitude of the averaged signal at frequency *f* over the defined baseline.

The non-phase-locked component of ERSP (*NPL*) is then computed by subtracting the *PL* from the amplitude change in single trial, as indicated by the following formula:

$$NPL(t,f) = \frac{\sum_{n=1}^{N} \left(\frac{A_n(t,f) - \overline{A_{baseline,\ n}(f)}}{\overline{A_{baseline,\ n}(f)}} \times 100\% - PL(t,f) \right)}{N}$$

1.4.2 Electrophysiological techniques for studying brain functions

Electrophysiological studies in auditory processing and speech perception often use MEG and EEG to record brain activities to take advantage of their high temporal resolution.

MEG and EEG are two complementary techniques that measure, respectively, the magnetic field outside the head and the electric field on the scalp (Hämäläinen et al., 1993).

1.4.2.1 The neurophysiology of MEG and EEG

MEG and EEG are two complementary techniques that measure magnetic fields outside the head and the electric potential on the scalp, respectively (Nakasatp et al., 1994, Baillet et al., 2001, Sharon et al., 2007, Hansen et al., 2010). MEG measure the magnetic fields generated mainly by intracellular currents (also called primary currents) in cortical pyramidal cells at the location of each of the sensors fixed inside a helmet (Baillet et al., 2001, Hansen et al., 2010). Unlike MEG, EEG measures the electric potentials on the scalp produced by extracellular currents (also termed return currents) which travel through the volume conductor of the head at the location of each of the electrodes attached to the scalp (Baillet et al., 2001, Buzsaki et al., 2012). The primary currents are ionic currents that are generated within pyramidal cells as the result of synaptic activity. Due to the force caused by the electric field, the primary currents induce return currents outside the cells. Therefore MEG and EEG are measure different aspects of electromagnetic fields arising from ionic current flow in the neural cell assembly.

To be detected by EEG and MEG, the spatial arrangement of these neural cell assemblies is crucial. It has been suggested that the main generators for EEG and MEG signal are macro-columns consisting of tens of thousands of large pyramidal cortical neurons. When these pyramidal neurons are simultaneously activated, the dendritic trunks of these neurons form a coherence assembly that points perpendicularly to the cortical surface. This coherent assembly of neurons generate summated magnetic fields large enough to be measurable outside the head using MEG, as well electric fields measureable on the scalp with EEG (Hämäläinen et al., 1993, Baillet et al., 2001, Hansen et al., 2010). Figure 6 gives an illustration of the auditory evoked response elicited by a broadband noise recorded in concurrent MEG and EEG. As shown in Figure 6, the main generator (primary current) is quite similar in terms of location and orientation, despite the different field topographies.

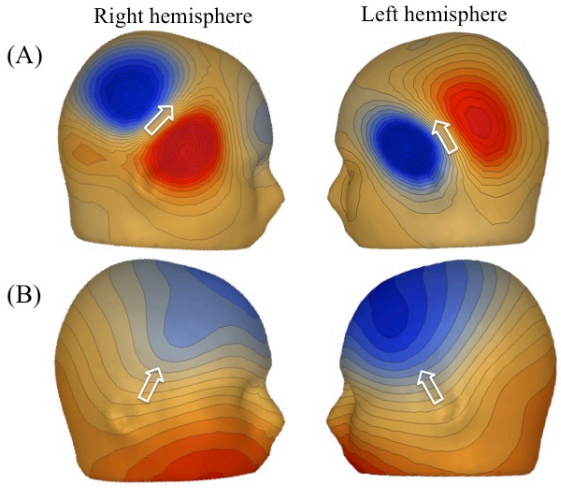

Figure 6. Topographies of auditory evoked magnetic fields and scalp potentials recorded in concurrent MEG and EEG, respectively, of a representative subject. (A) Magnetic flux of M100 measured by MEG sensors is projected onto the surface of the head; Blue indicates negative flux (entering head), red indicates positive flux (leaving head). (B) Scalp potential of N100 recorded with EEG. Blue indicates negative potential, and red indicates positive potential. The open arrows show the surface projection of the cortical primary current generating the magnetic field and electric field.

Both MEG and EEG have excellent temporal resolution, but they also have their own unique characteristics (Hämäläinen et al., 1993, Darvas et al., 2004, Nunez and Srinivasan, 2006, Hansen et al., 2010). As stated earlier, the return currents that reach the scalp are the electric potentials measured by EEG, whereas the magnetic field produced by the primary currents is the signal measured by MEG. Because of the complexity of brain structures and the strength of electric currents are highly influenced by the conductivity of the medium they travel through, the interpretation of EEG signals requires precise knowledge of the thicknesses and conductivity of different layers of the brain volume. Unlike electric fields, magnetic fields measured with MEG are not influenced by conductivity of the medium. This simplifies, at least to some extent, the source localization problem, and provides a relatively higher spatial resolution profile as compared with EEG.

Furthermore, MEG is more subject to the structure of the cortex, as compared with EEG (Hämäläinen et al., 1993, Hansen et al., 2010). The cortical surface is densely folded and thus forms different orientations of the primary currents. MEG is sensitive to the tangentially

(relative to the spherical model of the head) oriented primary currents, which arise in the sulci – the concave depths of the cortical folds; but MEG is insensitive to the radially oriented primary currents, which originate in the gyri – the convex tops of the cortical folds. The further away the source, the closer it approaches to the center of the head sphere, so the more radial it becomes (Hämäläinen et al., 1993). As a result, MEG is generally thought to be relatively insensitive to signals from deeper brain structures. In contrast, EEG can detect return currents induced by both tangential and radial primary currents (Jones and Byrne, 1998, Tonnquist-Uhlen et al., 2003, Nunez and Srinivasan, 2006).

In addition, MEG measure is also sensitive to the distance between the MEG sensors and the brain signal source. MEG sensors are fixed inside the MEG helmet, which locates with a distance from the subject's head. This characteristic of MEG measurement implies an inherent limitation of deep source imaging in MEG. First, the magnitude of the magnetic field decreases as the distance (r) from the location of measure to the field source increases in a manner of $1/r^2$ (Hansen et al., 2010). Thus, the amplitude of the magnetic field at a fixed sensor array depends strongly on the depth of the source of brain activity and the distance between the head and the sensors. Additionally, MEG is highly subject to environmental noise due to the small amplitude of neuromagnetic signals (Hansen et al., 2010). To reduce the noise, advanced techniques such as gradiometers are used. Gradiometers, however, increase the depth dependency of MEG measurements (de Jongh et al., 2005). Given the magnetic field generated by the small cortical current flow is at an extremely small scale, the detection of deep sources can be limited in MEG recording (Hillebrand and Barnes, 2002, Leijten et al., 2003).

In view of these considerations, it is reasonable to suggest that neuromagnetic signals detectable outside the head are mainly produced in the cerebral cortex, due to its proximity to the MEG sensors. On the other hand, deeper brain structures such as brainstem (Herdman et al., 2002, Purcell et al., 2004) have been reported in many EEG studies using scalp recordings. While the high sensitivity of MEG to cortical structure and distance has been thought as a

disadvantage of MEG, it can be an advantage for studying auditory processing given that the superficial and tangential sources are the main generators (Phillips et al., 2000, Houde et al., 2002, Chait et al., 2004, Miyazaki et al., 2013).

1.4.2.2 Paediatric MEG for preschool-aged children

As stated above, the measurement of neuromagnetic signals in MEG highly depends on the distance between the MEG sensor array fixed inside the MEG helmet and the source of the signal. It is therefore crucial to place the participant's head as close to the fixed sensor array installed inside the MEG helmet as possible. However, due to the much smaller sizes of young children's head, the distance between the head center and sensor array could be too large for accurate measurements (Johnson et al., 2010, He et al., 2014). What makes the task even harder is the fact that children's shoulders usually prevent them from inserting their heads completely into an adult MEG helmet (Johnson et al., 2010, He et al., 2014). Figure 7 illustrates the problem in using the conventional MEG system for measuring brain activity in young children. Hence, in addition to using standard MEG and EEG systems to record brain activity in adults, a custom-designed pediatric MEG system was employed to measure brain responses in preschool-aged children.

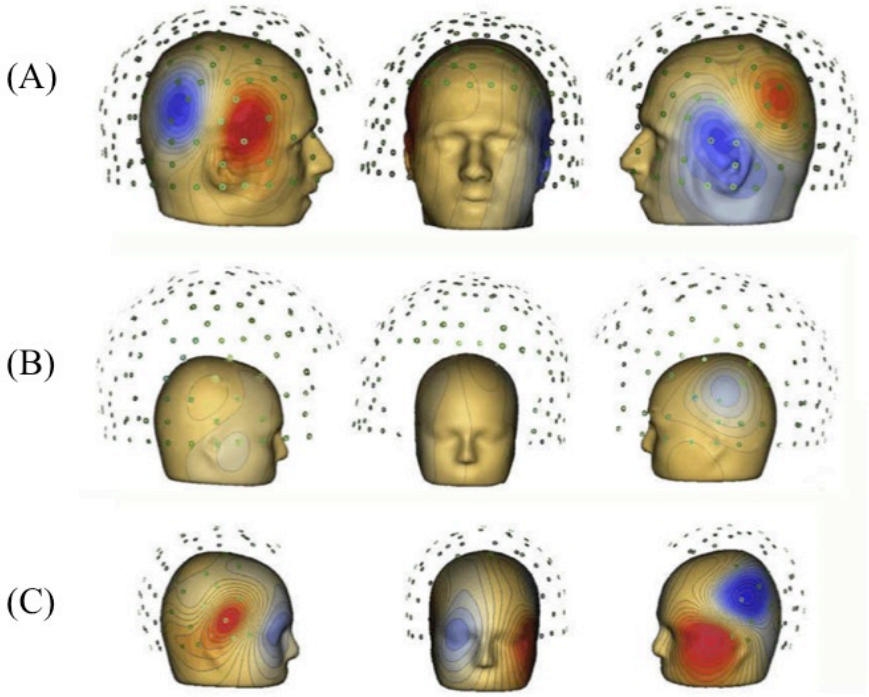

Figure 7. Sensor positions shown with respect to head for: (A) an adult in adult MEG. (B) A child in adult MEG. (C) A child in child MEG. M100 topography is shown in (A), P100m topography in (C). Magnetic flux measured by MEG sensors is projected onto the surface of the head. Blue indicates negative flux (entering head), red indicates positive flux (leaving head). Figure adapted from Johnson et al., (2010).

1.5 Research questions

Temporal modulations in the overall amplitude of sounds (the "sound envelope") contain critical information for the perception of speech and music (Drullman, 1995; Drullman et al., 1994; Rosen, 1992; Shannon et al., 1995). Interest in the topic of auditory temporal processing has been strongly stimulated by recent neurolinguistic models that advocate an essential role in speech perception for neural mechanisms that encode the speech envelope (Giraud & Poeppel, 2012; Goswami & Leong, 2013; Gross et al., 2013). How the temporal modulations are encoded in the human auditory cortex, however, is still unclear. Neurophysiological profiles based on the EEG/MEG-recorded ASSR/EFT show a low pass modulation transfer function with a high cutoff of about 50 Hz. However modulation rates well above this 50 Hz bound contain crucial temporal information for perception of environmental sounds, music and speech. This thesis was motivated by this gap in our understanding. Moreover, temporal

encoding models must account for the maturation of temporal processing during the critical periods in the development of speech perception and language acquisition. At present, there is much to learn about how the sound envelope is processed in the developing brain. Therefore, a second aim of the current thesis was to examine the temporal processing measured by EFRs in the brains of preschool-aged children. This thesis addressed the following questions:

1. How are temporal modulations encoded in the human auditory cortex? In particular, how are perceptually crucial temporal modulations at high temporal rates beyond the optimal upper limit of PL responses being encoded in the human auditory cortex? Using the ERSP analysis, **Chapter 2** demonstrates a dissociation between MEG-recorded PL and NPL following responses to the envelope of white noise amplitude modulated at the rate of 1-80 Hz. While both types of following responses were elicited by AM noise, the PL response was most prominent for modulation rates below 50 Hz; in contrast, NPL responses were most prominent in the higher frequency range (i.e. 50 – 80 Hz). These findings are difficult to accommodate within a framework that views the auditory cortex as a low pass temporal filter; however they are entirely consistent with the existence of a dual temporal encoding scheme as suggested by measurement of single unit responses in primary auditory cortex of nonhuman primates (Lu et al., 2001).

2. Can the MEG recorded PL and NPL following responses be observed using EEG? Since EEG measures both cortical and subcortical signals, we ask whether both the primary auditory cortex and the brainstem contribute to the EEG recorded following response. **Chapter 3** analyzes EEG and MEG data recorded concurrently in the same experiment. Dual PL and NPL following responses that are found in MEG data were also found in EEG data in the auditory cortical sources. In contrast, the brainstem source showed only PL following responses with no significant NPL responses. These analyses fit well with the explanation that PL responses predominate in subcortical auditory regions, with a progressive transformation of PL to NPL responding as information ascends in the central auditory system (Joris et al., 2004).

3. How does cortical temporal encoding develop in children? To our knowledge there are no previous EEG/MEG data on the ASSR/EFR for children older than infants or younger than school age. Yet, the preschool years between 3 and 5 years represent a crucial period in children's acquisition of language. Given the emerging importance of temporal processing in neurobiological models of language perception, it is imperative to provide objective neurophysiological data on cortical temporal processing capabilities in children during the course of language acquisition. Our results (**Chapter 4**) indicate that temporal processing is limited to slow modulation below about 25 Hz in synchronization-based temporal codes in preschool-aged children. These results support the hypothesis that temporal processing in the auditory cortex has a protracted maturational time course (Moore et al., 2001).

1.6 References

Aiken SJ, Picton TW (2008) Human cortical responses to the speech envelope. *Ear Hear 29:*139-157.

Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. *IEEE Signal Process Mag* 18:14-30.

Banai K, Sabin AT, Wright BA (2011) Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hear Res* 280:219-227.

Bartlett EL, Wang X (2007) Neural Representations of Temporally Modulated Signals in the Auditory Thalamus of Awake Primates. *J Neurophysiol 97*:1005-1017.

Basar E, Basar-Eroglu C, Karakas S, Schurmann M (2001) Gamma, alpha, delta, and theta oscillations govern cognitive processes. *Int J Psychophysiol* 39:241-248.

Başar E, Schürmann M, Başar-Eroglu C, Karakaş S (1997) Alpha oscillations in brain functioning: an integrative theory. *Int J Psychophysiol 26:*5-29.

Baumann S, Joly O, Rees A, Petkov CI, Sun L, Thiele A, Griffiths TD (2015) The topography of frequency and time representation in primate auditory cortices. *eLife 4*.

Ben-Yehudah G, Banai K, Ahissar M (2004) Patterns of deficit in auditory temporal processing among dyslexic adults. *NeuroReport* 15:627-631.

Bieser A, Muller-Preuss P (1996) Auditory responsive cortex in the squirrel monkey: neural responses to amplitude-modulated sounds. *Exp Brain Res* 108:273-284.

Bishop DV (2007) Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: where are we, and where should we be going? *Psychol Bull 133*:651-672.

Bishop DVM, Carlyon RP, Deeks JM, Bishop SJ (1999) Auditory Temporal Processing Impairment: Neither Necessary nor Sufficient for Causing Language Impairment in Children. *J Speech Lang Hear Res* 42:1295-1310.

Boets B, Wouters J, van Wieringen A, Ghesquiere P (2007) Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: a longitudinal study of the auditory temporal processing theory. *Neuropsychologia 45:*1608-1620.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. *J Neurophysiol 102*:2358-2374.

Buss E, Hall IIIJW, Porter H, Grose JH (2014) Gap Detection in School-Age Children and Adults: Effects of Inherent Envelope Modulation and the Availability of Cues Across Frequency. *J Speech Lang Hear Res 57:*1098-1107.

Buzsaki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. *Nat Rev Neurosci 13:*407-420.

Chait M, Simon JZ, Poeppel D (2004) Auditory M50 and M100 responses to broadband noise: functional implications. *Neuroreport 15:*2455-2458.

Cohen LT, Rickards FW, Clark GM (1991) A comparison of steady-state evoked potentials to modulated tones in awake and sleeping humans. *J Acoust Soc Am 90:*2467-2479.

Crone NE, Boatman D, Gordon B, Hao L (2001) Induced electrocorticographic gamma activity during auditory perception. *Clin Neurophysiol 112*: 565-582.

Crone NE, Korzeniewska A, Franaszczuk PJ (2011) Cortical gamma responses: searching high and low. *Int J Psychophysiol 79*:9-15.

Darvas F, Pantazis D, Kucukaltun-Yildirim E, Leahy RM (2004) Mapping human brain function with MEG and EEG: methods and validation. *Neuroimage 23* Suppl 1:S289-299.

David SV, Shamma SA (2013) Integration over Multiple Timescales in Primary Auditory Cortex. *J Neurosci 33*:19154-19166.

Dawes P, Bishop DVM (2008) Maturation of Visual and Auditory Temporal Processing in School-Aged Children. *J Speech Lang Hear Res 51*:1002-1015.

de Jongh A, de Munck JC, Gonçalves SI, Ossenblok P (2005) Differences in MEG/EEG epileptic spike yields explained by regional differences in signal-to-noise ratios. *J Clin Neurophysiol 22:*153-158.

Linden RD, Picton TW, Hamel G, Campbell KB (1987) Human auditory steady-state evoked potentials during selective attention. *Electroencephalogr Clin Neurophysiol 66:*145-159.

Ding N, Simon JZ (2012) Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. *J Neurophysiol 107*:78-89.

Ding N, Simon JZ (2013) Power and phase properties of oscillatory neural responses in the presence of background activity. *J Comput Neurosci 34:*337-343.

Doppelmayr M, Klimesch W, Pachinger T, Ripper B (1998) Individual di€erences in brain dynamics: important implications for the calculation of event-related band power. *Biol Cybern 79*:49-57.

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am 97:*585-592.

Drullman R, Festen JM, Plomp R (1994a) Effect of reducing slow temporal modulations on speech reception. *J Acoust Soc Am 95:*2670-2680.

Drullman R, Festen JM, Plomp R (1994b) Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am 95:*1053-1064.

Dugue L, Marque P, VanRullen R (2011) The phase of ongoing oscillations mediates the causal relation between brain excitation and visual perception. *J Neurosci 31:*11889-11893.

Durka PJ, Ircha D, Neuper C, Pfurtscheller G (2001) Time-frequency microstructure of event-related electro-encephalogram desynchronisation and synchronisation. *Med Biol Eng Comput 39*:315-321.

Edwards E, Soltani M, Kim W, Dalal SS, Nagarajan SS, Berger MS, Knight RT (2009) Comparison of time–frequency responses and the event-related potential to auditory speech stimuli in human cortex. *J Neurophysiol 102*:377-386.

Eggermont JJ (2001) Between sound and perception: reviewing the search for a neural code. *Hear Res 157:*1-42.

Eggermont JJ (2014) Animal models of auditory temporal processing. *Int J Psychophysiol 95*:202-215.

Fishman YI, Reser DH, Arezzo JC, Steinschneider M (2000) Complex tone processing in primary auditory cortex of the awake monkey. I. Neural ensemble correlates of roughness. *J Acoust Soc Am* 108:235.

Fox AM, Reid CL, Anderson M, Richardson C, Bishop DV (2012) Maturation of rapid auditory temporal processing and subsequent nonword repetition performance in children. *Dev Sci 15*:204-211.

Galambos R (1992) A comparison of certain gamma band (40-Hz) brain rhythms in cat and man: Birkhäuser, Boston, MA.

Galambos R, Makeig S, Talmachoff PJ (1981) A 40-Hz auditory potential recorded from the human scalp. *Proc Indian Natl Sci Acad 78:*2643-2647.

Gaucher Q, Huetz C, Gourevitch B, Laudanski J, Occelli F, Edeline JM (2013) How do auditory cortex neurons represent communication sounds? *Hear Res* 305:102-112.

Ghitza O, Giraud A, Poeppel D (2013) Neuronal oscillations and speech perception: critical-band temporal envelopes are the essence. *Front Hum Neurosci* 6:1-4.

Giraud A-L, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A (2000) Representation of the temporal envelope of sounds in the human brain. *J Neurophysiol 84:*1588-1598.

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci 15:*511-517.

Goldstein MH, Kiang NYS, Brown RM (1959) Responses of the Auditory Cortex to Repetitive Acoustic Stimuli. *J Acoust Soc Am* 31:356-364.

Goswami U, Leong V (2013) Speech rhythm and temporal structure: Converging perspectives? *Lab Phonol* 4:67-92.

Gross J (2014) Analytical methods and experimental approaches for electrophysiological studies of brain oscillations. *J Neurosci Methods 228:*57-66.

Hall JW, Grose JH (1994) Development of temporal resolution in children as measured by the temporal modulation transfer function. *J Acoust Soc Am* 96:150-154.

Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev Mod Phys* 65:413-497.

Hansen PC, Kringelbach ML, Salmelin R (2010) MEG: An Introduction to Methods. New York: Oxford University Press

He W, Brock J, Johnson BW (2014) Face-sensitive brain responses measured from a four-year-old child with a custom-sized child MEG system. *J Neurosci Methods 222*:213-217.

Helfrich RF, Knepper H, Nolte G, Struber D, Rach S, Herrmann CS, Schneider TR, Engel AK (2014) Selective modulation of interhemispheric functional connectivity by HD-tACS shapes perception. *PLoS Biol 12:*e1002031.

Herdman AT, Lins O, Van Roon P, Stapells DR, Scherg M, Picton TW (2002) Intracerebral sources of human auditory steady-state responses. *Brain Topogr 15:*69-86.

Hill PR, Hartley DEH, Glasberg BR, Moore BCJ, Moore DR (2004) Auditory processing efficiency and temporal resolution in children and adults. *J Speech Lang Hear Res 47:*1022-1029.

Hillebrand A, Barnes GR (2002) A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. *NeuroImage* 16:638-650.

Hoechstetter K, Bornfleth H, Weckesser D, Ille N, Berg P, Scherg M (2004) BESA source coherence: a new method to study cortical oscillatory coupling. *Brain Topogr 16*:233-238.

Houde JF, Nagarajan SS, Sekihara K, Merzenich MM (2002) Modulation of the Auditory Cortex during Speech: An MEG Study. J Cogn Neurosci 14:1125-1138.

Howard MF, Poeppel D (2012) The neuromagnetic response to spoken sentences: Co-modulation of theta band amplitude and phase. *NeuroImage* 60:2118-2127.

Irwin RJ, Ball AKR, Kay N, Stillman JA, Rosser J (1985) The Development of Auditory Temporal Acuity in Children. *Child Dev* 56:614-620.

Jansen BH, Brandt ME (1991) The effect of the phase of prestimulus alpha activity on the averaged visual evoked response. *Electroencephalogr Clin Neurophysiol* 80:241-250.

John MS, Brown DK, Muir PJ, Picton TW (2004) Recording auditory steady-state responses in young infants. *Ear Hear 25:*539-553.

John MS, Picton TW (2000) Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. *Hear Res* **141**:57-79.

Johnson BW, Crain S, Thornton R, Tesan G, Reid M (2010) Measurement of brain function in pre-school children using a custom sized whole-head MEG sensor array. *Clin Neurophysiol 121:*340-349.

Jones SJ, Byrne C (1998) The AEP T-complex to synthesised musical tones: left-right asymmetry in relation to handedness and hemisphere dominance. *Electroencephalogr Clin Neurophysiol* 108:355-360.

Jorgens S, Biermann-Ruben K, Kurz MW, Flugel C, Daehli Kurz K, Antke C, Hartung HP, Seitz RJ, Schnitzler A (2008) Word deafness as a cortical auditory processing deficit: a case report with MEG. *Neurocase 14:*307-316.

Joris P, Schreiner C, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev 84*:541-577.

Kayser C (2009) Phase Resetting as a Mechanism for Supramodal Attentional Control. *Neuron 64:*300-302.

Kraus N (2012) Atypical brain oscillations: a biological basis for dyslexia? Trends Cogn Sci 16.

Kubanek J, Brunner P, Gunduz A, Poeppel D, G S (2013) The Tracking of Speech Envelope in the Human Cortex. *PLoS One* 8:e53398.

Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science 320:*110-112.

Lehongre K, Ramus F, Villiermet N, Schwartz D, Giraud AL (2011) Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron 72:*1080-1090.

Leijten FSS, Huiskamp G-JM, Hilgersom I, van Huffelen AC (2003) High-resolution Source Imaging in Mesiotemporal Lobe Epilepsy: A Comparison Between MEG and Simultaneous EEG. J Clin Neurophysiol 20:227-238.

Levi EC, Folsom RC, Dobie RA (1995) Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hear Res* 89:21-27.

Liang L, Lu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *J Neurophysiol 87:*2237-2261.

Liegeois-Chauvel C, Lorenzi C, Trebuchon A, Regis J, Chauvel P (2004) Temporal envelope processing in the human left and right auditory cortices. Cereb Cortex 14:731-740.

Lins OG, Picton TW (1995) Auditory steady-state responses to multiple simultaneous stimuli. *Electroencephalogr Clin Neurophysiol* 96:420-432.

Lu T, Liang L, Wang X (2001a) Neural Representations of Temporally Asymmetric

Stimuli in the Auditory Cortex of Awake Primates. *J Neurophysiol 85*:2364-2380.

Lu T, Liang L, Wang X (2001b) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat Neurosci 4:*1131-1138.

Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron 54:*1001-1010.

Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr Clin Neurophysiol 86:*283-293.

Makeig S, Debener S, Onton J, Delorme A (2004) Mining event-related brain dynamics. *Trends Cogn Sci 8:*204-210.

Makela JP, Hari R (1987) Evidence for cortical origin of the 40 Hz auditory evoked response in man. *Electroencephalogr Clin Neurophysiol* 66.

Mauk MD, Buonomano DV (2004) The neural basis of temporal processing. *Annual Rev Neurosci 27:*307-340.

Mazaheri A, Jensen O (2008) Asymmetric amplitude modulations of brain oscillations generate slow evoked responses. *J Neurosci 28:*7781-7787.

Mijares Nodarse E, Perez Abalo MC, Torres Fortuny A, Vega Hernandez M (2011) Electrophysiological Characterisation of Envelope-Following Responses. *Acta Otorrinolaringol Esp* 62:425-431.

Miyazaki T, Thompson J, Fujioka T, Ross B (2013) Sound envelope encoding in the auditory cortex revealed by neuromagnetic responses in the theta to gamma frequency bands. *Brain Res 1506*:64-75.

Moore DR (2002) Auditory development and the role of experience. *Br Med Bull* 63:171-181.

Moore DR, Cowan JA, Riley A, Edmondson-Jones AM, Ferguson MA (2011) Development of Auditory Processing in 6- to 11-Yr-Old Children. *Ear Hear 32:*269–285.

Moore JK, Linthicum Jr FH (2007) The human auditory system: a timeline of development. *Int J Audiol* 46:460-478.

Muhler R, Rahne T, Mentzel K, Verhey JL (2014) 40-Hz multiple auditory steady-state responses to narrow-band chirps in sedated and anaesthetized infants. *Int J Pediatric Otorhinolaryngol 78:*762-768.

Nakasatp N, Levesque MF, Barth DS, Baumgartner C, Rogers RL, Sutherling WW (1994) Comparisons of MEG, EEG, and ECoG source localization in neocortical partial epilepsy in humans. *Electroencephalogr Clin Neurophysiol* 91:171-178.

Niedermeyer E (1997) Alpha rhythms as physiological and abnormal phenomena. *Int J Psychophysiol* 26:31-49.

Nodarse EM, Abalo MaCPr, Fortuny AT, Hernández MV, Castellanos AL (2012) Maturational Changes in the Human Envelope-following Responses. *Acta Otorrinolaringol Esp* 63:258-264.

Nourski KV, Brugge JF (2011) Representation of temporal sound features in the human auditory cortex. *Rev Neurosci 22*:187-203.

Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard III MA (2013) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *J Neurophysiol 109:*1283-1295.

Nourski KV, Reale RA, Oya H, Kawasaki H, Kovach CK, Chen H, Howard MA, Brugge JF (2009) Temporal envelope of time-compressed speech represented in the human auditory cortex. *J Neurosci 29*:15564-15574.

Nunez PL, Srinivasan R (2006) Electric fields of the brain: the neurophysics of EEG: Oxford university press.

Obleser J, Weisz N (2012) Suppressed alpha oscillations predict intelligibility of speech and its acoustic details. *Cereb Cortex 22:*2466-2477.

Pantev C (1995) Evoked and induced gamma-band activity of the human cortex. *Brain Topogr 7:*321-330.

Pantev C, Makeig S, Hoke M, Galambos R, Hampson S, C.Gallen (1991) Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Sciences 88*:8996-9000.

Pasley BN, David SV, Mesgarani N, Flinker A, Shamma SA, Crone NE, Knight RT, Chang EF (2012) Reconstructing speech from human auditory cortex. *PLoS Biol 10:*e1001251.

Patterson RD (1994) The sound of a sinusoid: Time-interval models. *J Acoust Soc Am* 96:1419-1428.

Peelle JE, Davis MH (2012) Neural oscillations carry speech rhythm through to comprehension. *Front Psychol 3*.

Peña M, Melloni L (2012) Brain oscillations during spoken sentence processing. *J Cogn Neurosci 24:*1149-1164.

Pethe J, Specht Hv, Mühler R, Hocke T (2001) Amplitude modulation following responses in awake and sleeping humans - a comparison for 40 Hz and 80 Hz modulation frequency. *Scand Audiol* 30:152-155.

Pfurtscheller G, Silva FHLd (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin Neurophysiol 110:*1842-1857.

Phillips C, Pellathy T, Marantz A, Yellin E, Wexler K, Poeppel D, McGinnis M, Roberts T (2000) Auditory cortex accesses phonological categories: an MEG mismatch study. *J Cogn Neurosci 12*:1038-1055.

Phillips DP, Farmer ME (1990) Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. *Behav Brain Res* 40:85-94.

Picton T (2013) Hearing in Time: evoked Potential Studies of Temporal Processing. *Ear Hear 34:*385-401.

Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int J Audiol* **42**:177-219.

Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun 41:*245-255.

Poeppel D, Idsardi WJ, van Wassenhove V (2008) Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Lond* **363**:1071-1086.

Poulsen C, Picton TW, Paus T (2007) Age-related changes in transient and oscillatory brain responses to auditory stimulation in healthy adults 19-45 years old. *Cerebl Cortex 17:*1454-1467.

Preuß A, Müller-Preuss P (1990) Processing of amplitude modulated sounds in the medial geniculate body of squirrel monkeys. Experimental Brain Research 79:207-211.

Purcell DW, John SM, Schneider BA, Picton TW (2004) Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am* **116**:3581-3593.

Regan D (1989) Human brain electrophysiology : evoked potentials and evoked magnetic fields in science and medicine. New York Elsevier.

Rhode WS, Greenberg S (1994) Encoding of amplitude modulation in the cochlear nucleus of the cat.

Rickards FW, Tan LE, Cohen LT, Wilson OJ, Drew JH, Clark GM (1994) Auditory steady-state evoked potential in newborns. *Br J Audiol* 28:327-337.

Rosen S (1992) Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects. *Philos Trans R Soc Lond B Biol Sci 336*:367-373.

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

Roß B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones. *J Acoust Soc Am 108:*679-691.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones *J Acoust Soc Am* 108:679-691.

Ross B, Herdman A, Pantev C (2005) Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cereb Cortex 15:*2029-2039.

Roß B, Picton TW, Pantev C (2002) Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field. *Hear Res* 165:68-84.

Sauseng P, Klimesch W, Gruber W, Hanslmayr S, Freunberger R, Doppelmayr M (2007) Are event-related potential components generated by phase resetting of brain oscillations? A critical discussion. *Neurosci* 146:1435-1444.

Schurmann M, Basar E (2001) Functional aspects of alpha oscillations in the EEG. *Int J Psychophysiol 39*:151-158.

Sharon D, Hämäläinen MS, Tootell RBH, Halgren E, Belliveau JW (2007) The advantage of combining MEG and EEG: Comparison to fMRI in focally stimulated visual cortex. *NeuroImage* 36:1225-1235.

Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. *Nature 416.*

Souza PE (2000) Older Listeners' Use of Temporal Cues Altered by Compression Amplification. *J Speech Lang Hear Res* 43.

Steinschneider M, Fishman YI, Arezzo JC (2008) Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. *Cereb Cortex 18*:610-625.

Steinschneider M, Nourski KV, Fishman YI (2013) Representation of speech in human auditory cortex: is it special? *Hear Res* 305:57-73.

Strauß A, Henry MJ, Scharinger M, Obleser J (2015) Alpha phase determines successful lexical decision in noise. *J Neurosci 35:*3256 –3262.

Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *J Neurosci 16*:4240-4249.

Thut G (2014) Modulating brain oscillations to drive brain function. *PLoS Biol 12*:e1002032.

Thut G, Schyns PG, Gross J (2011) Entrainment of perceptually relevant brain oscillations by non-invasive rhythmic stimulation of the human brain. *Front Psychol 2*.

Tlumak AI, Durrant JD, Delgado RE, Boston JR (2011) Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: profile in adults. *Int J Audiol* 50:448-458.

Tlumak AI, Durrant JD, Delgado RE, Boston JR (2012) Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: Profile in children vs. adults. *Int J Audiol* 51:480-490.

Tonnquist-Uhlen I, Ponton CW, Eggermont JJ, Kwong B, Don M (2003) Maturation of human central auditory system activity: the T-complex. *Clin Neurophysiol* 114:685-701.

Trehub SE, Schneider BA, Henderson JL (1995) Gap detection in infants, children, and adults. *J Acoust Soc Am* 98:2532-2541.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am 66:*1364-1380.

Wang X (2007) Neural coding strategies in auditory cortex. *Hear Res* **229**:81-93.

Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. *Neurosci*. *154:*294-303.

Wang X, Lu T, Liang L (2003) Cortical processing of temporal modulations. *Speech Commun* **41**:107-121.

Whitfield I, Evans E (1965) Responses of auditory cortical neurons to stimuli of changing frequency. *J Neurophysiol*.

Whitfield IC (1980) Auditory cortex and the pitch of complex tones. *J Acoust Soc Am* 67:644-647.

Wightman F, Allen P, Dolan T, Kistler D, Jamieson D (1989) Temporal Resolution in Children. *Child Dev* 60:611-624.

Yin P, Johnson JS, O'Connor KN, Sutter ML (2011) Coding of amplitude modulation in primary auditory cortex. *J Neurophysiol* 105:582-600.

Zeng F-G, Nie K, Stickney GS, Kong Y-Y, Vongphoe M, Bhargave A, Wei C, Cao K (2005) Speech recognition with amplitude and frequency modulations. *Proc Natl Acad Sci* 102:2293-2298.

Chapter 2 – Evidence From MEG

The study presented in this chapter was submitted for publication in *NeuroImage* on 08/02/2015:

Tang, H., Crain, S. and Johnson, W. B. (submitted). Evidence for dual temporal encoding mechanisms in human auditory cortex. *NeuroImage.*

Evidence for dual temporal encoding mechanisms in human auditory cortex

Huizhen Tang[a,b,1,2] **Stephen Crain** [a,b,3] **Blake W. Johnson** [a,b,2]

[a]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW 2109, Australia. http://www.ccd.edu.au, [b]Department of Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia. [c]Department of Linguistics, Macquarie University, Sydney, NSW 2109, Australia

[1] To whom correspondence may be addressed: joann.tang@mq.edu.au

[2] Present address: Department of Cognitive Science, Macquarie University, Sydney, NSW 2109

[3]Present address: Department of Linguistics, Macquarie University, Sydney, NSW 2109

Abstract

Current hypotheses about language processing advocate an integral relationship between encoding of temporal information and linguistic processing in the brain. All such explanations must accommodate the evident ability of the perceptual system to process both slow and fast time scales in speech. However most cortical neurons are limited in their capability to precisely synchronise to temporal modulation at rates faster than about 50 Hz. Hence, a central question in auditory neurophysiology concerns how the full range of perceptually relevant modulation rates might be encoded in the cerebral cortex. Here we show with noninvasive magnetoencephalography (MEG) measurements that the human auditory cortex transitions between a phase-locked (PL) mode of responding to modulation rates below about 50 Hz, and a non phase-locked (NPL) mode at higher rates. Precisely such dual response modes are predictable from the behaviours of single neurons in auditory cortices of non-human primates. Our data point to a common mechanistic explanation for the single neuron and MEG results and support the hypothesis that two distinct types of neuronal encoding mechanisms are employed by the auditory cortex to represent a wide range of temporal modulation rates. This dual encoding model allows slow and fast modulations in speech to be processed in parallel and is therefore consistent with theoretical frameworks in which slow temporal modulations (such as rhythm or syllabic structure) are akin to the contours or edges of visual objects, whereas faster modulations (such as periodicity pitch or phonemic structure) are more like visual texture.

Keywords: auditory cortex; auditory steady state response; envelope following response; temporal encoding; gamma band response

2.1 Introduction

At the present time there is considerable neuroscientific interest in how the human auditory system may be organized to encode, process and utilize various classes of temporal information contained in sounds, to aid in the construction of accurate perceptions of the auditory world and in the derivation of meaningful linguistic representations from the speech signal (Eggermont, 2001; Giraud et al., 2000; Giraud and Poeppel, 2012; Herdener et al., 2013). In particular many investigators are interested in how and where information about the sound envelope -- a class of temporal modulations at both relatively slow and fast rates between 1 Hz to 100 Hz – is represented in the human auditory system (Eggermont, 2001; Giraud et al., 2000; Peelle et al., 2012). Interest in the mechanisms of sound envelope encoding has been strongly stimulated in recent years by new hypotheses of language perception that posit an integral relationship between encoding of temporal information in speech and linguistic processing in the brain (Poeppel, 2003; Poeppel et al., 2008).

The most straightforward way to encode temporal modulations is to synchronize neuronal activity with the periodicity of a sound via phase-locking (PL) (Eggermont, 2014; Joris et al., 2004). Such PL activity is evident throughout the auditory system and is prevalent at lower levels of the auditory system, with auditory nerve fibres showing precise PL up to rates of 4 kHz (Eggermont, 2014; Wang et al., 2008). However the upper limit of stimulus-synchronization in single units declines progressively and markedly as one ascends the auditory pathway (Eggermont, 2014; Wang et al., 2008). By the level of the primary auditory cortex, PL in single units is limited to relatively low modulation rates, showing a low-pass temporal modulation transfer function (TMTF) with a high cutoff of about 50 Hz (Eggermont, 2001; Wang et al., 2003). This TMTF profile of phase-locked (PL) temporal coding is also observed in noninvasive measurements of human auditory cortical activity by electroencephalography/magnetoencephalography (EEG/MEG) studies measuring auditory steady-state responses (ASSR) to constant periodic stimuli (John and Picton, 2000; Picton et al,

2003; Roß et al, 2000), and envelope following responses (EFR) to stimuli in which amplitude modulation (AM) is swept continuously over a range of rates (Nodarse et al., 2011; Miyazaki et al., 2013; Picton et al., 2003; Purcell et al., 2004). These neurophysiological profiles roughly match the psychophysical TMTF for detection of AM in white noise (Viemeister, 1979). Such results suggest that the auditory system can be modeled as a simple low pass temporal filter (Viemeister, 1979; Rees et al., 1986; Ross et al., 2000) that is organized to capture and emphasize relatively low frequency temporal envelope information such as syllable rate periodicities in speech, and to filter or deemphasize temporal content that is communicated at higher rates (Doelling et al., 2014).

However higher modulation rates (above about 40 Hz, ranging as high as 500 Hz) contain temporal information crucial for perception of voiced speech sounds (e.g. /d/-/t/) as well as periodicity pitch (Joris et al., 2004; Langner, 1992; Rosen, 1992). Further, ablation studies in cats have shown that primary auditory cortex is essential for detection of periodicity pitch (Whitfield, 1980). Given the limited capacity of cortical neurons to precisely synchronise to temporal information at faster rates, a central question in auditory neurophysiology concerns how the full range of perceptually relevant modulation rates might be encoded in the cerebral cortex. At present it is not clear how higher modulations are represented at the level of the cortex in humans. Studies of single neurons in the auditory cortex of marmoset monkeys have shown that high modulation rates are associated with increases in non phase-locking (NPL) neuronal activity, leading to the hypothesis that the perceptually-relevant range of temporal periodicities is represented in cortex by a dual encoding scheme (Liang et al., 2002; Lu et al., 2001; Wang et al., 2003). On this model, the nature of envelope encoding changes: from an "explicit" or "isomorphic" mode of PL "temporal coding" at slow modulation rates, to an "implicit" or "nonisomorphic" mode of NPL "rate coding" (or possibly, "rate-place coding") of faster modulations (Joris et al., 2004; Liang et al., 2002; Lu et al., 2001; Wang et al., 2003).

To our knowledge, a NPL type of auditory response has not previously been described in the macroscopic measurements of EEG/MEG recordings of the ASSR/EFR in humans. It is notable that conventional methods for measuring the ASSR/EFR emphasise PL activity with continuous stimulation paradigms and time domain averaging of trials (Miyazaki et al., 2013; Purcell et al., 2004). These procedures improve the signal-to-noise ratio of PL activity within the recordings, but essentially discard any information about NPL activity. If the auditory cortex employs NPL representation to any significant extent, then the existing picture of human auditory cortical tuning provided by these measures must be incomplete. The current study was therefore designed to characterise the total magnitude of both PL and NPL MEG responses to AM across a range of perceptually-relevant modulation rates (see Figure 1 for details). Motivated by the dual encoding model from single unit neurophysiology (Liang et al., 2002; Lu et al., 2001; Wang et al., 2003), we predicted there should be substantial NPL activity in the envelope-following response of the MEG, particularly at rates that exceed the 50 Hz maximum that is optimal for PL responses (Picton et al., 2003).

2.2 Methods

2.2.1 Subjects

Twenty-eight adults (seventeen females) aged 22 to 36 years (mean = 29 years) participated in the study. Hearing thresholds were measured using an Otovation Amplitude T3 series audiometer (Otovation LLC, King of Prussia, PA). All subjects showed normal hearing thresholds ($\leq$ 20 dB HL) for octave frequencies from 500 to 2000 Hz. All procedures were approved by the Human Subjects Ethics Committee of Macquarie University.

2.2.2 Acoustic stimulation

The acoustic stimulus was created in MATLAB (Mathworks: Natickm, MA) by modulating a 9-second white noise with a 9-second sweep changing exponentially from 1 Hz to 80 Hz. The

modulation depth was 100%. The exponential function used to create the stimulus was: $f(t) = f_0 \times \left(\frac{f_1}{f_0}\right)^{\frac{t}{t_1}}$, where $f(t)$ is the frequency changing by time, t, f_0 = 1 (Hz) is the starting frequency, f_1= 80 (Hz) is the ending frequency, $t_1 = 9\ (second)$ is the sweep duration. The function therefore can be simplified as: $f(t) = 80^{t/9}$.

In order to separate the envelope following response from the responses to sound onset/offset, an unmodulated 0.3 second white noise was inserted both at the beginning and the end, resulting a total stimulus duration of 9.6 seconds. The temporal waveform of the complete stimulus and spectrogram of the envelope of the AM sweep are presented in Figure 1.

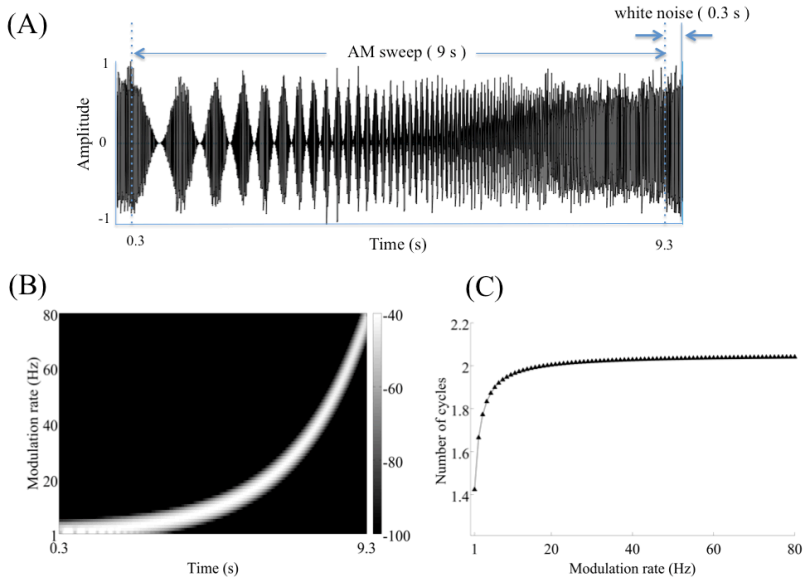

Figure 1. Characteristics of the acoustic stimulus. *(A) Stimulus waveform. Total sound duration was 9.6 s, consisting of 0.3 s of unmodulated noise, followed by AM sweeping from 1-80 Hz over 9.0 sec, and concluding with another 0.3 seconds of unmodulated white noise. The modulation rate of the AM sweep changed logarithmically from 1 Hz to 80 Hz over time from 0.3 second to 9.3 second. Modulation depth was 100%. Stimuli were repeated for 200 trials with a mean interstimulus interval of 950 ms. (B) Spectrogram of the envelope of the AM sweep. (C) The stimulus was designed to deliver about 2 cycles of each modulation rate.*

2.2.3 Procedure

Participants listened to the stimuli passively while watching a muted movie. Active attention has a relatively small effect on the magnitude of the ASSR (Ross and Pantev, 2004) and passive listening paradigms have been used in most previous studies of the auditory ASSR/EFR (see for example, Purcell et al., 2004; Ross et al., 2000). In addition we wished to employ the same

methods used by us in studies of the EFR in preschoolers (Tang et al., submitted for publication) which necessitated the use of procedures that place minimal demands on the attentional and motivational capacities of the children.

Stimulus presentation was controlled using Experiment Builder 1.10.165 (SR Research: Mississauga, Ontario, Canada). The same stimulus was presented repeatedly for 200 trials with an inter-stimulus interval (ISI) of ~950 ms (ISI was randomly selected from a rectangular distribution between 900 ms and 1000 ms) separately in two blocks. The total duration of the experiment was about 35 minutes. The sounds were delivered binaurally using insert earphones (Model ER-30, Etymotic Research Inc., Elk Grove Village, IL) at a level of 70 dB SPL.

2.2.4 MEG Recording

Brain activity was recorded continuously using a whole-head MEG system consisting of 160 axial gradiometers with a 50 mm baseline (Model PQ1160R-N2, KIT, Kanazawa, Japan) located in a magnetically shielded room (Fujihara Co. Ltd., Tokyo, Japan). All measurements were carried out with participants in a supine position. MEG data were acquired with a sampling rate of 1000 Hz and band-pass filtered between 0.03 Hz and 200 Hz using an analogue filter.

2.2.5 Structural MRI Scans

Anatomical images were acquired at the Macquarie University Hospital, Sydney, using a 3 Tesla Siemens Magnetom Verio scanner with a 12-channel head coil. Images were acquired using an MP-RAGE sequence (208 axial slices, TR = 2000 ms, TE = 3.94 s, FOV = 240 mm, voxel size = 0.9 mm^3, TI = 900, flip angle = 9°).

2.2.6 Analyses

MEG data were analysed using BESA Research Version 6.0 (BESA Research GmbH: Grafelfing, Germany). Structural MRIs were processed, warped into Talairach space, and co-registered with the MEG data using BESA MRI 2.0 (BESA Research GmbH: Grafelfing,

Germany). Data were analysed using a spatial filter consisting of bilateral dipoles fitted to the M100 component of auditory evoked fields (AEFs) elicited by the onsets of the acoustic stimuli. AEFs were epoched from -100 to 400 ms with respect to the onset of each stimulus. The epoched data was averaged across trials and band pass filtered between 2 and 20 Hz. Semi-symmetric dipoles were fitted within the M100 latency window of 100 - 200 ms (see Figure 2.A) for each subject (Schoonhoven et al., 2003). Here semi-symmetric refers to the left-right symmetric location but independent orientation of the two dipoles (Schoonhoven et al., 2003). Dipoles for all participants were located in the posterior superior temporal gyri (Figure 2), clustered around Brodmann areas 22 (primary auditory cortex) and 41 (auditory association cortex). The group mean dipole locations was (-50.3 ± 4.8, -20.2 ± 6.4, 4.2 ± 6.9) in the left hemisphere and (50.3 ± 4.8, -20.2 ± 6.4, 4.2 ± 6.9) in the right hemisphere.

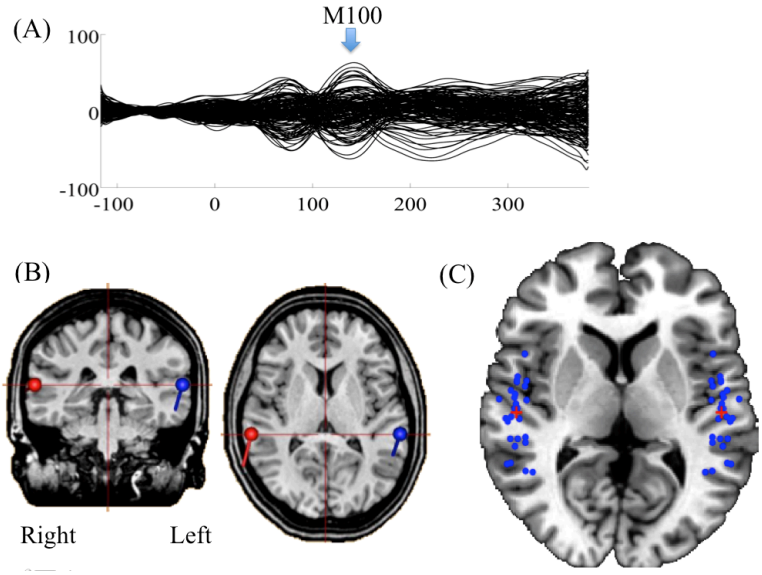

Figure 2 Individual dipole sources and group mean. (A) Overlaid sensor waveform epoched from -100 to 400 ms with respect to the onset and averaged across trials of a representative participant. (B) Semi-symmetric bilateral dipoles localized based on M100 of a representative participant using individual MRI. The bar of the dipole represents the source strength and its orientation represents the orientation of the equivalent dipole source. (C) Semi-symmetric bilateral dipoles of all the participants (blue dot) projecting on the axial plane defined by the group mean location (red cross) in the MNE MRI.

A bilateral two-dipole source montage was applied to the raw MEG data of each individual to reconstruct the source waveform from each dipole in each hemisphere. Time-frequency analyses were then carried out on the continuous source waveforms over a frequency

range of 1 to 80 Hz within an epoch window from 600 ms pre-stimulus to 9600 ms post-stimulus. ERSPs were computed using the complex demodulation method (Hoechstetter et al., 2004), with a frequency step of 1 Hz and a time step of 50 ms. The ERSP (Brugge et al., 2009; Makeig, 1993; Pfurtscheller and Silva, 1999) quantifies the increase or decrease in amplitude/power in a given frequency or frequency band with reference to the ongoing background brain activity recorded in the defined baseline. Amplitude was normalized independently in each frequency or frequency band. Specifically, the ERSP index is given by the following formula.

$$ERSP(t,f) = \frac{\sum_{n=1}^{N}\left(\frac{A_n(t,f) - \overline{A_{baseline,\ n}(f)}}{\overline{A_{baseline,\ n}(f)}} \times 100\%\right)}{N}$$

where $A_n(t,f)$ is the absolute amplitude at time t and frequency f in trial n; and $\overline{A_{baseline,\ n}(f)}$ is the mean absolute amplitude at frequency f over the defined baseline in trial n.

The phase-locked component of ERSP (*PL)* was computed by applying spectral analysis on averaged waveform, which is obtained from averaging the MEG signals across trials, using the following formula:

$$PL(t,f) = \frac{\bar{A}(t,f) - \overline{\overline{A_{baseline}}(f)}}{\overline{\overline{A_{baseline}}(f)}} \times 100\%$$

where $\bar{A}(t,f)$ is the absolute amplitude at time t and frequency f of the averaged signal; and $\overline{\overline{A_{baseline}}(f)}$ is the mean absolute amplitude of the averaged signal at frequency f over the defined baseline.

The non-phase-locked component of ERSP (*NPL*) was then computed by subtracting the *PL* from the amplitude change in single trial as indicated by the following formula:

$$NPL(t,f) = \frac{\sum_{n=1}^{N}\left(\frac{A_n(t,f) - \overline{A_{baseline,\ n}(f)}}{\overline{A_{baseline,\ n}(f)}} \times 100\% - PL(t,f)\right)}{N}$$

The inter-trial phase locking (*ITPC*), which quantifies phase consistency cross trials for each frequency and time point, was computed using the following formula.

$$ITPC(t,f) = \frac{1}{N}\sum_{n=1}^{N} e^{i\varphi_k(t,f)}$$

with N = total number of trials; and $\varphi_k(f,t)$ = phase in trial k.

Permutation tests were then carried out on the time-frequency data based on previously developed algorithms (Maris and Oostenveld, 2007) using custom MATLAB scripts to examine whether the envelope following pattern was statistically significant. The permutation tests were applied to response between 0.3 second after the stimulus onset to 9.3 second, during which time the temporal modulation of the stimulus swept from 1 Hz to 80 Hz. During permutation, a primary threshold was used to identify the top 5% values. The clustering algorithm used the sum of activities within a cluster (instead of the size of a cluster). Only clusters surviving a permutation with 1000 iterations at the significant level of 0.05 were accepted. To better visualize the relationship between stimulus modulation rate and brain response, a linearization procedure was applied to convert the time axis to modulation rate, based on the stimulus spectrogram function $f(t) = 80^{t/9}$. Model fitting was then performed in MatLab R2014a (the MathWorks Inc., Natick, MA, USA) using the Curve Fitting Toolbox (version 3.4.1) to model the correlation between modulation rate and frequency of brain response. A linear model ($f = a * m + b$; here f is the frequency of brain response, m is the rate of modulation, a and b are two parameters set for free fitting) was used to fit the permuted data. EFR to AM at rate *m*, was identified as the mean magnitude within the bin [*f-1 f+1*]. Vector strength of the defined EFR was then re-plotted against the frequency of the AM to visualize the TMTFs. A 3 Hz smoothing window was applied to TMTF plots for visualization only.

Being the first to use a discrete stimulation with an inter-stimulus interval using a rapid sweeping stimulus like the one we employed, one concern would be whether the current paradigm is sufficient for eliciting modulation-specific following responses. To this end, we

divided the total number of trials (i.e. ~ 180 - 200) into four subsets chronically, with about 45-50 trials per subset. All time-frequency analyses described above were conducted for each subsets and difference between subsets were tested for statistically significance by performing permutation tests on the difference between two spectrograms of the same measure. If the rapid varying AM sweep employed in present study was sufficient to elicit following response in signal trial and the reset of auditory cortex during the ITI was not an interruption of temporal encoding, then no time effect should be observed and the dual temporal representations should be consistent across the four subsets.

2.3 Results

Computation of the PL activity, by averaging single trials in the time domain before computing the event-related spectral perturbations (ERSP) (Brugge et al., 2009; Makeig, 1993; Pfurtscheller and Silva, 1999), shows that our data replicate the low-pass TMTF reported in many previous EEG/MEG studies (Picton et al., 2003; Purcell et al., 2004), with a prominent peak around 40 Hz followed by a marked decline in magnitude above 50 Hz (see Figure 3.a-c). However the shape of the TMTF is considerably altered when total ERSP magnitude is computed (Figure 3.d-f). Here the TMTF profile measured with ERSP reflecting both PL and NPL activity, shows strong responses following the temporal patterns of the AM throughout the range of frequencies from 30-80 Hz (see Figure 3.d-f). The ERSPs at lower frequencies (1-30 Hz) were much weaker and accompany with a significant beta band (13-25 Hz) desynchronization. This behaviour is difficult to accommodate within a model that views the auditory cortex as a simple low pass filter for temporal information.

When both the NPL and PL components of the total ERSP were separated and plotted against modulation rate (Figure 4.c), we obtained a dual response profile that strongly resembles that described for single unit activity in auditory cortex of non human primates (Liang et al., 2002; Lu et al., 2001; Wang et al., 2003). This profile shows a predominantly PL

mode of responding at rates below 50 Hz, transitioning to a predominantly NPL mode of responding at rates above 50 Hz, and with an intermediate transition zone between 30 and 50 Hz where both PL and NPL modes are prominent (see Figure 4).

When the total number of trials was partitioned chronically to four subsets, the PL and NPL dual responses were consistently demonstrated in all the four subsets of data (Figure 5 and 6). No significant difference was found based on permutation tests on the spectrogram of the difference between subsets (Figure 6.A).

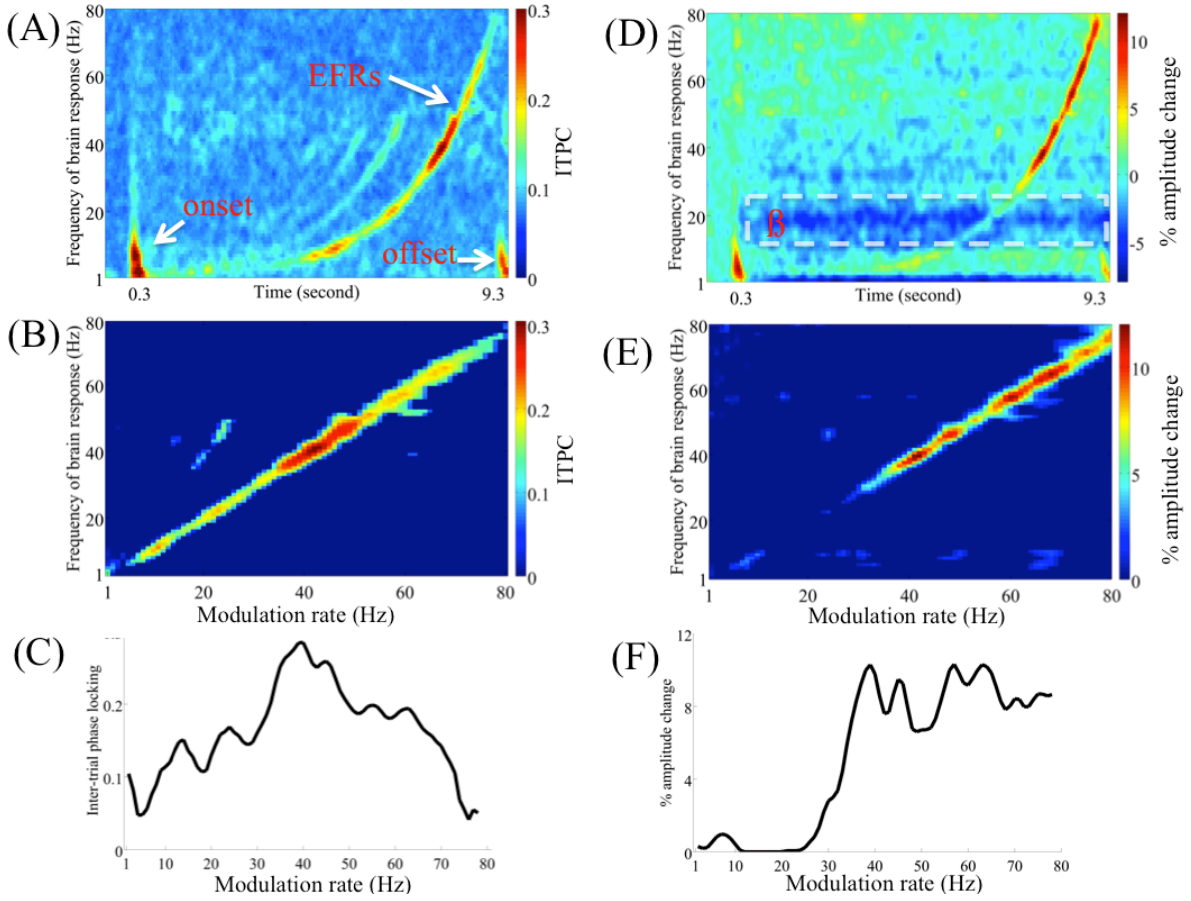

Figure 3. Event-related spectral perturbations in auditory cortex in response to the range of AM rates. *Left plots show inter-trial phase coherence (ITPC) values, right plots show total amplitude change from the prestimulus baseline. (a) ITPC over the analysis epoch. Responses to onset and offset of the stimulus are seen at about 100 ms and 9400 ms respectively. (b) Linearised and permuted ITPC values. (c) Temporal modulation transfer function (TMTF) for ITPC as a function of AM rate. (d) Total amplitude change from baseline over the analysis epoch. (e) Linearised and permuted total amplitude values. (f) Temporal modulation transfer function (TMTF) for total amplitude change as a function of AM rate.*

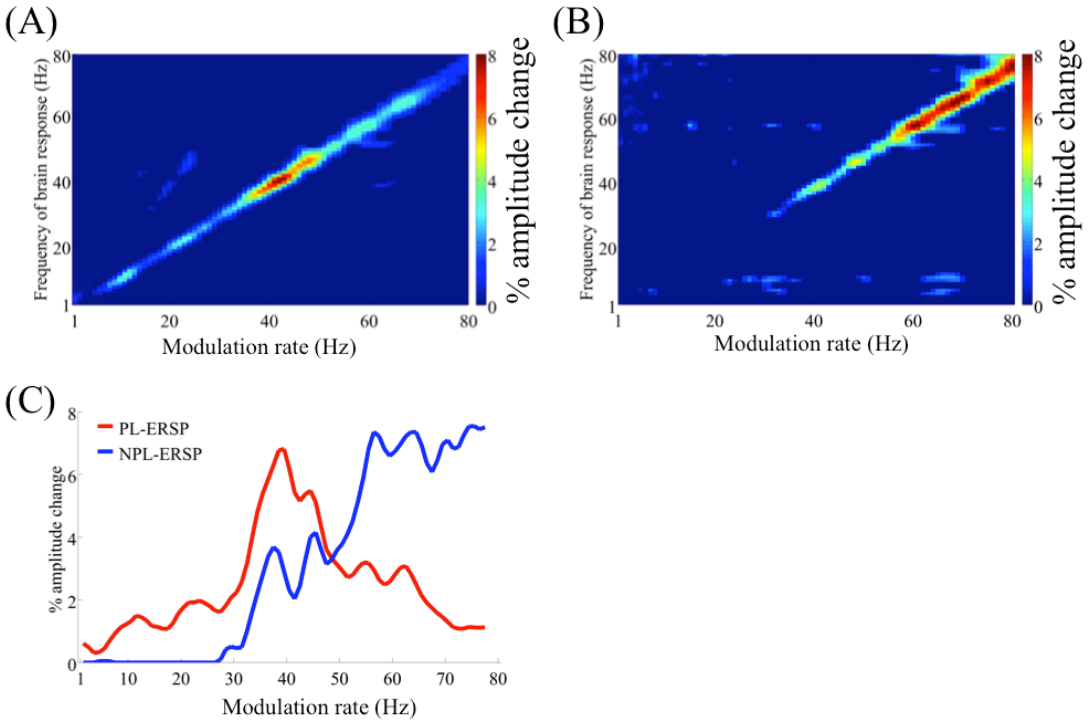

Figure 4. Phase-locked (PL) and non-phase-locked (NPL) components of event-related spectral perturbations (ERSPs). *All plots show group and hemisphere averaged MEG responses from bilateral auditory cortical sources. (a) Permuted and linearised spectrogram of PL component obtained by averaging individual trials in the time domain. Maximum amplitude changes are in the 30 - 50 Hz range of modulation rates. (b) Permuted and linearised spectrogram of NPL component obtained by subtracting PL component from individual trials. Maximum amplitude changes are in the ~55 - 80 Hz range of modulation rates. (c) PL and NPL components of the ERSP TMTF function. All non-zero data points are statistically significant ($p < 0.05$).*

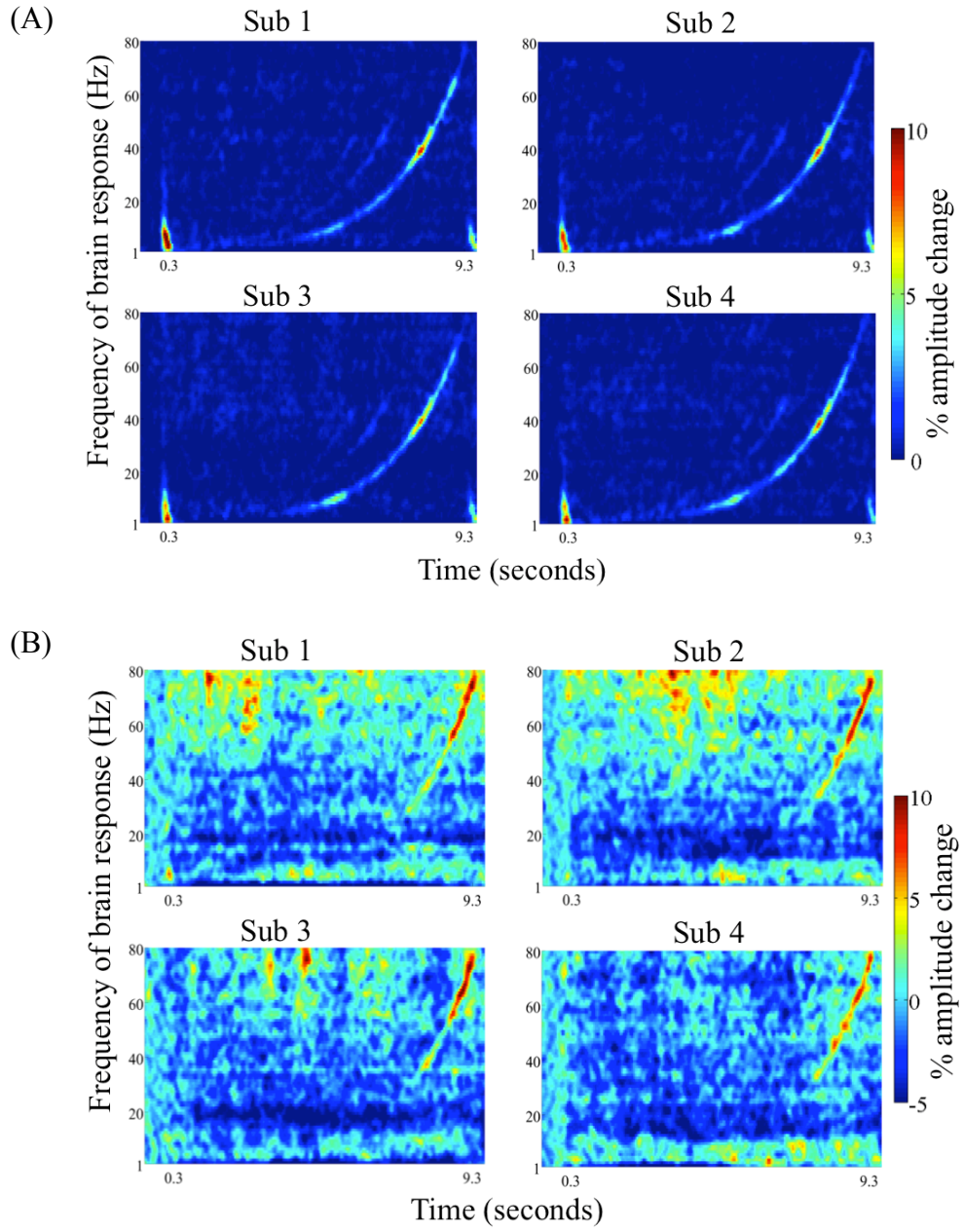

Figure 5. Phase-locked (PL) and non-phase-locked (NPL) event-related spectral perturbations (ERSPs) of four subsets partitioned chronically from the total dataset. *(A) PL component of ERSPs in four subsets. (B) NPL component of ERSPs in four subsets. Part 1-4 stands for the first ¼, second ¼, third ¼, and the four ¼ respectively.*

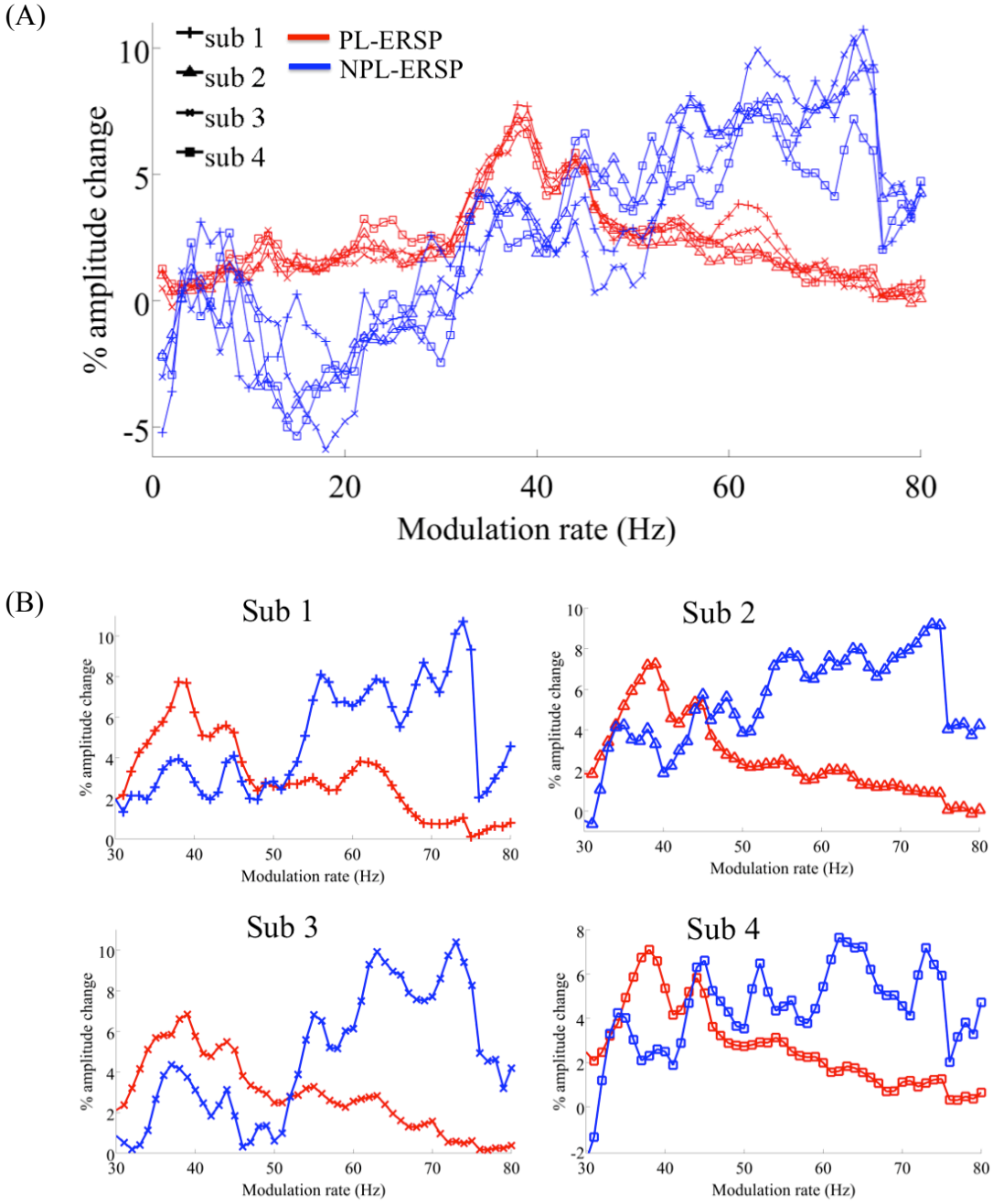

Figure 6. TMTFs of phase-locked (PL) and non phase-locked (NPL) evented-related spectral perturbations (ERSPs) of four subsets partitioned chronically from the total dataset. *(A) Overlaid TMTFs of four subsets. (B) Dual PL and NPL responses in gamma band range (30 – 80 Hz) in individual subsets.*

2.4 Discussion

The human auditory SSR/EFR (hereafter, referred to by the more general term of EFR) has been the subject of considerable neuroscientific and clinical interest since interest in these responses was stimulated by Galambos et al.'s 1981 report on the auditory 40 Hz response. Since then these responses have been extensively characterised with EEG and MEG

measurements over a wide range of stimulation rates. There is now substantial data to show that the conventional phase-locked auditory EFR -- elicited with continuous stimulation and time domain averaging – falls off steeply in magnitude at stimulation rates above 50 Hz. Below 50 Hz, the frequency characteristic shows a maximal amplitude of response at rates of about 40 Hz (the auditory 40 Hz response). Using non-continuous stimulation (to provide a reference baseline of prestimulus activity) and averaging in the frequency domain (to preserve non phase locked as well as phase locked responses), we show that the total amplitude change from baseline does not decline for rates above 50 Hz and is maintained at high levels to the maximal stimulation rate of 80 Hz. On the other hand, the phase-locked portion of this amplitude change precisely replicates the conventional TMTF of the auditory EFR described above (see reviews by Picton, 2003; Ross, 2103).

These results are interesting first because they show that the conventional phase-locked auditory EFR which has attracted substantial neuroscientific interest and clinical application (e.g., Galambos et al.'s 1981 paper has been cited nearly 1000 times to date, and steady-state auditory EFRs are now widely used in clinical audiometric testing) provides only a partial and incomplete picture of how the human auditory cortex responds to rapidly presented sounds. They are also interesting because they imply that the auditory cortex uses two distinct kinds of neuronal mechanisms to encode temporal modulations in high versus low temporal modulation rates in sounds. This possibility has been proposed in a variety of contexts in the last several decades and is consistent with a substantial amount of data from human psychophysics, functional MRI, and invasive studies in animals and humans.

2.4.1 A novel response property of the auditory EFR

Our results describe a novel property of the MEG brain response to rapidly amplitude-modulated sounds, a non-phase locked oscillatory response following the AMs at rates higher than 30 Hz. This observation stands in contrast to a considerable and long-standing literature on auditory steady-state responses, auditory envelope following responses, and auditory transient

gamma band responses that all emphasize strongly phase-locked oscillatory responses in the low gamma range of 30 – 50 Hz (Galambos et al., 1981, Makeig, 1993, Roß et al., 2000, Brosch et al., 2002, Picton et al., 2003, Purcell et al., 2004, Mijares Nodarse et al., 2011), with a rapid fall off in the magnitude at rates higher than about 50 Hz. Since the high rates EFR is not phase-locked its absence in previous descriptions of the ASSR is not surprising, since these have invariably emphasized phase-locked signals to increase signal to noise ratio by averaging of signals in time domain or by extracting phase-locked signals in the frequency domain. Both of these analyses will destroy signals that are not strongly phase-locked to the inducing stimulus.

The novel rate-sensitive, non phase-locked EFR described here provides a new look to the conventional auditory EFR, in several respects. First, this response strongly suggests a connection to the non-phase locked high gamma band auditory responses to *transient stimuli* that were initially described with invasive ECoG recordings in human patients (Brugge et al., 2009; Edwards et al., 2009; Crone et al., 2001; Crone et al., 2011; Nourski et al., 2009; Nourski et al., 2013) and in non-human primates (Brosch et al., 2002; Steinschneider et al., 2008) and more recently with non-invasive MEG recordings of auditory transient responses in healthy humans (Sedley et al., 2012). In addition to their non-phase-locked nature, the properties of the "high gamma band" responses differ importantly from those in the low (circa 40 Hz) gamma band, with stronger responses to linguistic than to non-linguistic stimuli (Crone et al., 2001) and with more susceptibility to top-down influences of higher level processes (Hermann et al., 2010). Thus, the NPL EFR can potentially provide a neural marker of higher level cognitive processes than are indexed by the PL EFR.

This more complete picture of auditory EFRs may also bear importantly on our understanding of the neural generators of both the new NPL EFR and those of the conventional PL EFR, which have been extensively debated but never completely resolved. In particular, the dual PL/NPL response profile of the EFT is well explained by mechanisms recently described

in *in vitro* preparations of auditory cortex of the rat. Ainsworth and colleagues (2012) reported two fundamentally different distinct local circuit generators of low and high gamma rhythms. Low levels of cortical stimulation (i.e. increases in firing rate caused by application of the glutamate agonist kainite) resulted in a frequency-stable low gamma frequency (30-45 Hz) rhythm originating in layers 2/3 supragranular layers, while higher levels of stimulation resulted in a "frequency-labile" high gamma frequency (50 – 80 Hz) rhythm originating in granular layer 4. Notably, the frequency of the local field potential (LFP) high gamma rhythm increased as a direct function of the firing rates of principal cells. The rate-sensitivities of these two gamma rhythms match remarkably with those of our PL (phase-stable and maximal at stimulus repetition rates of 40 Hz) and NPL (non phase stable and increasing as a function of repetition rate greater than 30 Hz). If we accept that rate coding is employed for higher temporal modulation rates (see discussion below) then it is apparent that both the high gamma rhythm of rat auditory cortex and the high rates NPL response in human cortex are both modulated directly by increases in firing rate.

One further point of interest is that both the high gamma rhythm of the rat cortex and the NPL response are narrow band phenomena: that is, they both exhibit a continuous range of gamma frequencies as a direct response to increasing driving input. In contrast, the high gamma responses to transient acoustic stimulation described above are typically responses of increased gamma power in fairly broad frequency bands of 60-80 Hz or much higher (Edwards et al., 2009; Crone et al., 2001; Nourski et al., 2009; Brugge et al., 2009; Nourski et al., 2013). Ainsworth et al. (2012) reconciled these gamma responses by suggesting that different components of a complex acoustic stimulus will drive subsets of layer 4 neurons at different gamma frequencies (Kaiser et al., 2008), resulting in a broadband response across a nearly continuous range of high gamma frequencies with naturalistic stimuli (Uhlhaas et al., 2011).

One may argue whether the novel findings of the present study are due to the fact we used a non-conventional paradigm for eliciting EFRs. We first would like to point out that both

the ITPC and PL metrics, which provide different measures phase-locked responses, closely replicated the typical low-pass filter profile with a cut off point at higher rates reported in previous studies (Ross et al., 2000, Picton et al., 2003, Purcell et al., 2004, Miyazaki et al., 2013). Moreover, the PL metrics in the four chronically partitioned subsets of data showed a close resemblance, which indicated no time related enhancement or suppression. These together suggest that the current paradigm, consisting of a rapidly sweeping AM stimulus and a discrete stimulation procedure was sufficient for eliciting frequency-specific following responses that exhibit the capacity to track the temporal dynamics of the acoustic signal. Furthermore, the insertion of an inter-trial-interval (ITI) between each sweep allowed us to investigate how the brains encode the rapidly changing temporal dynamics after being reset to a quasi-resting state.

2.4.2 Dual temporal encoding mechanisms

Lu and colleagues (2001) reported two distinct populations of neurons in auditory cortex of non-anesthetised marmoset monkeys: a "synchronised" population that responded to acoustic stimulation with by synchronised its firing to the stimulus repetition rate; and a "non-synchronised" population that responded to increases in stimulus rate with increases in firing rate that were not synchronised to the stimulus rate. The synchronised population was activated most strongly by low stimulus rates and responding fell off sharply at rates higher than about 50 Hz; In contrast the non-synchronised population responded most strongly to stimulus rates above 50 Hz (this 50 Hz inflection point is remarkably similar to that seen for our PL and NPL responses). These authors interpreted these results to reflect the operation of two distinct mechanisms for encoding temporal information in sounds at the level of the auditory cortex: a temporal code that explicitly represents relatively slowly changing sound sequences (less than 50 Hz), and a rate code that implicitly represents more rapidly occurring sequences.

The use of two fundamentally distinct methods of encoding fast and slow times scales provides a parsimonious neurophysiological explanation for how we actually perceive temporal phenomena (Moore et al., 2001). Sounds presented at slow repetition rates (below about 30 Hz)

are perceived as individual events, consistent with a high fidelity representation of timing (Miyazaki et al., 2013; Nourski and Brugge, 2011). This type of representation appears to be limited to modulation frequencies associated with the perception of rhythm and the syllable structure of speech (Giraud and Poeppel, 2012; Poeppel, 2003; Poeppel et al., 2008). On the other hand fast periodic repetitions above 25 - 30 Hz are perceived as a single continuous sound (Miyazaki et al., 2013), consistent with a nonisomorphic encoding scheme that captures overall temporal rate but discards the individual identities of events. Fast temporal rates in the range 40 - 500 Hz (Joris et al., 2004; Nourski and Brugge, 2011; Rosen, 1992) are important for the perception of periodicity pitch and underlie the discrimination of phonemes in speech (Crone et al., 2001; Eggermont, 2001; Langner, 1992). There are also intermediate perceptions like roughness, which has the strongest sensation at intermediate temporal rates of about 40 - 70 Hz, and may be explained by the co-existence of dual neuronal encodings (Langner, 1992; Wang et al., 2008).

2.4.3 Beta band suppression

Our data showed a striking suppression of lower beta band activity across all AM rates from stimulus onset to the end of the recording epoch (Fig. 3d). Very similar beta suppression by acoustic stimulation has been reported with intracranial recordings in human auditory cortex by Edwards et al. (2008; see Figure 2b); with noninvasive MEG recordings by Howard and Poeppel (2012; see Figure 3c,d and Figure 4b) and Kaiser et al. (2002); and noninvasive EEG recordings by Cacace and MacFarland (2003). Since beta suppression is classically associated with sensorimotor cortical activity (Engel and Fries, 2010), this effect suggests activation of the motor system during the presentation of the AM sound in our experiment, although no motor response was required in our passive listening task (nor in that of Edwards et al, 2008). A possible explanation comes from intracranial recordings from both humans and non-human primates implicating the dorsal auditory pathway in the automatic activation of articulatory representations and more generally in the preparation of behavioural responses to sounds

(Warren et al. 2005) and predictive coordination between auditory and motor systems (Fujioka et al. 2009; Fujioka et al. 2012). During a rhythm perception task, beta-band "rebound" (referring to the stronger beta rhythm which appears after the period of stimulus-induced beta suppression) adapts to rhythm timing in both the auditory and motor cortices (Fujioka et al. 2009; Fujioka et al. 2012). For humans auditory-motor coordination may be particularly relevant for coordination of speech perception and speech production mechanisms, and rhythm perception and movement production mechanisms. However since no behavioural response was required and attention was directed away from the sounds in our passive listening task (and also in the tasks of Fujioka et al. 2009, 2012) this may indicate that the auditory system maintains an "open-line" with the motor system, such that passive listening to sounds with potential biological relevance leads to an activation of motor systems associated with the automatic preparation of orienting (Kaiser et al., 2002).

2.4.4 Methodological considerations and limitations

On its own, the estimate of the NPL component should be viewed with some caution because this calculation is susceptible to artifactual inflation by trial-to-trial variability of PL activity (Brugge et al., 2009; David et al., 2006). We emphasize that *this problem does not apply to our calculations of total ERSP magnitude* (Brugge et al., 2009; Steinschneider et al. 2008), so our main observation – that total amplitude change from baseline does not decline for rates above 50 Hz and is maintained at high levels to the maximal stimulation rate of 80 Hz – is not affected by this consideration. Further, our calculations of PL magnitudes (Figure 4.a) are supported by an extensive literature on the human ASSR/EFR which have shown comparable or identical PL profiles (see reviews by Picton et al., 2003; Ross, 2014). Taken together, these two considerations strongly support our contention that the conventional phase-locked EFR provides only a partial picture of the human auditory system's response to repetitive high rates stimulation.

A reviewer of this article has pointed out that our auditory EFRs may reflect modulation from the visual system because participants were asked to attend to a muted movie while ignoring the acoustic stimulus. Indeed, studies of cross-modal selective attention have demonstrated that directing attention to a visual scene has inhibitory effects on auditory cortical processing (Foxe et al., 2014; Gomez-Ramirez et al., 2011). However, since the inhibitory effects reported by Foxe et al. (2014) and Gomez-Ramirez et al. (2011) were manifested in the low EEG frequencies (specifically, as increases in alpha band amplitude for ignored auditory stimuli during attention to the visual modality) they do not affect the main conclusions of our study which are based on the gamma frequency ranges. However such effects underline the necessity for caution in selection of experimental tasks and indicate that future studies should investigate the effects of attentive listening on the auditory EFR. As alluded to above, these effects may well be different for the PL and NPL components of the EFR (Hermann et al., 2010).

In addition, gamma band activity was observed in the NPL spectrogram (see Figure 5B) at higher frequencies (i.e. ~60 to 80 Hz). More specifically, there was substantial amount of gamma activity exhibited in both Sub 1 and Sub 2. This was reduced in Sub 3 and it almost completely disappeared in Sub 4. The nature of this high gamma activity might be related to the muted movie the subjects watched through the experiment. It has been found in a recent ECoG study that macaque monkeys had visual cortical gamma-band activity during free viewing of natural images (Brunet et al., 2015). The reduction across chronically defined subsets could be caused by the decrease of attention over time (Müller et al., 2000). Whether these broadband gamma activities would have enhanced the NPL in the corresponding frequencies is a question that needs to be addressed in future studies. This involve a comparison between conditions, with the same auditory stimilation being presented with or without visual input.

An important limitation of the present results is that the interpretation of the low frequency responses (below 50 Hz) are complicated by at least two factors. First, the 1/f frequency characteristic of the background EEG/MEG sharply reduces the signal to noise ratio of brain responses at low rates; second, responses in the low frequencies are further obscured by the sustained beta-band desynchronisation for the entire duration of the acoustic stimulus. Taken together, these factors strongly limited our ability to resolve brain responses to rates below about 30 Hz in the present experiment. Again however, these limitations do not affect our main observation that total amplitude change from baseline does not decline for rates above 50 Hz.

Finally, the scope of the present results is somewhat limited by the use of simple AM modulated noise stimuli, adopted because their spectral and temporal characteristics are well-controlled (Lu et al., 2003). However as indicated above the results have clear relevance to our understanding of neurophysiological responses elicited by naturalistic sounds (Uhlhaas et al., 2011; Ainworth et al., 2012). Further, we suggest that the results have strong relevance to current neurophysiological theories of language processing, which must explain our ability to simultaneously process temporal information at different times scales in the speech stream. A prominent and influential explanation is the asymmetric sampling in time (AST) theory proposed by Poeppel et al. (2003), which posits separate neural oscillatory mechanisms that are tuned to track and sample slow and fast temporal features (such as syllables and phonemes, respectively). The PL and NPL mechanisms reported by Ainsworth et al. (2012) and suggested by the present results are entirely consistent with such a temporal sampling framework.

2.5 Conclusions

These results demonstrate that our current picture of the auditory EFR measured by EEG/MEG (Picton, 2003; Ross et al., 2014) is incomplete, and further suggest that two distinct neural mechanisms are used to encode slow and rapid temporal modulations in sounds. This dual

temporal encoding framework provides an efficient explanation for how we actually perceive sounds and speech at different time scales, and shows important linkages between mechanisms reflected in macrosopic MEG measurements in healthy humans, micro-level measurements from in animal models, and meso-level ECoG measurements from human patients. This framework also opens new windows for research that posits a temporal processing deficit in certain language disorders, such as specific language impairment and dyslexia, where investigations to date have considered only the conventional phase-locked auditory following response (Lehongre et al., 2011).

Acknowledgements. We gratefully acknowledge the role of the Kanazawa Institute of Technology in establishing the KIT-Macquarie Brain Research Laboratory. This work was supported by the HearingCRC and Australian Research Council Grants CE110001021 and DP1096160.

2.6 References

Ainsworth, M., Lee, S., Cunningham, M. O., Roopun, A. K., Traub, R. D., Kopell, N. J., and Whittington, M. A. (2011). Dual gamma rhythm generators control interlaminar synchrony in auditory cortex. *Journal of Neuroscience*, *31*(47), 17040–17051.

Brosch, M., Budinger, E., and Scheich, H. (2002). Stimulus-related gamma oscillations in primate auditory cortex. Journal of Neurophysiology, 87(6), 2715–2725.

Brugge, J. F., Nourski, K. V., Oya, H., Reale, R. A., Kawasaki, H., Steinschneider, M., and Howard III, M. A. (2009). Coding of repetitive transients by auditory cortex on Heschl's gyrus. Journal of Neurophysiology, 102(4), 2358-2374.

Brunet N, Bosman CA, Roberts M, Oostenveld R, Womelsdorf T, De Weerd P, Fries P (2015) Visual cortical gamma-band activity during free viewing of natural images. *Cereb Cortex*, *25*, 918-926.

Cacace, A. T., and McFarland, D. J. (2003). Spectral dynamics of electroencephalographic activity during auditory information processing. *Hearing Research*, *176*(1-2), 25–41.

Crone, N. E., Boatman, D., Gordon, B., and Hao, L. (2001). Induced electrocorticographic gamma activity during auditory perception. *Neurophysiol Clin, 112*, 565-582.

Crone, N. E., Korzeniewska, A., and Franaszczuk, P. J. (2011). Cortical gamma responses: searching high and low. *Int J Psychophysiol, 79*(1), 9-15.

David, O., Kilner, J. M., and Friston, K. J. (2006). Mechanisms of evoked and induced responses in MEG/EEG. *NeuroImage, 31*(4), 1580-1591.

Doelling, K. B., Arnal, L. H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage, 85 Pt 2*, 761-768.

Echols CH, Crowhurst MJ, Childers JB (1997) The Perception of Rhythmic Units in Speech by Infants and Adults. *Journal of Memory and Language. 36*, 202–225.

Edwards, E., Soltani, M., Kim, W., Dalal, S. S., Nagarajan, S. S., Berger, M. S., and Knight, R. T. (2009). Comparison of time-frequency responses and the event-related potential to auditory speech stimuli in human cortex. *Journal of Neurophysiology*, *102*(1), 377–386.

Eggermont, J. J. (2001). Between sound and perception: reviewing the search for a neural code. *Hear Res, 157*(1), 1-42.

Eggermont, J. J. (2014). Animal models of auditory temporal processing. *Int J Psychophysiol*. doi: 10.1016/j.ijpsycho.2014.03.011

Foxe, J. J., Murphy, J. W., and de Sanctis, P. (2014). Throwing out the rules: anticipatory alpha-band oscillatory attention mechanisms during task-set reconfigurations. *European Journal of Neuroscience*, *39*(11), 1960–1972.

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2009). Beta and gamma rhythms in human auditory cortex during musical beat processing. *Annals of the New York Academy of Sciences*, *1169*, 89–92.

Fujioka, T., Trainor, L. J., Large, E. W., and Ross, B. (2012). Internalized timing of isochronous sounds is represented in neuromagnetic beta oscillations. *Journal of Neuroscience*, *32*(5), 1791–1802.

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.

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.

Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature neuroscience, 15*(4), 511-517.

Gomez-Ramirez, M., Kelly, S. P., Molholm, S., Sehatpour, P., Schwartz, T. H., and Foxe, J. J. (2011). Oscillatory sensory selection mechanisms during intersensory attention to rhythmic auditory and visual inputs: a human electrocorticographic investigation. *Journal of Neuroscience*, *31*(50), 18556–18567.

Herdener, M., Esposito, F., Scheffler, K., Schneider, P., Logothetis, N. K., Uludag, K., and Kayser, C. (2013). Spatial representations of temporal and spectral sound cues in human auditory cortex. *Cortex; a journal devoted to the study of the nervous system and behavior, 49*(10), 2822-2833.

Herrmann, C. S., Munk, M. H., and Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in cognitive sciences, 8*(8), 347-355.

Herrmann, C. S., FrÃ¼nd, I., and Lenz, D. (2010). Neuroscience and Biobehavioral Reviews. *Neuroscience and Biobehavioral Reviews*, *34*(7), 981–992.

Hoechstetter, K., Bornfleth, H., Weckesser, D., Ille, N., Berg, P., and Scherg, M. (2004). BESA Source Coherence: A New Method to Study Cortical Oscillatory Coupling. *Brain topography, 16*(4), 233-238.

Howard, M. F., and Poeppel, D. (2012). The neuromagnetic response to spoken sentences: Co-modulation of theta band amplitude and phase. *NeuroImage*, *60*(4), 2118–2127.

John, M. S., and Picton, T. W. (2000). Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. *Hear Res, 141*(1–2), 57-79.

Joris, P., Schreiner, C., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol Rev, 84*(2), 541-577.

Kaiser, J., Birbaumer, N., and Lutzenberger, W. (2002). Magnetic oscillatory responses to lateralization changes of natural and artificial sounds in humans. *European Journal of Neuroscience*, *15*(2), 345–354.

Kaiser, J., Heidegger, T., Wibral, M., Altmann, C. F., and Lutzenberger, W. (2008). Distinct gamma-gand components reflect the short-term memory maintenance of different sound lateralization angles. *Cerebral Cortex*, *18*(10), 2286–2295.

Langner, G. (1992). Periodicity coding in the auditory system. *Hear Res, 60*(2), 115–142.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., and Giraud, A. L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron, 72*(6), 1080-1090.

Liang, L., Lu, T., and Wang, X. (2002). Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. *Journal of Neurophysiology, 87*(5), 2237-2261.

Lu, T., Liang, L., and Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nature Neuroscience, 4*(11), 1131-1138.

Makeig, S. (1993). Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr Clin Neurophysiol, 86*, 283-293.

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

Mijares Nodarse, E., Perez Abalo, M. C., Torres Fortuny, A., and Vega Hernandez, M. (2011). Electrophysiological Characterisation of Envelope-Following Responses. *Acta Otorrinolaringologica, 62*(6), 425-431.

Miyazaki, T., Thompson, J., Fujioka, T., and Ross, B. (2013). Sound envelope encoding in the auditory cortex revealed by neuromagnetic responses in the theta to gamma frequency bands. *Brain Research, 1506*, 64-75.

Moore, D. R., Schnupp, J. W. H., and King, A. J. (2001). Coding the temporal structure of sounds in auditory cortex. Nature Neuroscience, 4(11), 1055-1056.

Müller MM, Gruber T, Keil A (2000) Modulation of induced gamma band activity in the human EEG by attention and visual information processing. *Int J Psychophysiol, 38*, 283-299.

Nourski, K. V., and Brugge, J. F. (2011). Representation of temporal sound features in the human auditory cortex. *Reviews in the Neurosciences, 22*(2), 187-203.

Nourski, K. V., Brugge, J. F., Reale, R. A., Kovach, C. K., Oya, H., Kawasaki, H., Jenison, RL Howard III, M. A. (2013). Coding of repetitive transients by auditory cortex on

posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. *Journal of Neurophysiology, 109*(5), 1283-1295.

Pantev, C., Makeig, S., Hoke, M., Galambos, R., Hampson, S., and Gallen, C. (1991). Human auditory evoked gamma-band magnetic fields. *Proceedings of the National Academy of Sciences*, *88*(20), 8996–9000.

Peelle, J. E., Gross, J., and Davis, M. H. (2012). Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. *Cerebral cortex*.

Pfurtscheller, G., and Silva, F. H. L. d. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clinical neurophysiology, 110*, 1842-1857.

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int J Audiol, 42*(4), 177-219.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Communication, 41*(1), 245-255.

Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Lond, 363*(1493), 1071-1086.

Purcell, D. W., John, S. M., Schneider, B. A., and Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am, 116*, 3581.

Rees, A., Green, G. G., and Kay, R. H. (1986). Steady-state evoked responses to sinusoidally amplitude-modulated sounds recorded in man. *Hearing Research*, *23*(2), 123–133.

Rosen, S. (1992). Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 336*(1278), 367-373.

Ross, B. (2013). The auditory steady state response. In: Celesia, G. (Ed.), Disorders of peripheral and central auditory processing. Elsevier, Amsterdam.

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. *J. Acoust. Soc. Am., 108*(2), 679-691.

Ross, B., Herdman, A.T. and Pantev, C. (2005). Right hemispheric laterality of human 40 Hz auditory steady-state responses. *Cerebral Cortex*, *15*, 2029–2039.

Ross B and Pantev C (2004). Auditory steady-state responses reveal amplitude modulation gap detection thresholds. *J Acoust Soc Am 115*, 2193-2206.

Sedley, W., Teki, S., Kumar, S., Overath, T., Barnes, G. R., and Griffiths, T. D. (2012). Gamma band pitch responses in human auditory cortex measured with magnetoencephalography. NeuroImage, 59(2), 1904–1911.

Schoonhoven R, Boden CJR, Verbunt JPA, de Munck JC (2003) A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. *Clinical Neurophysiology. 114,* 2096-2106.

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.

Tang H, Brock J, and Johnson BW (submitted for publication). Sound envelope processing in the developing brain: A MEG study.

Uhlhaas, P. J., Pipa, G., Neuenschwander, S., Wibral, M., and Singer, W. (2011). A new look at gamma? High- (>60 Hz) g-band activity in cortical networks: Function, mechanisms and impairment. *Progress in Biophysics and Molecular Biology*, *105*(1-2), 14–28.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am, 66*(5), 1364-1380.

Wang, X., Lu, T., Bendor, D., and Bartlett, E. (2008). Neural coding of temporal information in auditory thalamus and cortex. *Neurosci, 154*(1), 294-303.

Wang, X., Lu, T., and Liang, L. (2003). Cortical processing of temporal modulations. *Speech Commun, 41*(1), 107-121.

Warren, J., Wise, R., and Warren, J. (2005). Sounds do-able: auditory–motor transformations and the posterior temporal plane. *Trends in Neurosciences*, 8.

Whitfield, I. C. (1980). Auditory cortex and the pitch of complex tones. *J Acoust Soc Am, 67*(2), 644-647.

Chapter 3 – Evidence From EEG

Phase-locked and non-phase-locked envelope following response in the ascending auditory system: evidence from EEG

Huizhen Tang[a,b,1,2] Stephen Crain [a,b,3] Blake W. Johnson [a,b,2]

[a]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW 2109, Australia. http://www.ccd.edu.au, [b]Department of Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia. [c]Department of Linguistics, Macquarie University, Sydney, NSW 2109, Australia

[1] To whom correspondence may be addressed: joann.tang@mq.edu.au

[2] Present address: Department of Cognitive Science, Macquarie University, Sydney, NSW 2109

[3]Present address: Department of Linguistics, Macquarie University, Sydney, NSW 2109

3.1 Introduction

Temporal information delivered by the amplitude envelope of sounds is essential for speech perception and comprehension (Drullman, 1995, Shannon et al., 1995, Smith et al., 2002). Amplitude modulations (AMs) can be encoded in the auditory system through stimulus-neuron synchronization, which means that the spiking time of single neurons is aligned with the physical modulations of a sound (Joris et al., 2004, Eggermont, 2014). Such stimulus-neuron synchronisation is the most prominent encoding scheme for processing temporal information at lower levels of the auditory system but this capability declines progressively at higher levels (Joris et al., 2004, Eggermont, 2014). For example, the upper limit for phase synchronization in the auditory nerve is on the order of ~4 kHz, decreasing to ~256 Hz in the auditory brainstem (Joris et al., 2004, Wang et al., 2008, Eggermont, 2014). At the level of primary auditory cortex, the frequency range of precise phase synchronization in single neurons drops off sharply at rates above ~50 Hz (Giraud et al., 2000, Wang et al., 2003, Wang et al., 2008, Wang et al., 2012). Consistent with the single unit evidence from animal studies, EEG/MEG measurements show that the amplitude of the stimulus-synchonised auditory steady state response (ASSR) or envelope following response (EFR) drops off steeply at modulation rates greater than about 50 Hz (Picton et al., 2003, Purcell et al., 2004, Poulsen et al., 2007, Miyazaki et al., 2013).

Single unit studies in the marmoset monkey have identified two distinct classes of rate-sensitive neurons in primary auditory cortex: one that exhibits synchronization by phase-locking to low stimulus rates (below about 50 Hz) and another that shows non-synchronised increases in firing rate at higher stimulation rates (Lu et al., 2001, Liang et al., 2002, Wang et al., 2003). These results have been interpreted to suggest a dual encoding scheme for processing temporal modulations in sounds in the cortex (Lu et al., 2001, Liang et al., 2002). According to this dual encoding scheme, temporal modulation at low rates are represented “isomorphically” or “explicitly” by synchronised neuronal firing; whereas high rates are represented “non-

isomorphically" or "implicitly" with rate-based encoding that is not phase-locked to the stimulus periodicity (Lu et al., 2001, Liang et al., 2002, Wang et al., 2003). The results from the marmoset monkey (a new world monkey) have been confirmed in the macacque (an old world monkey) and with invasive electrophysiological recordings in auditory cortex of human patients (Steinschneider et al., 2008, Brugge et al., 2009), indicating that dual encoding is a general property of primate auditory cortex.

Motivated by these results from invasive recordings, we reasoned that conventional ASSR/EFR measured with noninvasive EEG/MEG recordings (Picton et al., 2003; Purcell et al., 2004; Poulsen et al., 2007; Miyazaki et al., 2013) may provide only one part of the story of how human auditory cortex responds to temporal modulations. These responses are typically analysed by first averaging across trials before conducting spectral or time-frequency analysis at a given frequency or frequency range. As such, previous reports have typically provided information only about phase-locked responses and have effectively discarded any information about non-phase locked responses. In Chapter 2 we characterised the total power change, phase-locked and non-phase locked power change, of MEG EFRs to a range of stimulation rates. Our results confirmed that there is substantially more total power than phase-locked power in EFRs to high modulation rates above about 50 Hz. These results are consistent with a dual encoding scheme and cannot be readily explained by a model that posits the auditory cortex acts as a simple low-pass filter for temporal information.

In the present chapter we analyze the EEG data collected concurrently with the MEG data described in Chapter 2. MEG and EEG are complementary techniques that measure the magnetic field outside the head and the electric potential on the scalp, respectively (Nakasatp et al., 1994, Baillet et al., 2001, Sharon et al., 2007, Hansen et al., 2010). MEG is based on the measurement of the magnetic field at the location of the sensors generated mainly by intracellular currents (also called primary currents) in cortical pyramidal cells (Baillet et al., 2001, Hansen et al., 2010). In contrast, EEG measures the electric potentials on the scalp

generated by the extracellular currents (also termed return currents) conducted through the volume of the head (Baillet et al., 2001, Buzsaki et al., 2012). To be measurable using EEG and MEG, the spatial arrangement of neural cell assemblies is crucial. It has been suggested that the main generators for EEG and MEG signal are the macrocolumns of tens of thousands of large pyramidal cortical neurons (Baillet et al., 2001).

EEG and MEG provide complementary but non-identical measures of brain activity. An important feature of MEG is that, because the magnitude of the magnetic field decreases drastically with distance (de Jongh et al., 2005, Hansen et al., 2010), the detection of deep sources is limited. The brain signals that are detectable using MEG are mainly produced in the cerebral cortex, due to its relative proximity to the MEG sensors (Hillebrand and Barnes, 2002, Leijten et al., 2003). In contrast, EEG detects volume currents induced by primary currents and so the distance of the source and electrode is not an formidable obstacle for the measurement of these brain responses (Jones and Byrne, 1998, Tonnquist-Uhlen et al., 2003, Nunez and Srinivasan, 2006).

A number of EEG studies have reported that both brainstem and cortical regions contribute to scalp recorded ASSRs or EFRs (Mauer and Döring, 1999, Herdman et al., 2002, Purcell et al., 2004). Mauer and Döring (1999) reported that both the brainstem and the temporal lobe were active during the ASSRs to noise that is amplitude modulated at rates between 24 and 120 Hz. Using dipole source analysis, they found that both brainstem and cortical regions (i.e. temporal lobe) were involved in generating the EFRs to modulation at low rates between 20 to 40 Hz. At higher rates between 70 to 110 Hz, the dominant responses were the brainstem following responses while the cortical responses decreased as the modulation rate increased. Similarly, an EEG study by Herdman and colleagues (2002) found that both the cortex and brainstem were simultaneously active and that the scalp recorded ASSRs were resulting from multiple generators with decreasing cortical participation as the frequency increased. Consistent with invasive recordings in cats (Kiren et al., 1994), Herdman and

colleagues (2002) reported that the brainstem was the primary source of the scalp recorded ASSRs at 88 Hz in humans. Using computational modeling on their experimental data, Purcell and colleagues (2004) investigated the multiple generators of the scalp recorded EFRs to amplitude-modulated at rates from 20 to 600 Hz. Model fitting of EFRs to rates between 20 to 100 Hz resulted in a model of the EFR consistent with two EFR generators in the brainstem and cortex. The brainstem EFR showed a constant amplitude of 35 nV across the entire modulation frequency range between 20 to 100 Hz. Meanwhile, the cortical EFR showed a constant amplitude of 85 nV at the frequencies between 20 to 50 Hz and decrease linearly to zero at 95 Hz.

On the other hand, the MEG-recorded ASSR/EFRs are reported to be primarily generated in auditory cortex with little if any contribution from the brainstem even at high AM rates (Ross et al., 2000). Consistent with these findings, Miyazaki and colleagues (2013) reported that only cortical sources were involved in generating the EFRs to two-beat complex sound with amplitude-modulation sweeping from 3 Hz to 60 Hz and back to 3 Hz. Similarly, MEG recorded EFRs to white noise amplitude-modulated with rates changing linearly from 10 Hz to 80 Hz exhibited cortical responses up to 80 Hz in both healthy adults and adults with dyslexia (Lehongre et al., 2011). Moreover, Schoonhoven and colleagues (2003) used MEG to record ASSRs to amplitude-modulated tones with modulation rate at around 40 Hz and around 80 Hz concentrating on estimating underlying dipole sources (Schoonhoven et al., 2003). Their dipole modeling showed clear bilateral activation of regions within or near the auditory cortex at both low and high modulation rates. They concluded that MEG-recorded ASSRs are dominated by activity in the bilateral auditory cortex at both the 40 Hz and 80 Hz frequency rates.

These seemingly contradictory results from EEG and MEG studies are actually expected from the different spatial sensitivities of the two techniques. Given that EEG and MEG provide complementary but non-redundant measures of brain functions, EEG was included in the

present study. The present chapter compares the EFRs measured by the two techniques, with two aims. First, since EEG and MEG both measure functioning of auditory cortex, we predicted that we should obtain similar a similar profile of PL and NPL cortical responses in the EEG as in the profile that was obtained using MEG. If this is correct, this will provide an independent validation of our Chapter 2 results, and generalise our MEG results to EEG, a technique that is much more widely available to auditory researchers and clinicians. The second aim was to exploit EEG's sensitivity to deep brain sources, in order to compare cortical and brainstem EFRs. As described above, the responses of subcortical auditory neurons are primarily phase-locked responses (Joris, 2004), and it has been well established that brainstem responses to modulation rates below about 250 Hz are strictly phase-locked (Eggermont, 2014). Accordingly, we predicted that EFRs from EEG cortical sources should consist of both PL responses (at low AM rates) and NPL responses (at higher AM rates) whereas in contrast, EFRs from EEG brainstem sources should comprised mainly or entirely of PL responses at all AM rates between 1-80 Hz.

3.2 Methods

3.2.1 Subjects

Subjects were sixteen adults (10 females) aged 22 to 36 years (M = 29.1, SD = 3.9), from the group of 28 subjects reported in Chapter 2. All subjects showed normal thresholds of hearing ($\leq$ 20 dB HL) for octave frequencies from 500 to 2000 Hz measured with an Otovation Amplitude T3 series audiometer (Otovation LLC, King of Prussia, PA). All procedures were approved by the Human Subjects Ethics Committee of Macquarie University.

3.2.2 Acoustic stimulation and procedure

The acoustic stimulus and procedures were the same as the first experiment described in **Chapter 2**. In short, the acoustic stimulus was an amplitude-modulated white noise. The

modulation depth was 100% with the rate swept exponentially from 1 Hz to 80 Hz in a 9-second duration (see Figure 1 in Chapter 2). A 300 ms segment of un-modulated white noise was presented right before and after the AM sweep followed by a silent inter-stimulus interval of about 950 ms.

3.2.3 EEG recording

Brain activity (reported in Chapter 2) was recorded continuously using a whole-head MEG system consisting of 160 axial gradiometers with a 50 mm baseline (Model PQ1160R-N2, KIT, Kanazawa, Japan) located in a magnetically shielded room (Fujihara Co. Ltd., Tokyo, Japan). Concurrently, brain activity was also recorded using a 64-electrode, MEG-compatible whole head EEG system (BrainProducts GmbH, Gilching, Germany) with a sampling rate of 1000 Hz. All measurements were carried out with participants in a supine position.

3.2.4 Structural MRI Scans

The subjects' structural MRI scans were obtained in a 3 Tesla Siemens Magnetom Verio scanner with a 12-channel head coil at the Macquarie University Hospital, Sydney. Anatomic images were processed using an MP-RAGE sequence (208 axial slices, TR = 2000 ms, TE = 3.94 s, FOV = 240 mm, voxel size = 0.9 mm3, TI = 900, flip angle = 9°).

3.2.5 Analyses.

BESA Research Version 6.0 (BESA Research GmbH: Grafelfing, Germany) was used to analyse both the MEG and EEG data. Structural MRIs were co-registered with the MEG and EEG data using BESA MRI 2.0 (BESA Research GmbH: Grafelfing, Germany) after being warped into Talairach space.

3.2.5.1 Source waveform reconstruction

Bilateral primary auditory cortex source model. To determine whether the dual temporal representations shown in our MEG data (reported in Chapter 2) could be seen in our EEG data, we applied the same bilateral dipole sources used for reconstructing source waveforms for

extracting EFRs in MEG data. The rationale underlying this procedure relies on the assumption that neurophysiological signals measured in EEG and MEG principally originate from the same pyramidal cell assembly (Baillet et al., 2001). Additionally, compared with EEG, source localisation based on single subject MEG data is usually more accurate than single subject EEG data with the same experimental setting due to a better spatial resolution of MEG. Having fixed the dipole location based on individual MEG data, the orientation of each dipole was re-adjusted based on the N100 of each subject's EEG data.

Cortical-subcortical 3-dipole source model. The bilateral dipoles were fixed in orientation and position, and we added a third dipole at the midline lower brainstem using the individual subjects anatomical MRI (Giraud et al., 2000, Herdman et al., 2002). The orientation of the brainstem dipole was fixed to the vertical-lateral plane (i.e. x=0, y=0.7, z=0.7) (Herdman et al., 2002).

3.2.5.2 Time-frequency analysis

The continuous source waveforms were first segmented with an epoch window from 600 ms pre-stimulus to 9600 ms post-stimulus. Time-frequency analyses were then carried out on the epoched source waveforms over a frequency range of 1 to 80 Hz using the complex demodulation method (Hoechstetter et al., 2004), with a frequency step of 1 Hz and a time step of 50 ms. Time-frequency representations of both amplitude and phase were generated for each waveform and further analyses were conducted on these time-frequency representations of amplitude or phase.

Inter-trial phase coherence (*ITPC*), which quantifies phase consistency cross trials for each frequency and time point, is usually computed from the single trial time-frequency representation using the following formula.

$$ITPC(t,f) = \frac{1}{N}\sum_{k=1}^{N} e^{i\varphi_k(t,f)}$$

where N = total number of trials; and $\varphi_k(f,t)$ = the phase in trial k.

Event-related spectral perturbation *(ERSP)*, which quantifies the increase or decrease in amplitude/power in a given frequency or frequency band with reference to the ongoing background brain activity recorded in the defined baseline, was computed. Specifically, the *ERSP* index is given by the following formula (Makeig, 1993, Pfurtscheller and Silva, 1999, Brugge et al., 2009).

$$ERSP(t,f) = \frac{\sum_{n=1}^{N}\left(\frac{A_n(t,f) - \overline{A_{baseline,\ n}(f)}}{\overline{A_{baseline,\ n}(f)}} \times 100\%\right)}{N}$$

where $A_n(t,f)$ is the absolute amplitude at time *t* and frequency *f* in trial *n*; and $\overline{A_{baseline,\ n}(f)}$ is the mean absolute amplitude at frequency *f* over the defined baseline in trial *n*.

The phase-locked component of ERSP (*PL)* was computed by applying spectral analysis on averaged waveform, which is obtained from averaging the MEG signals across trials, using the following formula:

$$PL(t,f) = \frac{\bar{A}(t,f) - \overline{\overline{A_{baseline}}(f)}}{\overline{\overline{A_{baseline}}(f)}} \times 100\%$$

where $\bar{A}(t,f)$ is the absolute amplitude at time *t* and frequency *f* of the averaged signal; and $\overline{\overline{A_{baseline}}(f)}$ is the mean absolute amplitude of the averaged signal at frequency *f* over the defined baseline.

The non-phase-locked component of ERSP (*NPL*) was then computed by subtracting the *PL* from the amplitude change in single trial as indicated by the following formula:

$$NPL(t,f) = \frac{\sum_{n=1}^{N}\left(\frac{A_n(t,f) - \overline{A_{baseline,\ n}(f)}}{\overline{A_{baseline,\ n}(f)}} \times 100\% - PL(t,f)\right)}{N}$$

3.2.5.3 Statistical analysis

Permutation tests were carried out on the time-frequency data based on previously developed algorithms (Maris and Oostenveld, 2007) using custom MATLAB scripts to examine whether

the envelope following pattern was statistically significant. The permutation tests were applied between 0.3 second after the stimulus onset to 9.3 second, during which time the temporal modulation of the stimulus swept from 1 Hz to 80 Hz. During permutation, a primary threshold was used to identify the top 5% of values. The clustering algorithm used the sum of activities within a cluster (instead of the size of a cluster). Only clusters surviving a permutation with 1000 iterations at the significant level of 0.05 were accepted. To better visualize the relationship between stimulus modulation rate and brain response, a linearization procedure was applied to convert the time axis to modulation rate, based on the stimulus spectrogram function $f(t) = 80^{t/9}$. Model fitting was then performed in MatLab R2014a (the MathWorks Inc., Natick, MA, USA) using the Curve Fitting Toolbox (version 3.4.1) to model the correlation between modulation rate and frequency of brain response. A linear model ($f = a \times m + b$; here f is the frequency of brain response, m is the rate of modulation, a and b are two parameters set for free fitting) was used to fit the permuted data. The EFR to AM at rate *m*, was identified as the mean magnitude within the bin [*f-1 f+1*]. The vector strength of the defined EFR was then re-plotted against the frequency of the AM to visualize the TMTFs. A 3 Hz smoothing window was applied to TMTF plots for visualization only.

For statistical comparison of EEG and MEG data, spectrograms were first normalised within each frequency of brain response, and z-scores were calculated. Then the absolute value of the difference between the normalised spectrogram data of EEG and MEG was computed to form a new spectrogram. The same permutation test and statistical procedures were performed on this new spectrogram to determine if the difference was statistically significant. Cortical responses were averaged across two hemispheres to maximise the signal-to-noise ratios of the EEG and MEG data.

3.3 Results

3.3.1 EFRs in the primary auditory cortex

3.3.1.1 EFRs quantified with ITPC and ERSP

Figure 1 (A) and (B) shows the EEG and MEG responses quantified with ITPC and ERSP. Clear tracking patterns during the presentation of the AM sweep was observed in both the EEG and MEG data, especially with an ITPC measure. From Figure 1 (C) and (D), we can see a close resemblance between the TMTF profiles extracted from the EEG data and the MEG data.

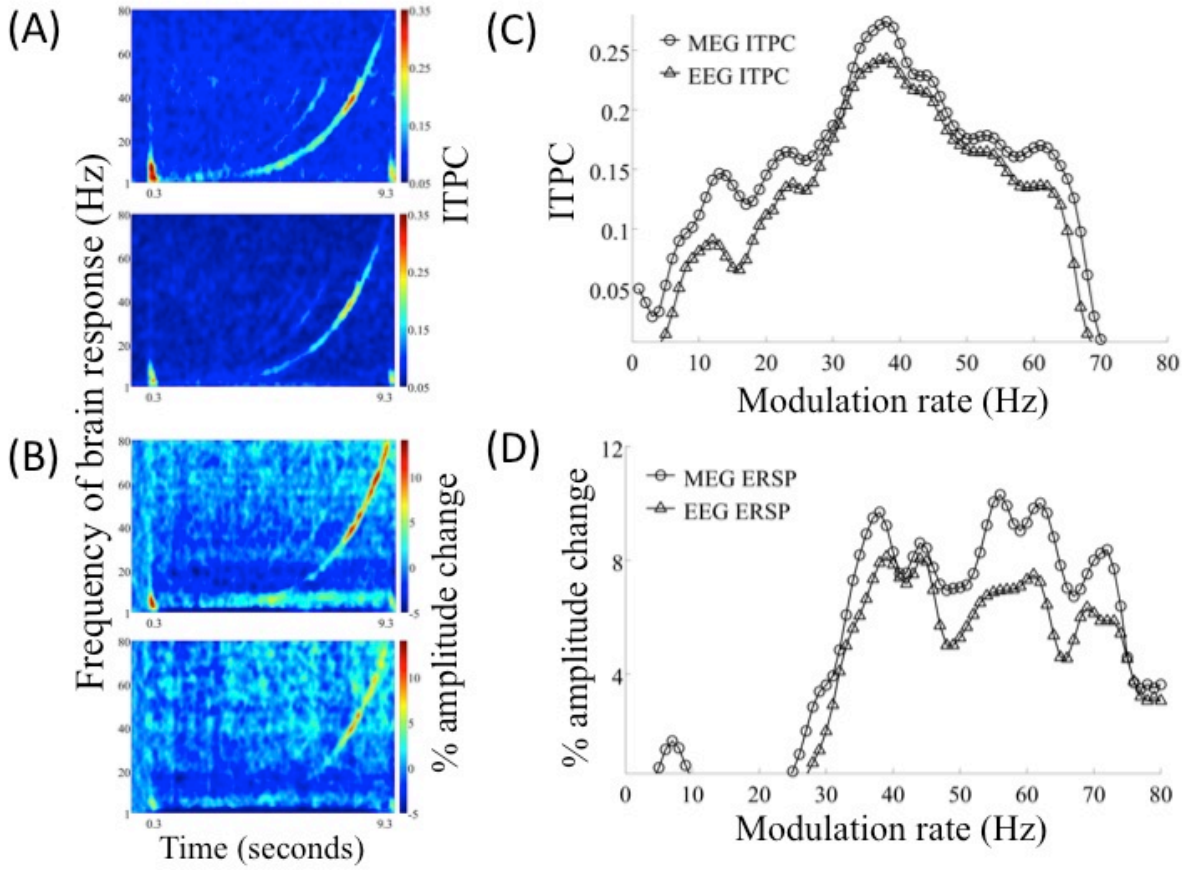

Figure 1. EEG and MEG responses quantified with ITPC and ERSP during stimulation. (A) Top and bottom panels show the ITPC for MEG and EEG, respectively. (B) Top and bottom panel show ERSP MEG and EEG, respectively. (C) TMTFs of ITPC recorded with EEG and MEG. (D) TMTFs of ERSP recorded with EEG and MEG.

3.3.1.2 EFRs quantified with PL- and NPL-ERSP

Figure 2 shows the spectrogram data of phase-locked and non phase-locked event-related spectral perturbation (ERSP in the primary auditory cortex recorded in both EEG and MEG.

From Figure 2, we can clearly see two distinct components of following response with the PL-ERSP being observable almost through the whole range of frequency tested and peaked around 40 Hz and the NPL-ERSP did not appear around 40 Hz and remained high in amplitude on frequencies above 40 Hz. These properties were also evident in the TMTF shown in Figure 3, which shows EFR profile plotted against the frequency of AMs. Clearly, the PL-ERSP and NPL-ERSP components demonstrated in our MEG data were also evident in the EEG data. The dominant frequency ranges for each of these two representations recorded with EEG were similar to the ones with MEG.

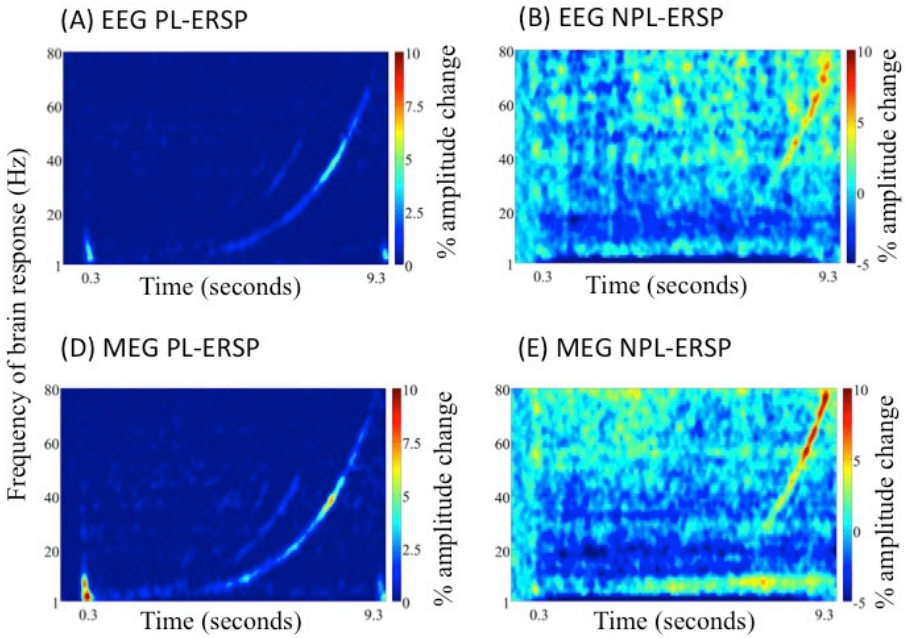

Figure 2. Phase-locked (PL) and non-phase-locked (NPL) components of EEG and MEG responses. (A) and (B) show PL and NPL components, repectively, of brain responses recorded with EEG. (C) and (D) show PL and NPL component, repectively, of brain responses recorded with MEG.

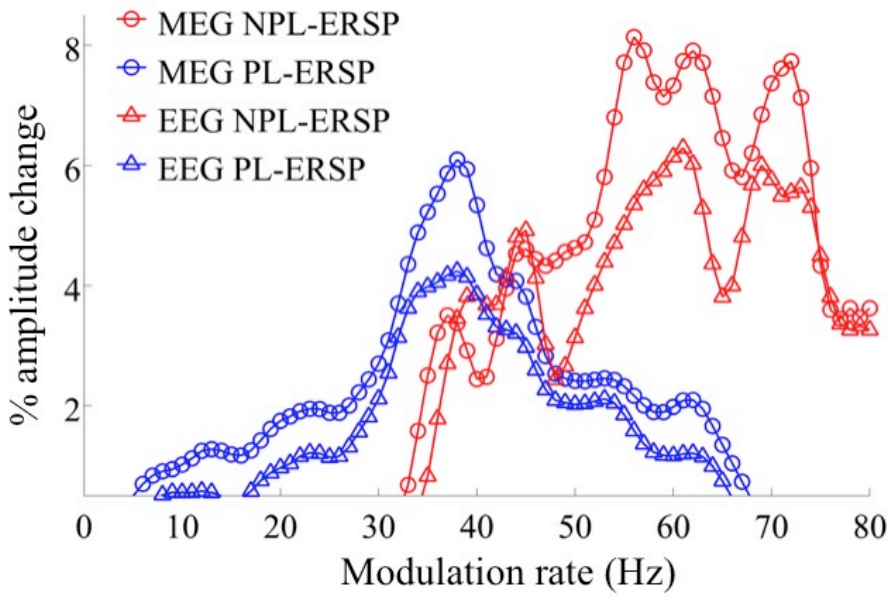

Figure 3. EEG and MEG recorded envelope following responses demonstrated with TMTFs.

3.3.2 EFRs in the cortical and subcortical sources

3.3.2.1 EFRs quantified with ITPC and ERSP

Figure 4 and 5 show the spectrograms of brain response measured with ITPC and ERSP when a third dipole was added to the bilateral dipole model when analysing the EEG and MEG data. Figure 6 presents the EEG and MEG recorded cortical and subcortical TMTF profiles of EFRs. From Figure 6(A), we can see both the cortical and subcortical ITPCs showed similarity in both the EEG and MEG data. However, when permutation tests (alpha = .05, see method for details) indicated MEG recorded EFRs measured with ITPC were larger than that recorded with EEG at frequencies between 3 Hz to 15 Hz and between 40 to 66 Hz. However, the permutation tests indicated that the EEG recorded subcortical EFRs were larger than the one recorded with MEG from 28 Hz to 40 Hz with a reversed pattern between 51 Hz to 58 Hz. Comparing cortical and subcortical responses shown in Figure 6(A), when the cortical ITPC showed a typical band-pass filter TMTF profile with a peak around 40 Hz, the subcortical ITPC, especially the one recorded in EEG showed a flat, sustained response between about 30 to 75 Hz. Permutation

tests indicated that MEG recorded cortical ITPC was larger than the subcortical ITPC at frequencies between 1 - 60 Hz and smaller between 66 Hz to 73 Hz. Meanwhile, similar cortical and subcortical differences were observed in the EEG data, cortical ITPC was larger between 9 Hz and 58 Hz and smaller between 66 Hz and 75 Hz.

Figure 6(B) shows that the ERSP recorded with EEG has a similar pattern to the MEG ERSP. Although the mean ERSP was lower in the EEG, permutation tests indicated the differences were not statistically significant, probably due to larger variance across trials as well as across subjects in the EEG. A significant subcortical ERSP was only observed in the MEG data at frequencies between 42 Hz to 66 Hz. The MEG recorded cortical ERSP was significantly larger than the brainstem ERSP.

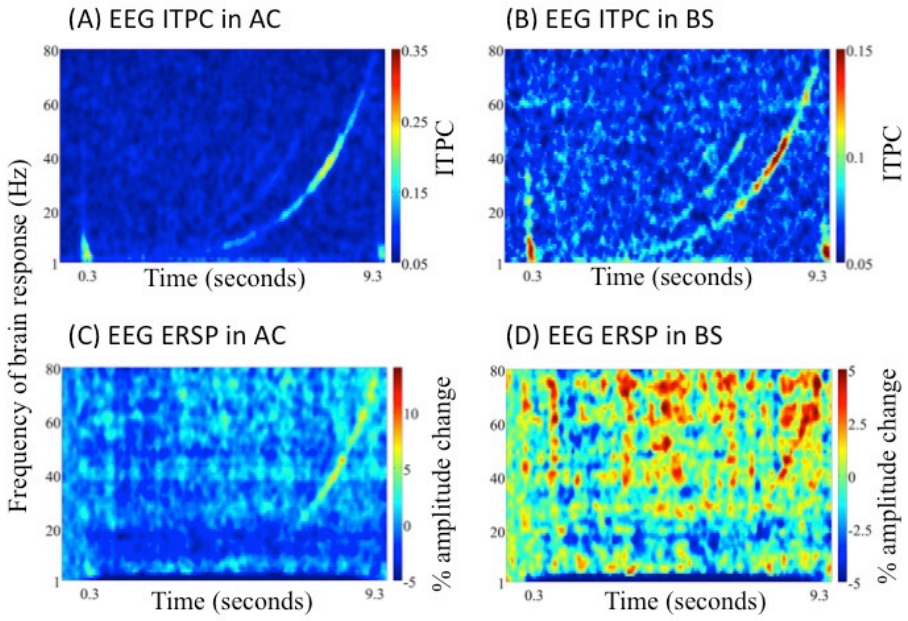

Figure 4. EEG recorded brain response in the auditory cortex (AC) and brainstem (BS). (A) and (B) shows the EEG recorded brain response quantified with ITPC in the AC and BS, respectively. (C) and (D) presents the EEG recorded brain response quantified with ERSP in the AC and BS, respectively.

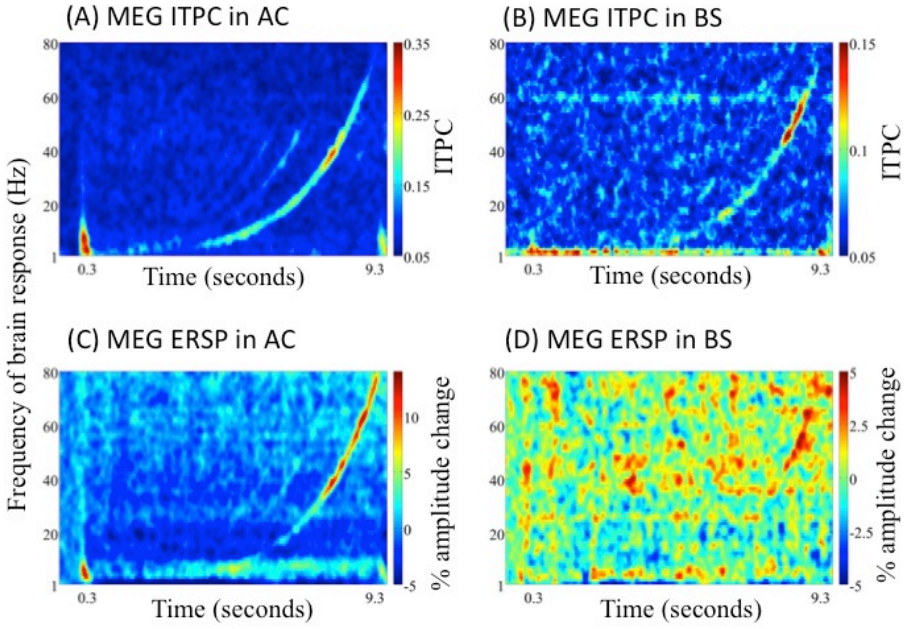

Figure 5. MEG recorded brain response in the auditory cortex (AC) and brainstem (BS). (A) and (B) shows the MEG recorded brain response quantified with ITPC in the AC and BS, respectively. (C) and (D) presents the MEG recorded brain response quantified with ERSP in the AC and BS, respectively.

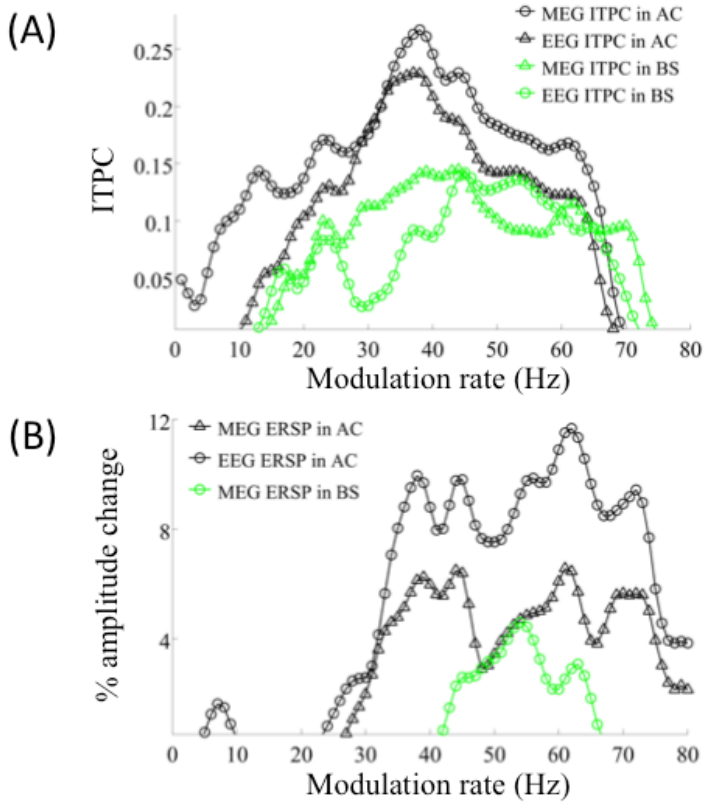

Figure 6. EEG and MEG recorded cortical and subcortical envelope following responses demonstrated with TMTFs. (A) TMTFs of ITPC in auditory cortex (AC) and brainstem (BS). (B) TMTFs of ERSP in AC and BS.

3.3.2.2 EFRs quantified with PL- and NPL-ERSP

Figure 7 and 8 show the cortical and subcortical PL and NPL components of the envelope following response. While both the PL-ERSP and NPL-ERSP were evident in the bilateral cortex, our analyses showed a significant PL response but no significant NPL response in the brainstem.

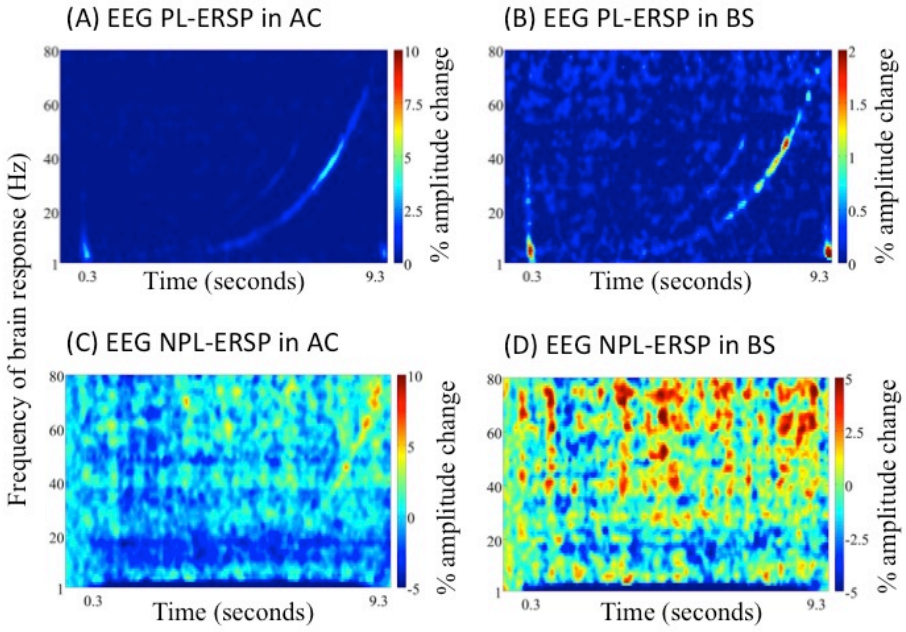

Figure 7. Cortical and subcortical PL and NPL component of brain response recorded with EEG. (A) and (B) presents the EEG recorded PL component generated in the auditory cortex (AC) and the brainstem (BS), repectively. (C) and (D) presents the EEG recorded PL component generated in the auditory cortex (AC) and the brainstem (BS), repectively. For better visualisation, the scaling for cortical and subcortical spectrogram data is different.

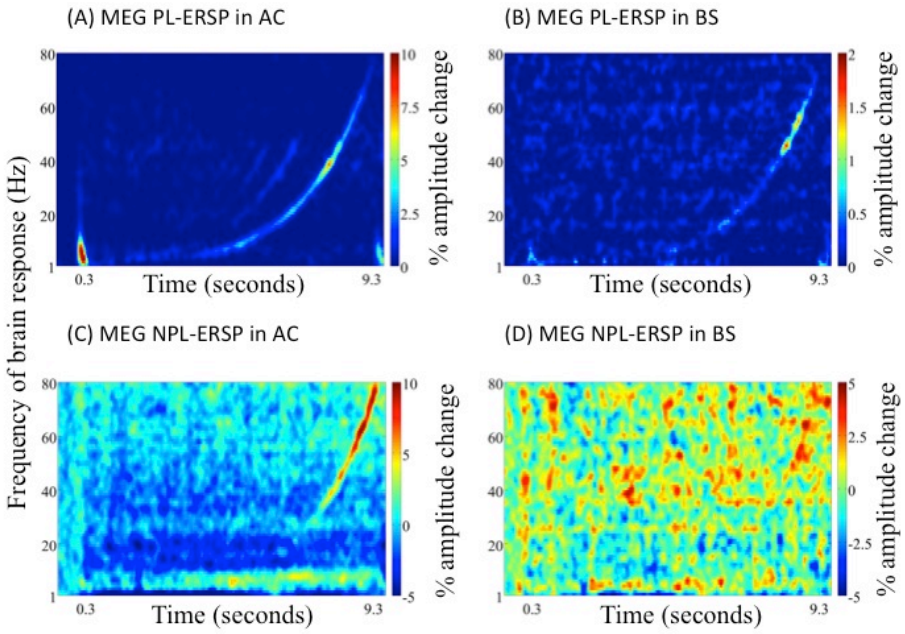

Figure 8. Cortical and subcortical PL and NPL component of brain response recorded with MEG. (A) shows the PL component in the auditory cortex (AC); (B) shows the PL component in the brainstem (BS); (C) shows the PL component in the auditory cortex (AC); (D) shows the PL component in the brainstem (BS). For better visualisation, the scaling for cortical and subcortical spectrogram data is different.

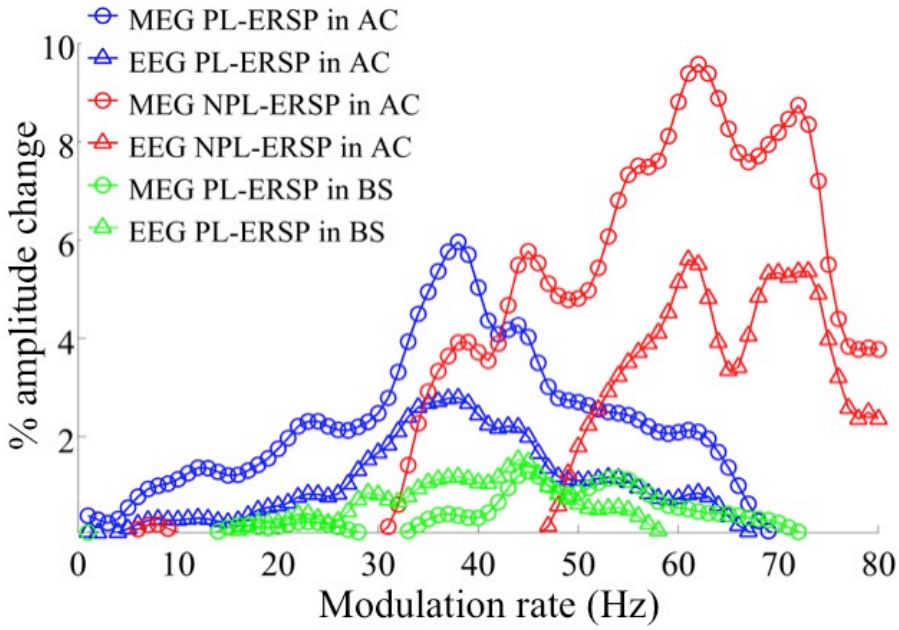

Figure 9. TMTFs of PL and NPL component of cortical and subcortical envelope following response recorded with EEG and MEG. AC stand for auditory cortex and BS stands for brainstem. No significant NPL following responses were observed in the brainstem source and therefore not plotted in this figure.

3.3.2.3 Effect of source modelling

Figure 10 shows the PL and NPL components of the cortical EFRs recorded in EEG and MEG. The MEG recorded cortical EFRs were not affected whether or not a third dipole in the brainstem was added to the source modeling for reconstructing the source waveforms, when the brainstem dipole resulted reduction of the EEG recorded PL following response from 36 Hz to 54 Hz and NPL following response between about 35 Hz and about 46 Hz.

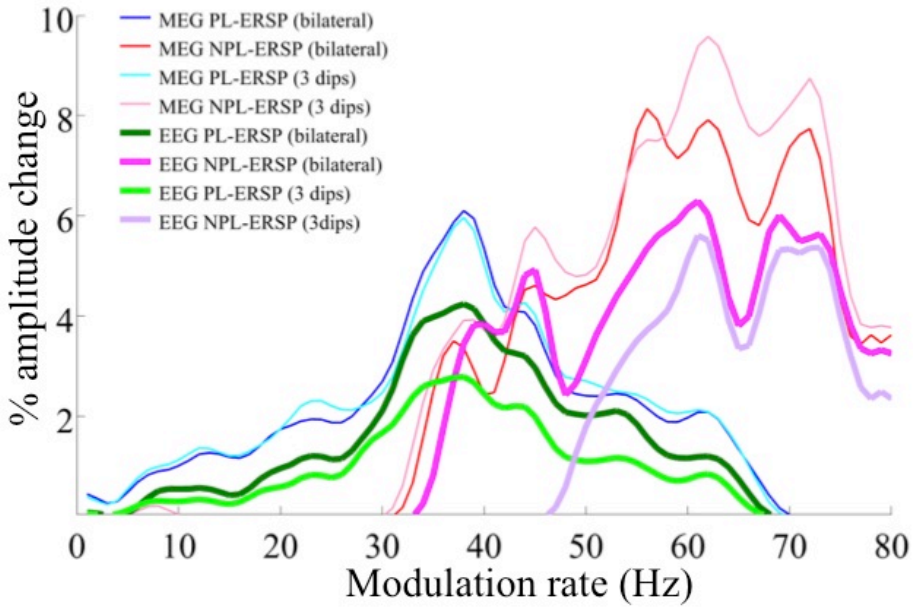

Figure 10. TMTFs of envelope following responses using different source models: '3 dips' stands for source model with bilaterally placed dipoles in the primary auditory cortex plus a dipole placed in the brainstem. 'bilateral' stands for source model with only bilaterally placed dipoles in the primary auditory cortex. The primary auditory cortical locations were identified with dipole fitting to the M100 response for each subject based on their MEG data.

3.4 Discussion

The present study was designed to investigate whether the PL and NPL responses observed in our MEG-recorded EFRs (Chapter 2) are also present in the concurrently-recorded EEG

responses; and to compare PL and NPL responses in auditory cortex to brainstem responses that are (in principle) accessible in the EEG measurements but not in the MEG measurements.

3.4.1 Phase-locked and non phase-locked components of the EEG-recorded envelope following response

The EEG EFRs modelled with 2 bilateral cortical dipoles (Figures 1-3) strongly resemble the MEG responses described in Chapter 2. When ITPC and total power of the ERSP were compared in the EEG and MEG, the overall TMTF profiles showed nearly the same pattern, with EEG showing only slightly smaller mean magnitudes of both measures, differences that were not statistically significant. Therefore it is clear that the EEG EFR shows the same PL and NPL components as in the EEG. These results provide confirmation that the MEG results of Chapter 2 cannot be attributed to methodological or analytical issues unique to the MEG, and also extend the generality of the MEG results to the EEG, an electrophysiological technique that is much more widely available to researchers and clinicians than the MEG. Given that the EEG and MEG are measuring complementary aspects of neuronal signals, the close resemblance between the EFRs responses recorded with MEG and EEG provides strong validation for a two-component view of the EFR and further support for the dual temporal encoding hypothesis.

Previous electrophysiological studies in humans have consistently demonstrated a TMTF profile that the cortical ASSR/EFR peaks around 40 Hz and declines steeply at frequencies higher than about 50 Hz (Ross et al., 2000, Picton et al., 2003, Purcell et al., 2004, Miyazaki et al., 2013). The extraction of ASSRs/EFRs in these studies were based on averaging methods in which the non-phase-locked activity is minimized through phase cancellation whereas the phase-locked responses survived as a result of their phase concentration (Makeig, 1993, Pfurtscheller and Silva, 1999, Picton et al., 2003). This is an effective method for increasing the signal to noise ratio of strongly phase-locked signals, but effectively discards any information about non phase-locked responses. Accordingly, the ITPC metric used in the

present study, which quantifies the degree of phase-locking to the acoustic periodicity, results in a TMTF that peaks at stimulus rates of about 40 Hz and falls off steeply at rates higher than about 50 Hz. This profile closely replicates the phase-locked properties of ASSRs/EFRs reported in the previous literature. However the ERSP power metric, which quantifies the total power of both phase-locked and non-phase locked responses, results in a much different profile with maintained power at rates above 50 Hz. This ERSP response profile indicates that strongly phase-locked responses (Joris et al., 2004, Picton, 2013) provide only a partial and incomplete picture of how the auditory cortex responds to AM sounds. The ERSP temporal pattern cannot be accommodated within our current understanding of cortical temporal processing in humans (e.g. Picton, 2013; Ross, 2001). The discrepancy between the two measures, ITPC and ERSP, indicates that in addition to the phase-locked responses, the non-phase-locked activity might be also involved in processing the temporal modulations. Specifically, the ERSP profile showed sustained EFRs at frequencies between about 30 Hz up to about 80 Hz, which clearly does not fit a view of reduced of auditory responsiveness at higher frequencies at the cortical level in the ascending auditory system

Our results do accord well with recent results obtained with invasive electrophysiological measurements of ERSPs in primary auditory cortex elicited by click trains in monkeys and humans. An intracranial EEG study in the awake monkey by Steinschneider and colleagues (Steinschneider et al., 2008) found that power increases at low click rates was primarily phase-locked to the acoustic stimuli whereas power increases at higher frequencies were primarily non-phase-locked. This transition from predominantly phase-locked activity to predominantly non-phase-locked activity during auditory temporal processing was more directly demonstrated in ECoG study on epilepsy patients undergoing surgical examination. Brugge and colleagues (Brugge et al., 2009) showed that the periodicity of click trains elicited only phase-locked FFRs for click rates below 50 Hz. For higher rates at 100 and 125 Hz, in addition to the phase-locked FFRs a non-phase-locked broadband power increase was also

observed. For rates above 150 Hz, the phase-locked FFR was strongly attenuated and responses were dominated by a non-phase-locked broadband power increase. These results, taken together with the single unit results of Lu et al. (2001) and our own results with noninvasive EEG and MEG, all lead to the conclusion that the primate primary auditory cortex responds in (at least) two distinctly different ways to low and high temporal modulations in sounds.

3.4.2 Phase-locked and non phase-locked responses in the brainstem

In addition to establishing commonalities between MEG and EEG measurements, the present EEG study was also motivated by the possibility of exploiting the differential sensitivities of the two techniques. In particular, we drew on theoretical considerations that EEG measurements are sensitive to deep brain sources in the auditory brainstem (Jones and Byrne, 1998, Tonnquist-Uhlen et al., 2003, Nunez and Srinivasan, 2006), and on empirical studies that have reported EEG measurements of the EFR have strong contributions from the brainstem at higher temporal modulation rates (Mauer and Döring, 1999, Herdman et al., 2002, Purcell et al., 2004). If we could detect brainstem responses with the EEG, we reasoned that we could establish clear differences between responses at low levels of the auditory system and the cortical responses described in the preceding section.

Since the responses of subcortical auditory neurons are primarily phase-locked responses (Joris, 2004), and brainstem responses to modulation rates below about 250 Hz are strictly phase-locked responses (Eggermont, 2014), we predicted that, unlike the cortical sources, brainstem sources should show only PL responses over the 1-80 Hz range of modulation rates used in our study. As predicted, a dipole source in the brainstem showed clear EEG PL responses over a wide range of stimulation rates but showed no significant NPL activity at any stimulus rate.

This dissociation between cortical and subcortical responsiveness to AM is in good agreement with results from single units at different levels of the ascending auditory pathway (Joris et al., 2004, Wang, 2007, Wang et al., 2008). Stimulus-neuron synchronized temporal

coding is prominent at lower levels of the auditory system but declines progressively at higher levels, with a transformation from predominantly synchronised responding to a predominantly nonsynchonised responding with increased in firing rate. This implies that in low frequencies (i.e. <~50 Hz) both subcortical and cortical regions are involved in processing AMs but only subcortical structures have the temporal resolution to encode AMs in higher rates in a phase-locked manner (Schreiner and Urbas, 1986, Eggermont and Wang, 2011, Eggermont, 2014). Kuwada et al. (2002) suggested that the scalp-recorded amplitude modulation following response is a composite response from multiple brain generators, with the responses to high and low frequencies reflecting strong contributions from subcortical and cortical sources, respectively. More comprehensively, Giraud and colleagues investigated the cortical representation of AM frequencies using fMRI (Giraud et al., 2000). They found that multiple cortical and subcortical regions, including right lower brainstem and Heschl's gyrus, were specifically responsive to AM.

A surprising result of the current analyses was the observation of significant PL EFR (albeit smaller in magnitude and over a smaller range of modulation rates) in the MEG brainstem source. Because the MEG has a steeper dependence on sensor to source distance (de Jongh et al., 2005, Hansen et al., 2010) we expected there should be little or no significant response of any kind in the MEG brainstem source. One explanation for this result is that we have mismodeled the actual physiological situation: i.e. the addition of a brainstem source in the model where there is none in reality, resulting in the appropriation of some of the cortical activity by the third source. However midmodeling does not explain why only the PL and none of the NPL activity is misattributed from the cortical sources to the subcortical source. This leaves the possibility that the MEG does in fact have some sensitivity to deep auditory brainstem sources and in fact this has been demonstrated in a number of previous reports (Erné and Hoke, 1990, Iramina and Ueno, 1995, Lütkenhöner et al., 2000, Parkkonen et al., 2009). It is not possible to adjudicate between these two explanations on the present evidence and this

issue merits further exploration in future studies. However whichever explanation holds these results provide good evidence that the PL and NPL responses have different spatial sensitivity profiles, because PL responses are detectable with both superificial and deep source models, while NPL responses are detectable only with superficial sources. This pattern is entirely consistent with the hypothesis that NPL responses are of cortical origin while PL responses are of cortical and subcortical origin.

3.4.3 Limitations of the present study

Our MEG results showed that there were two EFRs – NL and NPL. However, an important consideration is whether these results may be a consequence of energy leaking from evoked (PL) response to induced (NPL) response (David et al., 2006). In other words, the reported NPL following response in our MEG data could be a result of misrepresentation due to the spectral analysis within a short-time window. As discussed in Chapter 2, this issue is intrinsic to the time-frequency analysis for extracting the EFRs in electrophysiological recorded brain signals (David et al., 2006).

Using simulated data, David and colleagues (David et al., 2006) demonstrated that the subtraction-based method to separate evoked (PL) and induced (NPL) power could be problematic due to the effect of latency jitter. Normally, the latency jitter is small and negligible when only low frequencies are considered. However, small latency jitter can cause evoked responses appear as induced power, particularly at higher frequencies. The un-detected "evoked" energy would leak into the "induced" power spectrogram if a simple subtraction method was used to compute the induce response. Moreover, this energy leaking would become larger for higher frequencies given the same latency jitter. Since our NPL ERSP showed an increase of magnitude as the frequencies increased from about 30 Hz to 80 Hz, and the PL and NPL components are negatively correlated, as would be predicted from an "energy leaking" explanation. Therefore, it is important to question here whether the NPL following responses can be partially or even completely attribute to the energy leaking from PL power spectrogram

to the NPL power spectrogram. This question is technically challenging due to the low precision in estimating the latency for each frequency in an AM sweep paradigm. Therefore, it is difficult to precisely quantify the amount of energy leaking at each frequency caused by the current time-frequency analysis algorithm.

That being said, the absence of NPL in the brainstem when PL was clearly observed, at least, suggests that the energy leaking alone is unlikely to explain the NPL profile found in the auditory cortex. More importantly, if the NPL can be attributed to the energy leaking from PL, that means the actual (detected + undetected) phase-locked following responses would have a profile similar to the total ERSP, which shows a sustained following responses through the higher frequencies tested (40 through to 80 Hz). Such phase-locked following responses at high rates are unlikely according to converging evidence from electrophysiological studies in humans (Ross et al., 2000, Picton et al., 2003, Tlumak et al., 2011, Nodarse et al., 2012, Miyazaki et al., 2013) as well as single unit recordings in animals (Lu et al., 2001, Joris et al., 2004).

We emphasize that this mis-estimation problem *does not* apply to the total ERSP power computations used here (Steinschneider et al., 1998; Brugge et al., 2013). Therefore, *regardless* of whether our NPL ERSP results are attributable to energy leaking, or to the dual contributions of PL and NPL responses, they do demonstrate that our current understanding of the auditory ASSR/EFR is incomplete because it is based on analytic methods that emphasize phase-locking and essentially discard brain activity that is not tightly phase-locked to the periodicity of the acoustic stimulus.

3.5 Conclusions

The present EEG data confirmed the findings of both phase-locked (PL) and non phase-locked (NPL) components of EFRs in the MEG data reported in Chapter 2. This further supported the view that our current picture of the auditory EFR measured by EEG/MEG (Picton, 2003; Ross

et al., 2014) is incomplete and two distinct schemes are used to encode slow and rapid temporal modulations in sounds in the auditory cortex. Furthermore, the dissociation of PL and NPL in the cortical and subcortical structures in the auditory system was entirely consistent with the proposed transformation of encoding strategy along the auditory pathway (Joris et al., 2004, Wang et al., 2008).

Acknowledgements. We gratefully acknowledge the role of the Kanazawa Institute of Technology in establishing the KIT-Macquarie Brain Research Laboratory. This work was supported by the HearingCRC and Australian Research Council Grants CE110001021 and DP1096160

3.6 References

Baillet S, Mosher JC, Leahy RM (2001) Electromagnetic brain mapping. IEEE Signal Process Mag 18:14-30.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol 102:2358-2374.

Buzsaki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407-420.

David O, Kilner JM, Friston KJ (2006) Mechanisms of evoked and induced responses in MEG/EEG. Neuroimage 31:1580-1591.

de Jongh A, de Munck JC, Gonçalves SI, Ossenblok P (2005) Differences in MEG/EEG epileptic spike yields explained by regional differences in signal-to-noise ratios. J Clin Neurophysiol 22:153-158.

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. J Acoust Soc Am 97:585-592.

Eggermont JJ (2014) Animal models of auditory temporal processing. International journal of psychophysiology : official journal of the International Organization of Psychophysiology 95:202-215.

Eggermont JJ, Wang X (2011) Temporal coding in auditory cortex In: The Auditory Cortex (Winer, J. A. and Schreiner, C. E., eds), pp 309-328: Springer US.

Erné SN, Hoke M (1990) Short-latency evoked magnetic fields from the human auditory brainstem. Advances in neurology, 54, 167 54.

Giraud A-L, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A (2000) Representation of the temporal envelope of sounds in the human brain. J Neurophysiol 84:1588-1598.

Hansen PC, Kringelbach ML, Salmelin R (2010) MEG: An Introduction to Methods. New York: Oxford University Press

Herdman AT, Lins O, Van Roon P, Stapells DR, Scherg M, Picton TW (2002) Intracerebral sources of human auditory steady-state responses. Brain Topogr 15:69-86.

Hillebrand A, Barnes GR (2002) A quantitative assessment of the sensitivity of whole-head MEG to activity in the adult human cortex. NeuroImage 16:638-650.

Hoechstetter K, Bornfleth H, Weckesser D, Ille N, Berg P, Scherg M (2004) BESA source coherence: a new method to study cortical oscillatory coupling. Brain Topogr 16:233-238.

Iramina K, Ueno S (1995) Measurement of brainstem auditory evoked magnetic fields using a highly sensitive SQUID magnetometer with a variable baseline. Magnetics, IEEE Transactions on 31:4271-4273.

Jones SJ, Byrne C (1998) The AEP T-complex to synthesised musical tones: left-right asymmetry in

relation to handedness and hemisphere dominance. Electroencephalogr Clin Neurophysiol 108:355-360.

Joris P, Schreiner C, Rees A (2004) Neural processing of amplitude-modulated sounds. Physiol Rev 84:541-577.

Lehongre K, Ramus F, Villiermet N, Schwartz D, Giraud AL (2011) Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. Neuron 72:1080-1090.

Leijten FSS, Huiskamp G-JM, Hilgersom I, van Huffelen AC (2003) High-resolution Source Imaging in Mesiotemporal Lobe Epilepsy: A Comparison Between MEG and Simultaneous EEG. J Clin Neurophysiol 20:227-238.

Liang L, Lu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol 87:2237-2261.

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat Neurosci 4:1131-1138.

Lütkenhöner B, Lammertmann C, Ross B, Pantev C (2000) Brain stem auditory evoked fields in response to clicks. NeuroReport 11:913-918.

Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. Electroencephalogr Clin Neurophysiol 86:283-293.

Mauer G, Döring WH (1999) Generators of amplitude modulation following response (AMFR). In: XVI Biennial Meeting of the International Evoked Response Audiometry Study Group Trømsø, Norway.

Miyazaki T, Thompson J, Fujioka T, Ross B (2013) Sound envelope encoding in the auditory cortex revealed by neuromagnetic responses in the theta to gamma frequency bands. Brain Res 1506:64-75.

Nakasatp N, Levesque MF, Barth DS, Baumgartner C, Rogers RL, Sutherling WW (1994) Comparisons of MEG, EEG, and ECoG source localization in neocortical partial epilepsy in humans. Electroencephalogr Clin Neurophysiol 91:171-178.

Nodarse EM, Abalo MaCPr, Fortuny AT, Hernández MV, Castellanos AL (2012) Maturational Changes in the Human Envelope-following Responses. Acta Otorrinolaringol Esp 63:258-264.

Nunez PL, Srinivasan R (2006) Electric fields of the brain: the neurophysics of EEG: Oxford university press.

Parkkonen L, Fujiki N, Makela JP (2009) Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. Hum Brain Mapp 30:1772-1782.

Pfurtscheller G, Silva FHLd (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110:1842-1857.

Picton T (2013) Hearing in Time: evoked Potential Studies of Temporal Processing. Ear Hear 34:385-401.

Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. Int J Audiol 42:177-219.

Poulsen C, Picton TW, Paus T (2007) Age-related changes in transient and oscillatory brain responses to auditory stimulation in healthy adults 19-45 years old. Cerebl Cortex 17:1454-1467.

Purcell DW, John SM, Schneider BA, Picton TW (2004) Human temporal auditory acuity as assessed by envelope following responses. J Acoust Soc Am 116:3581-3593.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones J Acoust Soc Am 108:679-691.

Schoonhoven R, Boden CJR, Verbunt JPA, de Munck JC (2003) A whole head MEG study of the amplitude-modulation-following response: phase coherence, group delay and dipole source analysis. Clin Neurophysiol 114:2096-2106.

Schreiner CE, Urbas JV (1986) Representation of amplitude modulation in the auditory cortex of the cat. I. The anterior auditory field (AAF). Hear Res 21:227-241.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303-304.

Sharon D, Hämäläinen MS, Tootell RBH, Halgren E, Belliveau JW (2007) The advantage of combining MEG and EEG: Comparison to fMRI in focally stimulated visual cortex. NeuroImage 36:1225-1235.

Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. Nature 416.

Steinschneider M, Fishman YI, Arezzo JC (2008) Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. Cereb Cortex 18:610-625.

Tlumak AI, Durrant JD, Delgado RE, Boston JR (2011) Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: profile in adults. Int J Audiol 50:448-458.

Tonnquist-Uhlen I, Ponton CW, Eggermont JJ, Kwong B, Don M (2003) Maturation of human central auditory system activity: the T-complex. Clinical neurophysiology 114:685-701.

Wang X (2007) Neural coding strategies in auditory cortex. Hear Res 229:81-93.

Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. Neurosci 154:294-303.

Wang X, Lu T, Liang L (2003) Cortical processing of temporal modulations. Speech Commun 41:107-121.

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ (2012) Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. J Neurophysiol 107:2033-2041

Chapter 4 – Sound Envelope Processing in the Developing Brains

The study presented in this chapter has been fully accepted in *Clinical Neurophysiology* on 16/07/2015:

Tang, H., Brock, J. and Johnson W. B. (in press). Sound envelope processing in the developing human brain : A MEG study. *Clinical Neurophysiology*.

Sound envelope processing in the developing human brain: A MEG study

Huizhen Tang[a,b,1,2] **Jon Brock** [a,b,2] **Blake W. Johnson** [a,b,2]

[a]ARC Centre of Excellence in Cognition and its Disorders, Macquarie University, Sydney, NSW 2109, Australia. http://www.ccd.edu.au;

[b]Department of Cognitive Science, Macquarie University, Sydney, NSW 2109, Australia.

[1] To whom correspondence may be addressed:

Email:joann.tang@mq.edu.au; Tel: +61298506789; Fax: +61298506059

[2] Present address: Department of Cognitive Science, Macquarie University, Sydney, NSW 2109

Acknowledgements: This work was supported by the HearingCRC and Australian Research Council Grants CE110001021 and DP1096160. We thank Stephen Crain for helpful discussions and comments about this manuscript.

* None of the authors have potential conflicts of interest to be disclosed

Abstract

Objective: This study investigated auditory cortical processing of linguistically-relevant temporal modulations in the developing brains of young children.

Methods: Auditory envelope following responses to white noise amplitude modulated at rates of 1 - 80 Hz in healthy children (aged 3-5 years) and adults were recorded using a paediatric magnetoencephalography (MEG) system and a conventional MEG system, respectively.

Results: For children, there were envelope following responses to slow modulations but no significant responses to rates higher than about 25 Hz, whereas adults showed significant envelope following responses to almost the entire range of stimulus rates.

Conclusion: Our results show that the auditory cortex of preschool-aged children has a sharply limited capacity to process rapid amplitude modulations in sounds, as compared to the auditory cortex of adults.

Significance: These neurophysiological results are consistent with previous psychophysical evidence for a protracted maturational time course for auditory temporal processing. The findings are also in good agreement with current linguistic theories that posit a perceptual bias for low frequency temporal information in speech during language acquisition. These insights also have clinical relevance for our understanding of language disorders that are associated with difficulties in processing temporal information in speech.

Keywords: auditory cortex; auditory steady state response; development; envelope following response; language acquisition; temporal processing

4.1 Introduction

Temporal modulations in the overall amplitude of sounds ("sound envelope") contain critical information for the perception of speech (Drullman, 1995; Drullman et al., 1994; Rosen, 1992; Shannon et al., 1995). For example it is well known that the prosodic content of a spoken sentence is conveyed within the slow temporal fluctuations of the sound envelope (Peelle and Davis, 2012; Rosen, 1992). Conversely, impaired temporal processing has long been associated with language problems including word deafness (Jorgens et al., 2008; Phillips and Farmer, 1990), deficits in speech discrimination (Ali and Jerger, 1992; Souza, 2000), and dyslexia (Ben-Yehudah et al., 2004; Boets et al., 2007; Lehongre et al., 2011; Menell et al., 1999; Putter-Katz, 2005; Walker et al., 2002).

Interest in the topic of auditory temporal processing has been stimulated by recent neurolinguistic models that propose an essential role in speech perception for neural mechanisms that encode the speech envelope (Giraud and Poeppel, 2012; Goswami and Leong, 2013; Gross et al., 2013; Peelle and Davis, 2012). According to such models, intrinsic brain oscillations play a critical role in the analysis of speech, serving to partition the continuous speech signal down into manageable units and to align neural activity with the temporal rhythms of the speech stream (Giraud and Poeppel, 2012; Peele and Davis, 2012). Speech is a dynamic signal that delivers units of critical information at quite different time scales (e.g. intonation/prosody at 500-1000 ms, syllables at 150-300 ms, and phonemic features at 20-80 ms). Therefore neural mechanisms are required for simultaneous sampling of these different speech units. The temporal sampling model proposed by Poeppel et al. (2003) proposes that intrinsic auditory cortical oscillations in the theta range (3-7 Hz) and gamma range (30 – 50 Hz) are "tuned" to track and sample the temporal features of syllables and phonemes respectively. Logically, such models of speech perception must account for the maturation of temporal processing during the critical periods for language acquisition. However at the present time little is known about sound envelope processing in the developing brain.

Behavioural evidence from psychophysical studies suggests that temporal processing undergoes a protracted maturational time course. For example, performance in gap detection continues to improve with age from 3-6 years (Trehub et al., 1995; Wightman et al., 1989) and reaches adult levels of performance at about 8-10 years (Davis and McCroskey, 1980; Irwin et al., 1985). Likewise, detection of amplitude modulations (AM) does not reach mature levels until mid- to late-childhood (Hall and Grose, 1994; Moore et al., 2011). In general, children consistently demonstrate poorer auditory temporal resolution than adults (Banai et al., 2011; Fox et al., 2012; Hall and Grose, 1994; Hill et al., 2004; Moore et al., 2011; Trehub et al., 1995; Wightman et al., 1989). Nevertheless, the reported developmental profiles show considerable variability between studies. Some of this variability may result from children's failure to attend to repetitive stimuli and their failure to comply with experimental requirements. For these reasons, it is difficult to know how much of the difference in performance between children and adults is due to physiological differences in their auditory systems, and how much is due to attentional and motivational variables (Bishop et al., 2011). Consequently it is essential to obtain direct and objective neurophysiological data on temporal processing in the developing brain.

One direct and objective measure of temporal processing is the auditory steady-state response (ASSR), a brain response evoked by rapidly presented periodic stimuli and whose frequency components are stable in both amplitude and phase over a long temporal window (Regan, 1989). ASSRs recorded with electroencephalography (EEG) or magnetoencephalography (MEG) provide an objective and noninvasive metric of the extent to which a periodic stimulus "drives" a neurophysiological response, or conversely, the extent to which the brain response "follows" the temporal patterns of the physical stimulus (Picton et al., 2003). As an alternative to constant stimulation at individual modulation rates, one can characterise brain responses to stimuli that undergo rapid transition (e.g. sweeps) through a range of rates. Since the evoked brain responses to such sweeps are not "steady", these evoked

brain responses are referred to as envelope following responses (EFRs) (e.g. Lehongre et al., 2011; Miyazaki et al., 2013; Purcell et al., 2004).

Both ASSRs and EFRs can be characterised using their associated temporal modulation transfer functions (TMTFs). These transfer functions depict a system's sensitivity to amplitude modulations of an acoustic signal as a function of the modulation rate (Viemeister, 1979). TMTFs for the adult auditory cortex exhibit a low-pass filter profile characterised by a best modulation frequency (BMF) occurring at about 40 Hz and an upper cut-off point at about 50 – 55 Hz, above which the strength of the response begins to decline steadily (Lehongre et al., 2011; Miyazaki et al., 2013; Picton et al., 2003; Poulsen et al., 2007; Poulsen et al., 2009; Purcell et al., 2004). This neurophysiological profile corresponds reasonably well to the psychophysical TMTF for the detection of the amplitude modulation of white noise (Viemeister, 1979).

While ASSRs/EFRs have been extensively characterised in the adult brain, the development of these responses during childhood is less well understood (Picton et al., 2003). Considering the abundant psychophysical evidence suggesting lower temporal resolution in children, it is somewhat paradoxical that EEG measurements of the ASSR/EFR in sleeping infants have identified the BMF to be at about 80 Hz, which is much higher than the BMF for adults (Picton et al., 2003; John et al., 2004; Rickards et al., 1994; Nodarse, et al., 2012). However, EEG responses to rates at about 40 Hz reside primarily in the auditory cortex, whereas adult responses in the 80-100 Hz range primarily involve subcortical (brainstem) responses (Purcell et al., 2004). This suggests that the difference between children's BMF at higher frequency, as compared to the lower BMF for adults, may simply reflect the earlier maturation of brainstem responses in children, as compared to their cortical response (Joris et al., 2004; Moore, 2002; Moore and Linthicum, 2007). The contribution of cortical responses in the infant studies may have been further reduced in previous studies because measurements were made when infants were either in sleep or sedated, both of which may have had

suppressive effects on auditory processing in the cortex (Cohen et al., 1991; Goldstein et al., 1959; Lu et al., 2001). This suggestion is supported by a recent EEG study that examined the ASSR in children aged 6-9 years who were asked to maintain their attention to trains of tone-burst at a range of repetition rates during the recording (Tlumak et al., 2012). The data from this study showed that the children's ASSRs were adult-like at 80 Hz. Nonetheless, in contrast to previous studies in which infants and young children were asleep or sedated, the ASSRs at 20 and 40 Hz in children were smaller than those of adults. Because the EEG ASSRs are dominated by brainstem responses at 80 Hz, but dominated by cortical response below about 50 Hz (Purcell et al., 2004), these findings can be taken as an indication of different developmental trajectories of temporal processing in the cortex and in peripheral regions (Moore, 2002). Interestingly, the magnitudes of responses at low repetition rates (i.e. 0.75 Hz, 1.25 Hz, 2.5 Hz, and 5 Hz) school-aged children were larger than those obtained in adults (Tlumak et al., 2011; Tlumak et al., 2012).

To our knowledge there are no EEG/MEG data on the ASSR/EFR for children who are older than infants or younger than school age. Yet, the preschool years between 3 and 5 years represent a crucial period for the acquisition of language. Given the emerging importance of temporal processing in neurobiological models of language perception (Giraud and Poeppel, 2012; Goswami and Leong, 2013; Gross et al., 2013), it has become imperative to obtain objective neurophysiological data on cortical temporal processing capabilities in children during the years in which they are rapidly mastering the local language. To this end, the current study characterised the EFRs to a range of linguistically-relevant AM rates in a group of healthy, awake preschool-aged children and compared these responses to those of a reference group of healthy adults. In order to focus our measurements on the auditory cortex, EFRs were measured using MEG, which is relatively insensitive to subcortical signals, as compared to EEG (Baillet et al., 2001; Johnson et al., 2010; Nakasatp et al., 1994).

4.2 Methods

4.2.1 Subjects

Twelve children (4 female) aged 3 to 5 years (mean = 49.3 months) participated in the study. All children were right-handed and showed normal hearing based on their parents' reports. Twelve adult participants (7 female) were aged from 22 to 36 (mean = 28.8 years). All the adults were right-handed. Their hearing thresholds were measured using an Otovation Amplitude T3 series audiometer (Otovation LLC, King of Prussia, PA). All adults showed normal hearing thresholds ($\leq$ 20 dB HL) for octave frequencies from 500 to 2000 Hz. All procedures were approved by the Human Subjects Ethics Committee at Macquarie University.

4.2.2 Acoustic stimulation

The acoustic stimulus were created in MATLAB (Mathworks: Natick, MA) and consisted of a 9-second AM sweep in the middle; and 0.3-second segments of unmodulated white noise both at the beginning and the end of the AM sweep. The unmodulated segments of white noise served to separate the envelope following response from the participants' responses to sound onset and offset. The AM sweep was generated by modulating a 9-second white noise with an equal duration sweep changing exponentially from 1 Hz to 80 Hz. A noise carrier was used to eliminate any spectral cues from the stimulus, and the modulation rates encompassed the range of linguistically-relevant AM rates (Rosen, 1992). The modulation depth was 100%. The exponential function used to create the stimulus was: $f(t) = f_0 \times \left(\frac{f_1}{f_0}\right)^{\frac{t}{t_1}}$, where $f(t)$ is the frequency changing by time, t, f_0 = 1 (Hz) is the starting frequency, f_1= 80 (Hz) is the ending frequency, $t_1 = 9\ (second)$ is the sweep duration. The function therefore can be simplified as: $f(t) = 80^{t/9}$.

The temporal waveform of the complete stimulus and spectrogram of the sweep envelope are presented in Figure 1. The sounds were delivered binaurally using insert

earphones (Model ER-30, Etymotic Research Inc., Elk Grove Village, IL) at a level of 70-75 dB SPL.

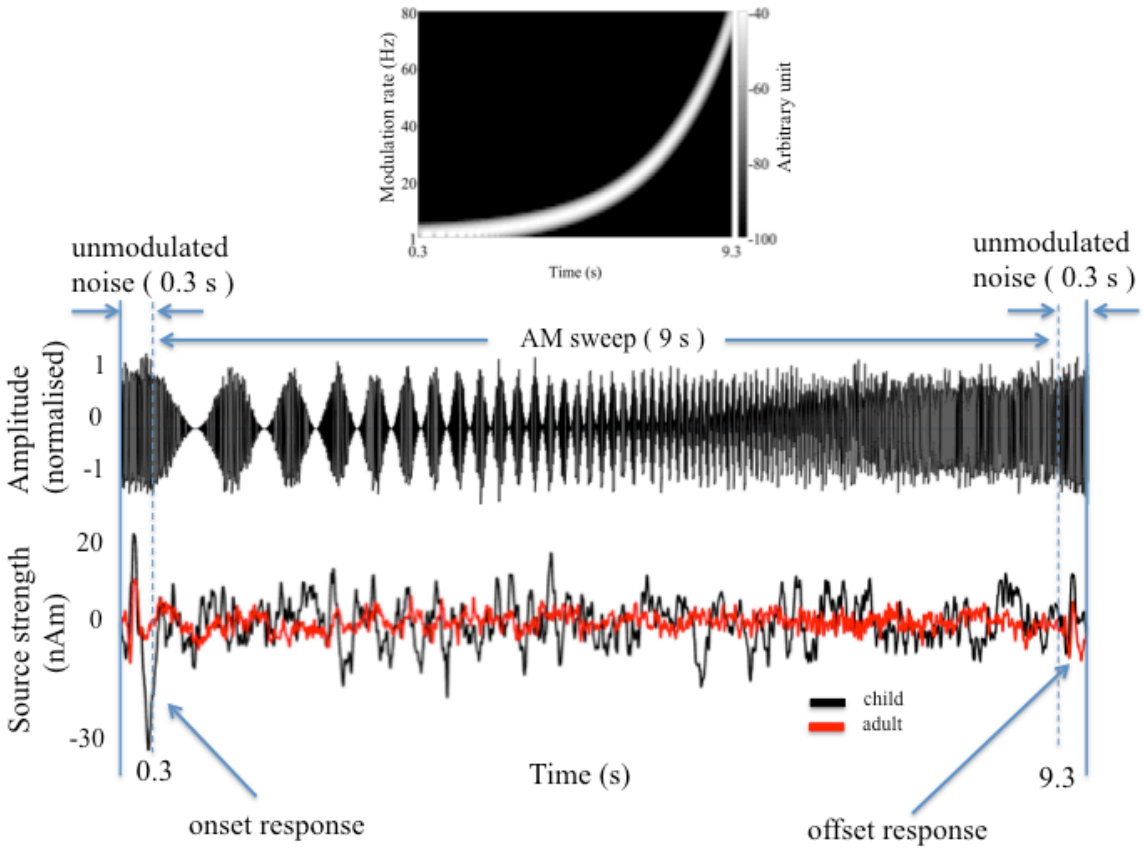

Figure 1. *Experimental paradigm. The acoustic stimulus (middle panel) had a total duration of 9.6 sec, beginning and ending with 300 ms of unmodulated white noise. From 0.3 to 9.3 sec the white noise carrier was amplitude modulated (100% modulated depth) with a logarithmic sweep over modulation rates from 1-80 Hz. The top panel shows the spectrogram of this logarithmic sweep. The bottom panel shows auditory cortex source waveforms (average of both hemispheres) for a representative child (black) and adult (red).*

4.2.3 Procedure

Participants listened to the stimuli passively while watching a muted movie. We adopted this passive listening strategy in order to make the experimental setup appealing enough for the child participants to remain in the MEG scanner for the duration of the experiment. Passive listening paradigms have been used in many previous EFR and ASSR studies (see Purcell et al., 2004; Picton et al., 2003).

Stimulus presentation was controlled using Experiment Builder 1.10.165 (SR Research: Mississauga, Ontario, Canada). The acoustic stimulus was presented with a mean inter-stimulus interval of 950 ms, randomly selected from a rectangular distribution between 900 ms and 1000 ms. 100 trials were presented in one block to children, while 200 trials were presented in two blocks to adults. For adult, there was a 5-minute break between two blocks. In the present analyses, only the first 100 trials of adult data were included for comparison to the child data. The full set of adult data will be presented in a separate report. To effectively convey instructions to young children and minimise head movement artifacts during MEG recordings, a child-friendly data acquisition protocol was employed (Johnson et al., 2010). The whole experiment took about half an hour including the MEG set up for children; the more comprehensive recordings in adults required about 1 hour.

4.2.4 MEG recordings

Prior to data acquisition, the participants were fitted with five head position marker coils. The position of the five coils and the head shape of the subject were digitised using a Polhemus Fastrak digitiser (Colchester, VT). Head positions were recorded before and after each block to measure head movement. Brain activity was recorded using a whole-head 64-channel paediatric MEG system (Model PQ1064R-N2m, KIT, Kanazawa, Japan) for children; and a whole-head 160-channel MEG system (Model PQ1160R-N2, KIT, Kanazawa, Japan) for adults. Both MEG systems consisted of first-order axial gradiometers with a 50 mm baseline and located in the same magnetically shielded room (Fujihara Co. Ltd., Tokyo, Japan). Sensor configurations for the two systems are described in detail in Johnson et al. (2010). All measurements were carried out with subjects in a supine position. MEG data were acquired continuously with a sampling rate of 1000 Hz and band-pass filtered between 0.03 Hz and 200 Hz using an analogue filter.

4.2.5 Analyses

MEG data were analysed using BESA Research Version 6.0 (BESA Research GmbH: Grafelfing, Germany). For both groups, the MEG data were coregistered to the template structural MRI implemented in BESA Research 6.0.

4.2.5.1 Dipole fitting

Data were analysed using a spatial filter consisting of bilateral dipoles fitted to auditory evoked fields (AEFs) elicited by the onsets of the acoustic stimuli (the P100m in children and the P50m in adults). AEFs were epoched from -100 to 400 ms with respect to the onset of each stimulus. The epoched data was averaged across trials and bandpass filtered between 2 and 20 Hz (Aiken and Picton, 2008). Symmetric dipoles were fitted within a latency window of 100 – 200 ms for each child subject and 50 – 100 ms for each adult subject. For both groups, mean dipole locations were in the superior temporal gyri in the vicinity of the primary auditory cortex in each hemisphere (Figure 2 and Table 1).

Table 1. Talairach coordinates for grand mean dipole locations (mm)

Child	Mean	45.76	-14.53	9.13
	STD	4.79	10.93	12.93
Adult	Mean	51.12	-21.62	4.4
	STD	5.51	6.30	6.91

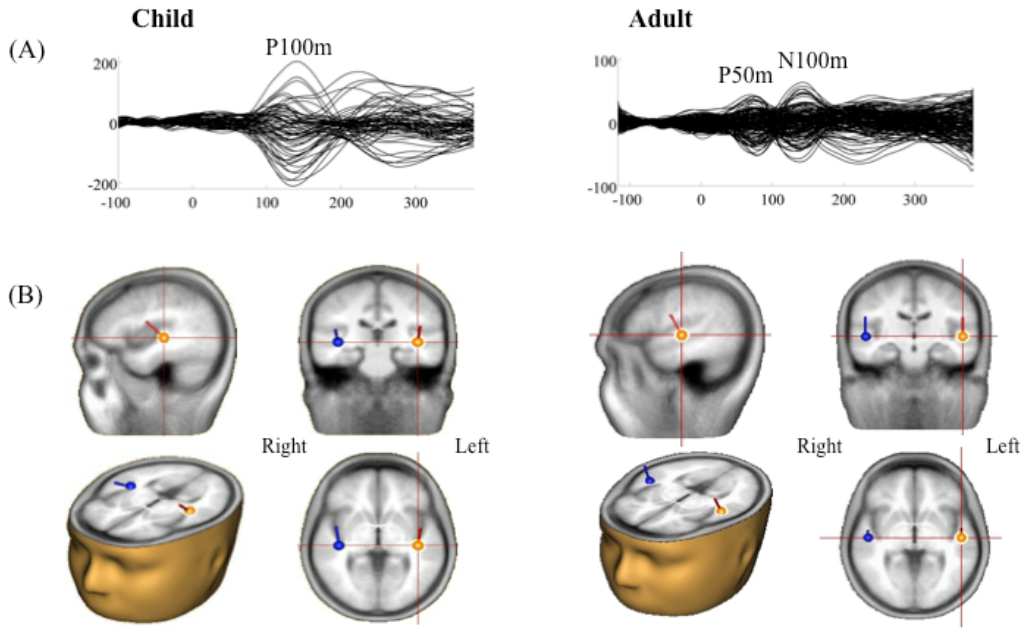

Figure 2. *(A) Overlaid sensor waveforms for a representative child and adult participant. (B) Grand mean dipole source locations for P100m (children) and P50m (adults).*

Further analyses were conducted on the waveforms reconstructed from each dipole source. The dipole source analysis procedure takes into account the physical geometry of the measurement sensors with respect to individual head shapes (see Irimia et al., 2014 for a discussion of this issue), generating a stable and comparable source estimate from each data set and permitting comparison of brain responses recorded from the two different MEG systems (He et al., 2015).

4.2.5.2 Spontaneous cortical oscillations

To depict the ongoing brain activities in the absence of auditory stimulation, the power spectrogram of part of the inter-stimulus interval (-600 to 0 ms) were computed using the two bilateral source montage derived from the dipole fitting procedure.

4.2.5.3 Auditory evoked fields

Source waveforms were constructed from each dipole in each hemisphere for each subject.

4.2.5.4 Time-frequency analyses

Using the source montage of the bilateral auditory cortical sources derived from dipole fitting to the AEFs, spectral analyses were then carried out on the continuous source waveforms over a frequency range of 1 to 80 Hz within a new epoch window from 600 ms pre-stimulus to 9600 ms post-stimulus. Time-frequency analyses were then conducted using the complex demodulation method implemented in BESA Research 6.0 (Hoechstetter et al., 2004), with a frequency step of 1 Hz and a time step of 50 ms.

Inter-trial phase coherence (ITPC), which quantifies phase consistency cross trials for each frequency and time point, was computed from the single trial time-frequency representation using the following formula.

$$ITPC(t,f) = \frac{1}{N}\sum_{k=1}^{N} e^{i\varphi_k(t,f)}$$

where N = total number of trials; and $\varphi_k(f,t)$ = the phase in trial k.

The ITPC was used as an index of the EFR, because it offers a much more sensitive measure of stimulus-synchronised brain activity than response power (Ding and Simon, 2013).

Permutation tests (Maris and Oostenveld, 2007) were then carried out on the time-frequency data for each group using custom MATLAB scripts to determine the statistical significance of the envelope following pattern. The permutation tests were applied to the 0.3 - 9.3 second interval during which the temporal modulation rate of the stimulus swept from 1 Hz to 80 Hz. A primary threshold was used to identify the top 5% of values. This clustering algorithm used the sum of activities within a cluster. Only clusters surviving a permutation with 1000 iterations at the significance level of 0.05 were accepted.

To better visualise the relationship between stimulus modulation rate and brain response, a linearisation procedure was applied to convert the time axis to modulation rate m, based on the logarithmic function $m(t) = 80^{t/9}$ used to generate the AM. Model fitting was then performed in MatLab (MathWorks, Natick, MA) using the Curve Fitting Toolbox (version

3.4.1) to model the correlation between modulation rate and frequency of brain response. A linear model ($f = a * m + b$; here f is the frequency of brain response, m is the rate of modulation, a and b are two parameters set for free fitting) was used to fit the permuted data. This linear modeling was conducted separately for each group. Individual EFRs to AM at rate *m* were identified as the mean magnitude within the bin [*f-1 f+1*]. The vector strength of the defined EFR was then re-plotted against the frequency of the AM, in oder to visualises the TMTF. Within each group, the same linear model generated from the permuted data was used to compute TMTF for each subject.

The ITPC values are proportional data (valued between 0 and 1). The distributions are not Gaussian and do not meet the distributional assumptions required for linear tests, such as ANOVA or t-tests. Therefore whenever linear tests were performed on ITPC values, we the values were first converted to normalised values using the rationalized arcsine transform function (Studebaker, 1985).

4.3 Results

4.3.1 Spontaneous oscillations

Figure 3 shows the average power of ongoing oscillations (combined hemispheres) as a function of its frequency in the absence of auditory stimulation. Both children and adults showed a prominent peak in the alpha band, with children showing a lower peak alpha frequency compared to adults (child: 8.67 ± 0.67 Hz, adult: $10.7 \pm$ **1.27 Hz, p <** 0.001), consistent with previous EEG measurements of peak alpha frequency in toddlers (Marshall et al., 2002; Saby and Marshall, 2012). Children showed significantly higher alpha amplitudes than adults (child: 28.77 ± 6.89 fT; adult: 16.82 ± 9.83 fT, $p < 0.001$). Figure 3 shows that children's oscillations were larger in magnitude compared to adults throughout the low frequency range (< 10 Hz). While mean alpha amplitude was slightly larger in the right

hemisphere in both groups (Figure 3, inset), these differences were not statistically significant in either group.

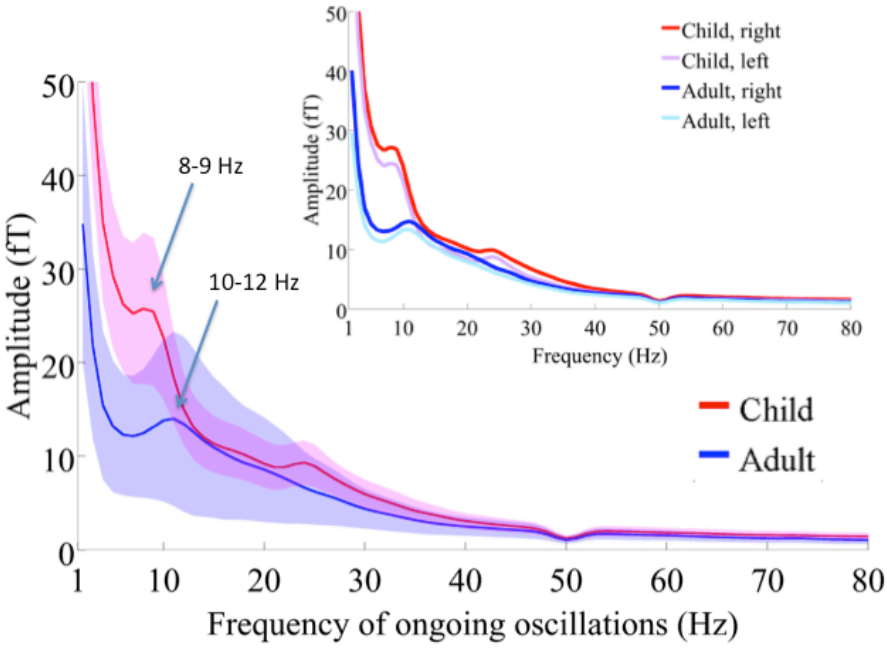

Figure 3. *Group mean profiles of spontaneous brain oscillations from auditory cortical sources measured in the absence of acoustic stimulation. Shading shows one standard deviation. Main figure shows average of two hemispheres. Inset shows profiles for each hemisphere.*

4.3.2 AEFs

Figure 4.A shows grand averaged AEF source waveforms in the two groups. In children the most prominent auditory AEF is a circa 100 ms peak (mean latency, 117 ms) termed the P100m, believed to be the precursor of the adult P50m response (Johnson et al., 2010; Johnson et al., 2013; Lippe et al., 2009). Surface topographic maps of the child P100m and adult P50m are shown in Figure 4. Statistical analyses showed no significant hemispheric difference in amplitude or latency for any of the adult or child AEF peaks.

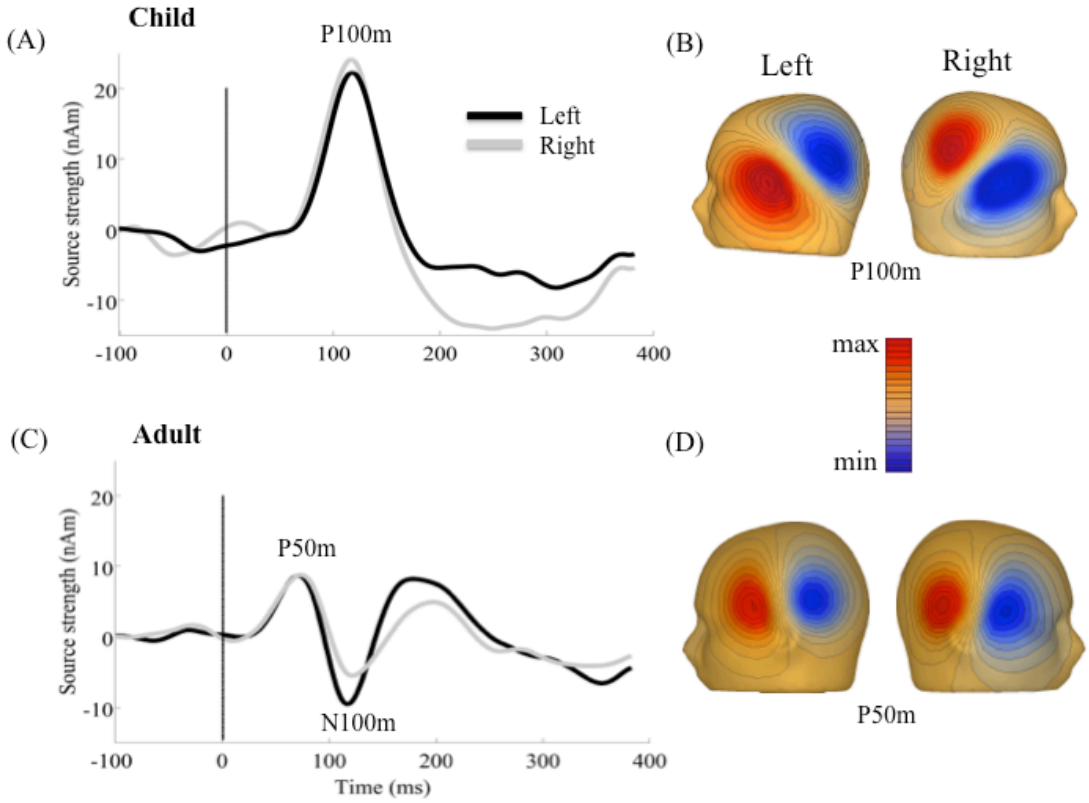

Figure 4. *Grand mean auditory event-related fields elicited by stimulus onsets and surface topographic distributions of P100m.*

4.3.3 EFRs

Figure 5 shows ITPC spectrograms and TMTFs for children and adults. From Figure 5B and D, significant phase-locking was obtained for the full range of modulation rates with strongest phase-locking values in the range of 35-50 Hz (corresponding to the "auditory 40 Hz response"; see Galambos et al., 1981; Picton et al., 2003) in adults. In contrast, significant phase-locking was obtained in children only at modulation rages below ~25 Hz with maximal ITPC values at rates of 12-18 Hz (see Figure 5A, C, E).

Figure 5E presents the grand mean and standard error of TMTF in both participant groups. Greater ITPC values for adults were visible in two frequency ranges (7 – 11 Hz, and 15 –80 Hz). Mean ITPC value within each frequency range were then calculated for each group and t-tests (one-tailed) on these mean values confirmed that the these differences were statistically significant: 7-11 Hz: adults $\overline{ITPC}_{(7\text{-}11\ Hz)}$ = 0.16, $SD_{(7\text{-}11\ Hz)}$ = 0.05, children $\overline{ITPC}$ (7-

$_{11\ Hz)}$ = 0.10, $SD_{(7\text{-}11\ Hz)}$ = 0.02, p = 0.006; 15-80 Hz: adults $\overline{ITPC}_{(15\text{-}80\ Hz)}$ = 0.19, $SD_{(15\text{-}80\ Hz)}$ =0.10, children $\overline{ITPC}_{(15\text{-}80\ Hz)}$ = .10, $SD_{(15\text{-}80\ Hz)}$ = 0.01, p = 0.008. These were post hoc t-tests based on prior observations. Because the test direction was known in advance, the use of one-tailed t-tests was justified.

Figure 6A and B show ITPC spectrograms computed separately for right and left hemispheres in the two groups. Figure 6C depicts the grand mean and standard error of hemispheric TMTFs in the two groups. Visual inspection of the children's plots indicates slightly greater ITPC values in the right hemisphere (Figure 6A and C), at least within the 1-25 Hz range where EFRs were statistically significant in this group (Figure 5C). T-test (two-tailed) confirmed that this difference is statistically significant: (mean left $\overline{ITPC}_{(1\text{-}25Hz)}$ = 0.11, $SD_{(1\text{-}25Hz)}$ = 0.02; mean right $\overline{ITPC}_{(1\text{-}25Hz)}$ = 0.13, $SD_{(1\text{-}25Hz)}$ = 0.03, p = 0.004).

Figure 6D explicitly compares hemispheric lateralisation in the two age groups. For this plot, a lateralization index (LI) was computed as (right ITPC – left ITPC)/ (right ITPC + left ITPC), so that negative values indicate left lateralization and positive values indicate right lateralization. This plot indicates that children are more right lateralised than adults over a range of modulation rates between about 10-20 Hz. T-tests (two-tailed) were computed for the mean ITPC values over this range, and the results confirmed a significant difference in lateralization between the two groups (children $\overline{LI}_{(10\text{-}20Hz)}$ = 0.03, $SD_{(10\text{-}20Hz)}$ = 0.03; adults $\overline{LI}_{(10\text{-}20Hz)}$ = 0.01, $SD_{(10\text{-}20Hz)}$ = .0.04, p = 0.005).

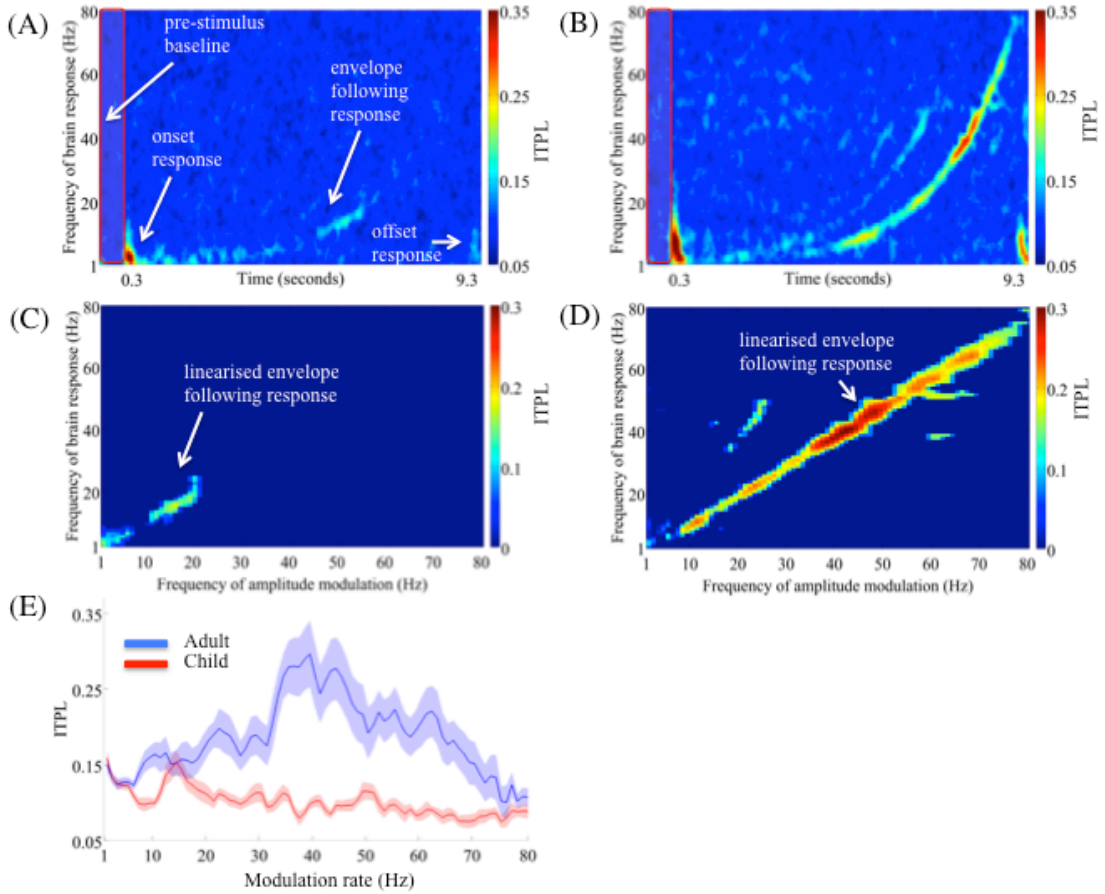

Figure 5. *Grand mean ITPC values measured from auditory cortical sources (average of two hemispheres) and indexed in (A) children and (B) adults. The pre-stimulus baseline is highlighted with a red rectangle. (C) ITPC values that survived permutation testing in children and (D) adults. Permutation tests were performed on data in A and B between 0.3 – 9.3 s (during the amplitude modulated portion of the stimulus). The exponential patterns in A and B are the envelope following responses plotted against time. Over the 0.3 - 9.3 sec time epoch the stimulus was amplitude modulated over rates of 1-80 Hz in an exponential fashion. Consequently, the EFR follows the exponential pattern of modulation rates (Fig 1 top). The linear patterns in Figure 5 C and D are the same data as Figure 5A and B, thresholded with permutation testing and replotted by converting the time axis to amplitude modulation rate. This has the effect of linearizing the EFR pattern to directly show the relationship between AM rate and frequency of the following response. (E) Comparison of child and adult TMTFs. Shading indicates standard errors.*

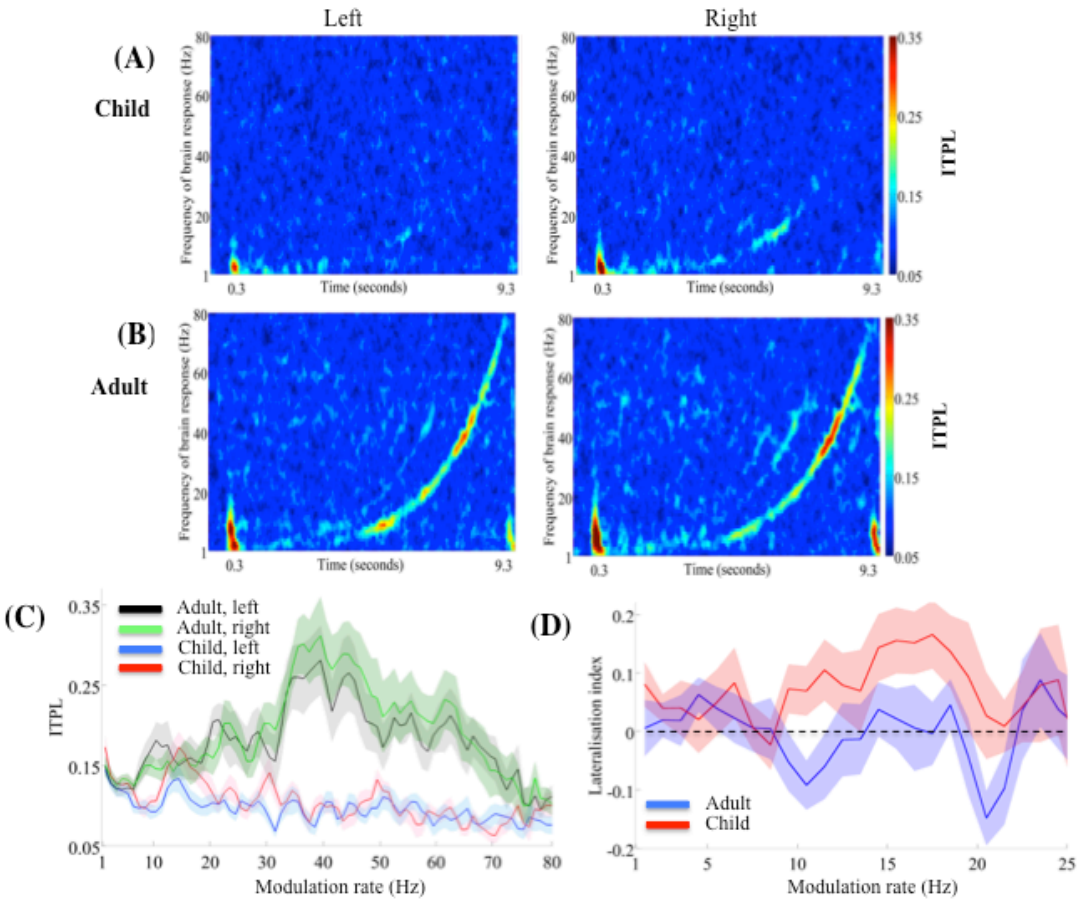

Figure 6. *Hemispheric comparison of EFRs for children and adults. (A) Grand mean spectrogram of ITPC in each hemisphere in children and (B) in adults. (C) Hemispheric comparison of ITPC values for both groups. (D) Between-group comparison of hemispheric lateralization index.*

4.4 Discussion

The present results are the first comprehensive neurophysiological description of temporal envelope processing in the auditory cortex of preschool-aged children. The capacity to process temporal modulation is essential for speech perception and is presumably an important component of language acquisition. In agreement with psychophysical evidence for limited temporal processing capabilities in children, our data show that EFR phase-locking in children's auditory cortex drops off sharply at AM rates higher than about 15 Hz, with no detectable phase-locking at rates higher than about 25 Hz. The immature phase-locking profile of children stands in sharp contrast to the adult profile which shows robust phase-locking throughout the range of AM rates, with maximal ITPC values in the range of 30-50 Hz. Tying

our results to the previous literature, we found that in the absence of acoustic stimulation, slow oscillatory activity of children's auditory cortex demonstrated higher power than adults, consistent with previous EEG work (Marshall et al., 2002; Saby and Marshall, 2012; Yordanova, 2008). In addition, while EFRs were relatively weak, stimulus onsets elicited robust AEFs in the same children, arguing against the possibility that there were deficiencies in either the delivery of the acoustic stimuli or in the neuromagnetic measurements of the children.

4.4.1 Spontaneous brain rhythms

Our finding of a greater magnitude of MEG slow oscillations in children during the inter-stimulus intervals of the present experiment is in agreement with previous EEG studies describing age-related changes of the brain oscillations in resting state (Gasser et al., 1988; Whitford et al., 2007; Yordanova, 2008). Whitford and colleagues (2007) examined the EEG power in subjects aged from 10 to 30 years old. They found that slow wave (sub-alpha) EEG power declined as a function of age, with the most rapid decline during adolescence, and that this decline mirrored the age-related reduction in gray matter measured with structural MRI. The authors concluded that both the structural and neurophysiological changes were due to elimination of active cortical synapses during development.

Our finding of a lower alpha peak frequency in children is also entirely consistent with previous EEG findings (Marshall et al., 2002; Saby and Marshall, 2012; Srinivasan, 1999). For example, Srinivasan (1999) measured alpha EEG in an older cohort of children (aged 6-9 years) and reported differences in peak alpha frequency between children (mean 8.9 Hz) and adults (mean 10.1 Hz) that are very similar to those obtained in the present MEG study (8.7 Hz and 10.7 Hz respectively). These findings imply that the increase of alpha peak frequency may continue until at least 9 years of age. In the EEG study, the measures of long-range alpha coherence also increased with age, indicating that the shift in peak alpha frequency is attributable to the establishment of long range cortico-cortical connections in the maturing neocortex. In general the maturation of the neocortex is characterised by a rearrangement of

cortical networks and an increase in temporal precision and synchrony of network interactions, resulting in changes of peak oscillatory frequencies within the network and an increase in magnitude of ongoing high frequency oscillations (Uhlhaas, 2010). Reduced temporal precision and synchrony of the immature cortex is likely to be responsible for our finding of sharply reduced EFR capacity in auditory cortex of children, discussed below.

It has been shown that the waking brain rhythms in the alpha range encompass at least three functionally and spatially different oscillations: the classic posterior alpha rhythm, the rolandic mu rhythm, and a lesser-known temporal rhythm (or possibly, set of rhythms) described by various investigators as "auditory alpha rhythm" (Weisz et al., 2011), "tau rhythm" (Lehtela et al., 1997), "third rhythm" (Neidermyer, 1990) or "breach rhythm" (Neidermyer, 1991). While the spatial filtering of the bilateral auditory cortical dipoles employed in this study suggests that our analyses should be biased to the auditory alpha rhythm, it is entirely possible that the alpha peaks in Figure 3 represent some weighted combination of auditory alpha and the much larger amplitude posterior alpha and mu rhythms (Weisz et al., 2011). Therefore the present data cannot specify if the alpha peaks in Figure 3 are modality specific (i.e. generated in auditory cortex) or not. This issue merits examination in future studies.

4.4.2 AEFs

AEFs elicited by sounds onsets in children's auditory cortex showed a P100m peak with a mean latency of about 120 ms. The P100m, and its EEG equivalent, the P100, have been reported in previous electrophysiological studies (Fujioka et al., 2006; Heim et al., 2003; Johnson et al., 2010; Paetau et al.,1995; Ponton et al., 2000; Sharma et al., 1997) and is considered to be an immature version of the P50m/P50 in the adult response. The latency and amplitude of this P100m/P100 component decrease with increasing age, achieving a peak latency of about 50 ms in the adult brain (Ponton et al., 2000; Sharma et al., 1997).

4.4.3 EFRs

Previous EEG studies of the ASSR/EFR in children have described a response that is at least partly and perhaps largely generated by subcortical sources (Purcell et al., 2004). MEG is considered to be relatively insensitive to subcortical sources, due to a much steeper dependence on distance from the measuring sensor compared to EEG (Baillet et al., 2001; Johnson et al., 2010; Nakasatp et al., 1994). While it is in principle possible to detect brainstem signals with MEG, it requires extraordinary efforts (e.g. thousands of trials) to achieve the signal-to-ratio (SNR) required to extract subcortical responses such as wave V of the auditory brainstem response (ABR; Parkkonen et al., 2009). The present MEG results therefore represent responses that originate mainly or entirely from the auditory cortex with little or no contribution from subcortical regions. The ability to focus our measurements specifically on cortical activity is important because the subcortical and cortical components of the auditory system have quite different temporal encoding capabilities in the mature brain. Single unit measurements show that the upper limit of synchronizing via phase-locking to the periodicity of a sound decreases progressively and markedly at higher levels of the auditory system, from about 4 kHz at the level of the auditory nerve to about 50 Hz at the level of primary auditory cortex (Joris et al., 2004). A further complication is that subcortical and cortical levels of the auditory system have quite different maturational trajectories (Moore, 2002; Moore and Linthicum, 2007).

Our results indicate that the immature auditory cortex has a limited capacity to encode and represent sound periodicities relative to the adult brain, with a sharp drop-off in phase-locking at rates higher than about 15 Hz and with no measurable following response to AMs faster than about 25 Hz. This indicates that sound envelope encoding, like other auditory cortical responses (Kraus et al., 1993; Moore, 2002; Moore and Linthicum, 2007; Ponton et al., 2000; Sharma et al., 1997; Sussman et al., 2008; Wunderlich and Cone-Wesson, 2006) has a prolonged maturational trajectory, with earlier development of responses to slow envelope periodicities and much later development of responding to high frequency modulations.

4.4.4 Right lateralised temporal processing in the developing brain

Children exhibited more consistent rightward dominance, whereas adults showed a combination of bilateral and leftward dominance within the alpha and beta bands. This developmental change in lateralization suggests a developmental delay in the functional specialization established at the level of sensory processing (Ross et al., 2005). In fact, the rightward dominance of EFRs in early development fits well with that hypothesis of prosody-dependent language acquisition. Specifically, the phonological bootstrapping hypothesis proposes that language acquisition in children relies on prosodic cues to segment the incoming auditory stream into linguistic units. In children, evidence has consistently pointed to a strong involvement of right hemisphere during speech prosody perception (Gandour et al., 2004; Homae et al., 2006; Plante, Holland, and Schmithorst, 2006; Wartenburger et al., 2007). For example, processing prosody in isolation elicits larger right hemisphere activation whereas normal speech is associated with greater activation in the left hemisphere (Wartenburger et al., 2007). Since prosodic cues is conveyed in the slow modulations in speech, our finding of the right dominance is consistent with previous studies using other methods to present prosodic information in isolation, such as hummed speech (Wartenburger et al., 2007). It has been suggested that altered patterns of auditory lateralization are responsible for pathological (e.g. dyslexia and schizophrenia) and supranormal (e.g. absolute pitch) cognitive function (Abrams et al., 2009; Johnson et al., 2013; Poelmans et al., 2012; Tervaniemi and Hugdahl, 2003)

4.4.5 Further implications

The current study was designed to assess the temporal processing capabilities of auditory cortex in young children. While we used a non-speech (noise) stimulus in order to systematically characterise brain responses across a range of linguistically-significant temporal modulation rates, the results are relevant to theoretical frameworks that suggest an important role for temporal encoding mechanisms in speech perception and in language acquisition.

4.4.5.1 Implications for speech perception

Recent neurolinguistic models suggest that the neural mechanisms of encoding temporal modulation of speech envelope play an essential role in speech perception (Giraud and Poeppel, 2012; Peelle and Davis, 2012). Application of such models of speech perception to dyslexia draws on the logic that abnormal temporal sampling of speech will result in inaccurate representations of speech features and impaired speech processing. However it remains unclear what sampling rate is specifically affected and how this might impair phonological representations in dyslexia. A number of authors have suggested that the perceptual and phonological problems associated with dyslexia are specifically associated with impaired oscillatory processing in the slow theta range (Abrams et al., 2009; Giraud et al., 2005; Goswami et al., 2011). In the temporal sampling framework proposed by Goswami and colleagues (2011), impaired phase-locking of theta rhythms results in impaired syllable parsing and difficulties in perceiving the components of syllables. On the other hand, others have emphasised problems in temporal sampling at higher frequency gamma band rates, resulting in abnormal representations of phonemes (Lehongre et al., 2011).

Notably, the temporal sampling models described above are presently based entirely on neurophysiological data obtained from adults. The present study contributes to this issue with evidence that children show quite different cortical phase locking profiles than adults: The immature human auditory cortex emphasises low AM rates, and, in contrast to the adult reference group, show no measureable phase-locking at rates higher than about 25 Hz. These neurophysiological results are entirely consistent with behavioural evidence that children are strongly biased towards low frequency temporal information in the speech stream during language acquisition (Christophe and Dupoux, 1996; Christophe et al., 2003; Mehler and Christophe, 1995); with the fact that temporal modulations below about 20 Hz are most crucial for speech recognition/intelligibility (Drullman, 1995; Shannon et al., 1995; Smith et al., 2002); and with explanations of dyslexia that posit problems in sampling of low temporal rates in

speech (Abrams et al., 2009; Giraud et al., 2005; Goswami et al., 2011). Temporal sampling models of language perception in general must accommodate the fact that children's temporal processing capabilities are sharply limited in comparison to adults.

4.4.5.2 Implications for language acquisition

Current views of language acquisition posit that the slow temporal rhythms associated with the prosodic content of speech are crucial for language acquisition (Peelle and Davis, 2012; Rosen, 1992). One proposal, known as the phonological bootstrapping hypothesis, contends that infants and young children depend on phrasal prosodic cues to initially segment continuous speech to form pre-lexical representations for early word learning (Christophe and Dupoux, 1996; Ramus et al., 1999). This provides a non-lexical method of finding word boundaries, which must exist for infants to acquire the words of a language (Christophe and Dupoux, 1996; Christophe et al., 2003; Ramus et al., 1999). On this view, prosodic cues are used by infants to segment the speech stream into prosodic units that are smaller than sentences but bigger than words. Lexical acquisition would then be performed on the basis of this prosody-segmented pre-lexical representation. Such a representation would be useful for the acquisition of phonology and syntax (Christophe and Dupoux, 1996; Ramus et al., 1999). In other words, prosody or rhythm plays an important role in speech processing and phonological bootstrapping is an essential ingredient of language acquisition (Christophe and Dupoux, 1996; Christophe et al., 2003; Mehler and Christophe, 1995). In this context our observed right hemispheric lateralization of EFRs in children is of interest because it has long been held that prosodic information is preferentially processed in the right hemisphere (Friederici and Alter, 2004; Gandour et al., 2004; Homae et al., 2006; Plante et al., 2006; Wartenburger et al., 2007).

4.5 Conclusions

These are the first MEG data on the auditory EFRs for healthy preschool-aged children, an age range that is under-represented in the neurophysiological and neuroimaging literatures. Our

results show that the auditory cortex of preschool-aged children has a sharply limited capacity to process amplitude modulations in sound that are faster than about 20 - 25 Hz. These neurophysiological results are consistent with psychophysical evidence for a protracted maturation time course for auditory temporal processing and with current linguistic theories that posit a perceptual bias for low temporal frequencies in speech during language acquisition. These insights also have clinical relevance for our understanding of language disorders associated with problems in processing temporal information in speech.

4.6 References

Abrams, D.A., Nicol, T., Zecker, S., and Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci., 29*, 7686–7693.

Aiken, S. J., and Picton, T. W. (2008). Human cortical responses to the speech envelope. *Ear Hear., 29*(2), 139-157.

Ali, A. A., and Jerger, J. (1992). Phase coherence of the middle-latency response in the elderly. *Scand. Audiol., 21*(3), 187-194.

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. *Hear. Res.*, *65*(1), 253-261.

Baillet, S., Mosher, J. C., and Leahy, R. M. (2001). Electromagnetic brain mapping. *Signal Process. Mag., IEEE, 18*(6), 14-30.

Banai, K., Sabin, A. T., and Wright, B. A. (2011). Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hear. Res., 280*(1-2), 219-227.

Ben-Yehudah, G., Banai, K., and Ahissar, M. (2004). Patterns of deficit in auditory temporal processing among dyslexic adults. *NeuroReport, 15*(4), 627-631.

Bishop, D. V. M., Hardiman, M. J., and Barry, J. G. (2011). Is auditory discrimination mature by middle childhood? A study using time-frequency analysis of mismatch responses from 7 years to adulthood. *Dev. Sci., 14*(2), 402-416.

Boets, B., Wouters, J., van Wieringen, A., and Ghesquiere, P. (2007). Auditory processing, speech perception and phonological ability in pre-school children at high-risk for dyslexia: a longitudinal study of the auditory temporal processing theory. *Neuropsychologia, 45*(8), 1608-1620.

Christophe, A., and Dupoux, E. (1996). Bootstrapping lexical acquisition: The role of prosodic structure. *Linguist. Rev., 13*(3-4), 383-412.

Christophe, A., Nespor, M., Guasti, M. T., and Ooyen, B. V. (2003). Prosodic structure and syntactic acquisition: the case of the head-direction parameter. *Dev. Sci., 6*(2), 211-220.

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. *J. Acoust. Soc. Am., 90*(5), 2467-2479.

Davis, S. M., and McCroskey, R. L. (1980). Auditory fusion in children. *Child Dev., 51*(1), 75-80.

Ding, N., and Simon, J. Z. (2013). Power and phase properties of oscillatory neural responses in the presence of background activity. *J. Comput. Neurosci., 34*(2), 337-343.

Drullman, R. (1995). Temporal envelope and fine structure cues for speech intelligibility. *J. Acoust. Soc. Am., 97*(1), 585-592.

Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of temporal envelope smearing on speech reception. *J. Acoust. Soc. Am., 95*(2), 1053-1064.

Fox, A. M., Reid, C. L., Anderson, M., Richardson, C., and Bishop, D. V. (2012). Maturation of rapid auditory temporal processing and subsequent nonword repetition performance in children. *Dev. Sci., 15*(2), 204-211.

Friederici, A. D., and Alter, K. (2004). Lateralization of auditory language functions: A dynamic dual pathway model. *Brain Lang., 89*(2), 267-276.

Fujioka, T., Ross, B., Kakigi, R., Pantev, C., and Trainor, L. J. (2006). One year of musical training affects development of auditory cortical-evoked fields in young children. *Brain, 129*, 2593-2608.

Galambos, R., Makeig, S., and Talmachoff, P. J. (1981). A 40-Hz auditory potential recorded from the human scalp. *PNAS, 78*(4), 2643-2647.

Gandour, J., Tong, Y., Wong, D., Talavage, T., Dzemidzic, M., Xu, Y. et al. (2004). Hemispheric roles in the perception of speech prosody. *NeuroImage, 23*(1), 344-357.

Gasser, T., Verleger, R., Bacher, P., and Sroka, L. (1988). Development of the EEG of school-age children and adolescents. I. Analysis of band power. *Electroencephalogr. Clin. Neurophysiol., 69*, 91-99.

Giraud, K., Demonet, J.F., Habib, M., Marquis, P., Chauvel, P., and Liegeois-Chauvel, C. (2005). Auditory evoked potential patterns to voiced and voiceless speech sounds in adult developmental dyslexics with persistent deficits. *Cereb. Cortex, 15*, 1524–1534.

Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: Emerging computational principles and operations. Nat. Neurosci., 15(4), 511-517.

Goldstein, M. H., Kiang, N. Y. S., and Brown, R. M. (1959). Responses of the Auditory Cortex to Repetitive Acoustic Stimuli. *J. Acoust. Soc. Am., 31*(3), 356-364.

Goswami, U. (2011). A temporal sampling framework for developmental dyslexia. *Trends Cogn. Sci., 15*(1), 3-10.

Goswami, U., and Leong, V. (2013). Speech rhythm and temporal structure: Converging perspectives? *Lab. Phonol., 4*(1), 67-92.

Gross, J., Hoogenboom, N., Thut, G., Schyns, P., Panzeri, S., Belin, P., and Garrod, S. (2013). Speech rhythms and multiplexed oscillatory sensory coding in the human brain. *PLoS Biol., 11*(12), e1001752.

Hall, J. W., and Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *J. Acoust. Soc. Am., 96*(1), 150-154.

He, W., Brock, J., and Johnson, B. W. (2015). Face processing in the brains of pre-school aged children measured with MEG. *NeuroImage, 106*, 317-327.

Heim, S., Eulitz, C., and Elbert, T. (2003). Altered hemispheric asymmetry of auditory P100m in dyslexia. *Eur. J. Neurosci., 17*(8), 1715-1722.

Hill, P. R., Hartley, D. E. H., Glasberg, B. R., Moore, B. C. J., and Moore, D. R. (2004). Auditory Processing Efficiency and Temporal Resolution in Children and Adults. *J. Speech Lang. Hear. Res., 47*(5), 1022-1029.

Hoechstetter, K., Bornfleth, H., Weckesser, D., Ille, N., Berg, P., and Scherg, M. (2004). BESA Source Coherence: A new method to study cortical oscillatory coupling. *Brain Topogr., 16*(4), 233-238.

Homae, F., Watanabe, H., Nakano, T., Asakawa, K., and Taga, G. (2006). The right hemisphere of sleeping infant perceives sentential prosody. *Neurosci. Res., 54*(4), 276-280.

Irimia, A., Erhart, M. J., and Brown, T. T. (2014). Variability of magnetoencephalographic sensor sensitivity measures as a function of age, brain volume and cortical area. *Clin. Neurophysiol., 125*(10), 1973-1984.

Irwin, R. J., Ball, A. K. R., Kay, N., Stillman, J. A., and Rosser, J. (1985). The development of auditory temporal acuity in children. *Child Dev., 56*(3), 614-620.

John, M. S., Brown, D. K., Muir, P. J., and Picton, T. W. (2004). Recording auditory steady-state responses in young infants. *Ear Hear.*, *25*(6), 539-553.

Johnson, B. W., Crain, S., Thornton, R., Tesan, G., and Reid, M. (2010). Measurement of brain function in pre-school children using a custom sized whole-head MEG sensor array. *Clin. Neurophysiol., 121*(3), 340-349.

Johnson, B. W., McArthur, G., Hautus, M., Reid, M., Brock, J., Castles, A., and Crain, S. (2013). Lateralized auditory brain function in children with normal reading ability and in children with dyslexia. *Neuropsychologia, 51*(4), 633-641.

Jorgens, S., Biermann-Ruben, K., Kurz, M. W., Flugel, C., Daehli Kurz, K., Antke, C., et al. (2008). Word deafness as a cortical auditory processing deficit: a case report with MEG. *Neurocase, 14*(4), 307-316.

Joris, P., Schreiner, C., and Rees, A. (2004). Neural processing of amplitude-modulated sounds. *Physiol. Rev., 84*(2), 541-577.

Kraus, N., McGee, T., Carrell, T., Sharma, A., Micco, A., and Nicol, T. (1993). Speech-evoked cortical potentials in children. *J. Am. Acad. Audiol., 4*(4), 238-248.

Lehongre, K., Ramus, F., Villiermet, N., Schwartz, D., and Giraud, A. L. (2011). Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. *Neuron, 72*(6), 1080-1090.

Lehtela, L., Salmelin, R., and Hari, R. (1997). Evidence for reactive magnetic 10-Hz rhythm in the human auditory cortex. *Neurosci. Lett.*, *222*(2), 111–114.

Lippe, S., Martinez-Montes, E., Arcand, C., and Lassonde, M. (2009). Electrophysiological study of auditory development. *Neuroscience, 164*(3), 1108-1118.

Lu, T., Liang, L., and Wang, X. (2001). Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. *Nat. Neurosci.*, *4*(11), 1131-1138.

Marshall, P. J., Bar-Haim, Y., and Fox, N. A. (2002). Development of the EEG from 5 months to 4 years of age. *Clin. Neurophysiol., 113*(8), 1199-1208.

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

Mehler, J., and Christophe, A. (1995). Maturation and learning of language in the first year of life. *The Cognitive Neurosciences* (pp. 943-954). Cambridge, MA: The MIT Press.

Menell, P., McAnally, K. I., and Stein, J. F. (1999). Psychophysical sensitivity and physiological response to amplitude modulation in adult dyslexic listeners. *J. Speech Lang. Hear. Res., 42*, 797-803.

Miyazaki, T., Thompson, J., Fujioka, T., and Ross, B. (2013). Sound envelope encoding in the auditory cortex revealed by neuromagnetic responses in the theta to gamma frequency bands. *Brain Res., 1506*, 64-75.

Moore, D. R. (2002). Auditory development and the role of experience. *Br. Med. Bull., 63*, 171-181.

Moore, D. R., Cowan, J. A., Riley, A., Edmondson-Jones, A. M., and Ferguson, M. A. (2011). Development of auditory processing in 6 to 11 yr old children. *Ear Hear., 32*, 269–285.

Moore, J. K., and Linthicum, F. H. (2007). The human auditory system: a timeline of development. *Int. J. Audiol., 46*(9), 460-478.

Nakasatp, N., Levesque, M. F., Barth, D. S., Baumgartner, C., Rogers, R. L., and Sutherling, W. W. (1994). Comparisons of MEG, EEG, and ECoG source localization in neocortical partial epilepsy in humans. *Electroencephalogr. Clin. Neurophysiol., 91*(3), 171-178.

Nodarse, E. M., Abalo, M. C. P., Fortuny, A. T., Hernández, M. V., and Castellanos, A. L. (2012). Maturational changes in the human envelope-following responses. *Acta Otorrinolaringol. Esp.*, *63*(4), 258-264.

Niedermeyer, E. (1990). Alpha-like rhythmical activity of the temporal lobe, *Clin. Electroencephalogr.*, *21*, 210–224.

Niedermeyer, E. (1991). The 'third' rhythm': further observations, *Clin. Electroencephalogr., 22*, 83–96.

Paetau, R., Ahonen, A., Salonen, O., and Sams, M. (1995). Auditory evoked magnetic fields to tones and pseudowords in healthy children and adults. *J. Clin. Neurophysiol., 12*(2), 177-185.

Parkkonen, L., Fujiki, N., and Makela, J. P. (2009). Sources of auditory brainstem responses revisited: contribution by magnetoencephalography. *Hum. Brain Mapp., 30*(6), 1772-1782.

Peelle, J. E., and Davis, M. H. (2012). Neural oscillations carry speech rhythm through to comprehension. *Front. Psychol., 3(*320), 1-17.

Phillips, D. P., and Farmer, M. E. (1990). Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. *Behav. Brain Res., 40*(2), 85-94.

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int. J. Audiol., 42*(4), 177-219.

Plante, E., Holland, S. K., and Schmithorst, V. J. (2006). Prosodic processing by children: an fMRI study. Brain Lang., 97(3), 332-342.

Poelmans H, Luts H, Vandermosten M, Boets B, Ghesquière P, Wouters J (2012) Auditory steady state cortical responses indicate deviant phonemic-rate processing in adults with dyslexia. *Ear Hear*, *33*, 134-143.

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. Speech Commun., 41(1), 245-255.

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

Poulsen, C., Picton, T. W., and Paus, T. (2007). Age-related changes in transient and oscillatory brain responses to auditory stimulation in healthy adults 19-45 years old. *Cereb. Cortex, 17*(6), 1454-1467.

Poulsen, C., Picton, T. W., and Paus, T. (2009). Age-related changes in transient and oscillatory brain responses to auditory stimulation during early adolescence. *Dev. Sci., 12*(2), 220-235.

Purcell, D. W., John, S. M., Schneider, B. A., and Picton, T. W. (2004). Human temporal auditory acuity as assessed by envelope following responses. *J. Acoust. Soc. Am., 116*, 3581-3593.

Putter-Katz, H., Kishon-Rabin, L., Sachartov, E., Shabtai, E. L., Sadeh, M., Weiz, R., et al., (2005). Cortical activity of children with dyslexia during natural speech processing: evidence of auditory processing deficiency. *J. Basic. Clin. Physiol. Pharmacol., 16*(2-3), 157-172.

Ramus, F., Nespor, M., and Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition, 73*, 265-292.

Regan, D. (1989). Human Brain Electrophysiology: Evoked Potentials and Evoked Magnetic Fields in Science and Medicine. Amsterdam: Elsevier.

Rosen, S. (1992). Temporal information in speech: Acoustic, auditory and linguistic aspects. *Proc. R. Soc. Lond. A. Math. Phys. Sci., 336*(1278), 367-373.

Saby, J. N., and Marshall, P. J. (2012). The utility of EEG band power analysis in the study of infancy and early childhood. *Dev. Neuropsychol., 37*(3), 253-273.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*(5234), 303-304.

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. *Electroencephalogr. Clin. Neurophysiol., 104*(6), 540-545.

Souza, P. E. (2000). Older listeners' use of temporal cues altered by compression Amplification. *J. Speech Lang. Hear. Res., 43*, 661-674.

Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature, 416*, 87-90.

Srinivasan, R. (1999). Spatial structure of the human alpha rhythm: global correlation in adults and local correlation in children. *Clin. Neurophysiol., 110*(8), 1351–1362.

Studebaker, G. A. (1985). A rationalized arcsine transform. *J. Speech Lang. Hear. Res., 28*(3), 455-462.

Sussman, E., Steinschneider, M., Gumenyuk, V., Grushko, J., and Lawson, K. (2008). The maturation of human evoked brain potentials to sounds presented at different stimulus rates. Hear. Res., 236(1–2), 61-79.

Tervaniemi M, Hugdahl K (2003) Lateralization of auditory-cortex functions. *Brain Res Rev*, *43*, 231-246.

Tlumak, A. I., Durrant, J. D., Delgado, R. E., and Boston, J. R. (2011). Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: profile in adults. *Int. J. Audiol., 50*(7), 448-458.

Tlumak, A. I., Durrant, J. D., Delgado, R. E., and Boston, J. R. (2012). Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: Profile in children vs. adults. *Int. J. Audiol., 51*(6), 480-490.

Trehub, S. E., Schneider, B. A., and Henderson, J. L. (1995). Gap detection in infants, children, and adults. *J. Acoust. Soc. Am., 98*(5), 2532-2541.

Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., and Singer, W. (2010). Neural synchrony and the development of cortical networks. *Trends Cogn. Sci.,* 14(2), 72–80.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. J. Acoust. Soc. Am., 66(5), 1364-1380.

Walker, M. M., Shinn, J. B., Cranford, J. L., and Givens, G. D. (2002). Auditory temporal processing performance of young adults with reading disorders. *J. Speech Lang. Hear. Res.*, *45*, 598–605.

Wartenburger, I., Steinbrink, J., Telkemeyer, S., Friedrich, M., Friederici, A. D., and Obrig, H. (2007). The processing of prosody: Evidence of interhemispheric specialization at the age of four. *NeuroImage, 34*(1), 416-425.

Weisz, N. (2011). Alpha rhythms in audition: cognitive and clinical perspectives, *Front. Psychol., 2*, 1–15.

Wightman, F., Allen, P., Dolan, T., Kistler, D., and Jamieson, D. (1989). Temporal resolution in children. *Child Dev., 60*(3), 611-624.

Whitford, T. J., Rennie, C. J., Grieve, S. M., Clark, C. R., Gordon, E., and Williams, L. M. (2007). Brain maturation in adolescence: concurrent changes in neuroanatomy and neurophysiology. *Hum. Brain Mapp., 28*(3), 228-237.

Wunderlich, J. L., and Cone-Wesson, B. K. (2006). Maturation of CAEP in infants and children: A review. *Hear. Res., 212*(1–2), 212-223.

Yordanova, J., and Kolev, V. (2008). Event-related brain oscillations in normal development. . In L. A. Schmidt and S. J. Segalowitz (Eds.), *Developmental Psychophysiology: Theory, Systems and Methods* (pp. 15-68). New York: Cambridge University Press.

Chapter 5 – General Discussion and Conclusions

Final Discussion and Conclusions

This chapter summarizes the general findings of the thesis and their implications for future research.

5.1 Overview of thesis findings

The purpose of this thesis was to investigate the encoding mechanisms associated with temporal modulation in the human auditory cortex both in the mature brain and in the developing brain. More specifically, the thesis was devoted to providing evidence both from a structural perspective and from a developmental perspective for the hypotheses that cortical temporal processing involves a dual encoding scheme, a scheme that includes both a temporal-based representation and a rate-based representation. There were four major findings reported in this thesis:

1) Using the inter-trial phase coherence (ITPC) measure, amplitude modulations at the rate of 1 – 80 Hz elicited envelope following responses (EFRs) with a temporal modulation transfer function (TMTF) profile that closely replicated the typical TMTF profiles of ASSRs/EFRs in previous studies (Figure 3 in chapter 1). However, using the event-related spectral perturbation (ERSP) measure, the TMTF profile of the EFRs showed strong sustained following responses between about 30 Hz and 80 Hz, with no sign of declining.

2) When phase-locked (PL) and non-phase locked (NPL) components of the ERSPs were separated, the PL component revealed a pattern of following responses that were phase-locked to the onset of each modulation. By contrast, the NPL component exhibited a pattern of following responses that were not phase-locked to the onset of modulations. The PL following

responses peaked at around 40 Hz and declined above 50 Hz, whereas the NPL following responses increased nearly monotonically from about 30 Hz onward. Below about 50 Hz, the PL component was dominant; between about 50 Hz and 80 Hz, the NPL following responses were dominant.

3) These dual PL and NPL following responses were validated using EEG recordings and were exhibited only at the cortical level. When the signal was reconstructed in the brainstem, the AMs only elicited PL responses.

4) In the developing brains of preschool children, neither the PL and NPL components of the ERSP were measureable in the current experimental setting. However, using ITPC, the phase concentration of phase-locked responses was revealed. The EFRs in preschool children were mainly limited to AMs with rates below about 25 Hz.

5.2 Cortical processing of AMs in the mature brain

This section discusses the implications of the findings from adults, recorded using MEG and EEG.

5.2.1 Limitations of current views of cortical processing of AMs

Current research in cortical processing of AMs have been largely limited to auditory evoked responses that are strictly phase-locked to the onsets of the stimuli or to the onset of each modulation in periodic stimuli (Ross et al., 2000, Joris et al., 2004, Eggermont and Wang, 2011, Nourski and Brugge, 2011). Stimulus-synchronised neuronal responses are evident and prevalent throughout the auditory system (Joris et al., 2004, Eggermont and Wang, 2011), but the upper limit of stimulus rates at which neurons can become synchronised declines as one ascends in the auditory system (Joris et al., 2004, Wang et al., 2008). Specifically, single unit

studies in the primary auditory cortex of awake marmoset monkeys showed that the stimulus-synchronisation representation is mostly restricted to slow modulations or click trains below about 50 Hz (Lu et al., 2001, Liang et al., 2002). Non invasive EEG and MEG measurements in humans also show a low-pass TMTF with a high cut off at about 50 Hz, reaching a trough at around 70 Hz. (Another amplitude peak in the EEG TMTF at rates of 80 to 110 Hz is attributable to brainstem responses (John and Picton, 2000, Ross et al., 2000, Picton et al., 2003, Purcell et al., 2004, Miyazaki et al., 2013)). These neurophysiological TMTF profiles are in good agreement with the psychophysically measured TMTF for detecting AMs in noise (Viemeister, 1979). Such results have suggested that the higher levels of the auditory system are organised to emphasize low frequency modulation rates in environmental sounds and speech (Drullman et al., 2014); some researchers have further posited that the temporal processing capabilities of the auditory system can be modelled as a simple low pass filter with a high cut-off frequency of about 50 Hz (Viemeister, 1979; see Ross, 2001 for similar MEG findings).

This typical low-pass filter profile with a cut off point at higher rates was replicated in our ITPC measures of the following responses to AMs. This suggests that our rapidly sweeping stimulus, which delivered only about 2 cycles of each modulation rate step, is sufficient for eliciting frequency-specific following responses that exhibit the capacity to track the temporal dynamics of the acoustic signal. Furthermore, the insertion of an inter-trial-interval (ITI) between each sweep allowed us to investigate how the brains encode the rapidly changing temporal dynamics after being reset to a quasi resting state. More importantly, the close replication of the ITPC and previous findings show that our data are comparable to conventional ASSRs/EFRs recorded with a continuous stimulation paradigm with peak phase-locking around 40 Hz and a rapid decline at rates above 50 Hz (Ross et al., 2000, Picton et al., 2003, Purcell et al., 2004, Miyazaki et al., 2013),

While the ITPC closely replicated previous findings in ASSR/EFR studies the ERSP

showed that the *total amplitude* of the response increased through the frequency range of ~30 Hz to 80 Hz (Figure 3 in chapter 2). The discrepancy between the ITPC and the ERSP measures can be attributed to the fact the ITPC solely quantifies phase-locked activity whereas the ERSP quantifies both PL and NPL activity (Makeig, 1993, Tallon-Baudry et al., 1996, Pfurtscheller and Silva, 1999, Tallon-Baudry and Bertrand, 1999, Brugge et al., 2009). In fact, when the PL and NPL components were separated from the ERSP, the PL component showed a TMTF profile similar to that of the ITPC.

The sustained amplitude increase at higher modulation rates indicates that phase-locked responses alone cannot account for the ERSP spectrogram, particularly at higher AM rates. These EEG and MEG results are entirely consistent with evidence from intracranial recordings in nonhuman primates (Lu et al., 2001; Steinschneider et al., 1998) and human patients (Brugge et al., 2009) and support the hypothesis that two distinct types of encoding schemes are employed to process low and high temporal modulation rates in auditory cortex (Lu et al., 2001). Regardless of the underlying mechanisms, these findings underscore the value of applying methods that allow researchers to examine both PL and NPL activity in the brain (Makeig, 1993, Tallon-Baudry and Bertrand, 1999, Nourski and Brugge, 2011). Conventional analyses that emphasise only PL activity provide only a partial and incomplete picture of brain function (see Joris, 2004 for a similar view regarding over-reliance on measures of strongly phase-locked responses in single unit studies).

5.2.2 Dual temporal representations

The distinct TMTF profiles of the PL and NPL components are in keeping with those of synchronised and non-synchronised neuronal populations in primary auditory cortex of marmoset monkeys (see Figure 2 in Chapter 1) (Lu et al., 2001, Liang et al., 2002), as well as the repetition rate dependent transition from phase-locked frequency following responses (FFRs) to non-phase-locked broadband gamma responses that have been found in primary

auditory cortex of macaque monkeys (Steinschneider et al., 1998) and primary auditory cortex of human epilepsy patients (see Figure 4 in Chapter 1) (Brugge et al., 2009, Nourski et al., 2013).

Lu and colleagues reported that temporal modulations at different rates in acoustic signals are processed in the auditory cortex by two different types of neuron populations (viz., synchronised and non synchronised) (Lu et al., 2001, Liang et al., 2002). One type of neuron exhibited a spiking pattern that was strictly synchronised to the onset of acoustic events, i.e., a typical temporal representation. A second type of neuron showed no stimulus-synchronisation, but the firing rate increased monotonically as the repetition rate increased, so this can be considered to be a rate-based representation (Figure 1 in Chapter 1). Examination of the percentages of activity of these two types of neuron populations as a function of the modulation or repetition rate indicated that the synchronised population was dominant when the rate was below about 50 Hz (inter click interval = ~ 20 ms) and the non-synchronised population increased almost monotonically as the rate increased above about 20 Hz (ICI = ~ 50 ms) (Figure 2 in Chapter 1).

In line with the single unit findings, recent ECoG studies in surgical patients has also revealed a clear transition from phase-locked FFRs to non-phase-locked broadband gamma responses in processing click trains (see Figure 4 in Chapter 1) (Brugge et al., 2009, Nourski et al., 2013). When the repetition rate of click-trains was below about 125 Hz, Brugge and colleagues (2009) were able to observe clear FFRs in the event-related band power spectrogram. Although weak FFRs were still visible at higher repetition rates (i.e. 150 Hz and 200 Hz), these were strongly attenuated. The attenuated FFRs were accompanied by an increase in the broadband gamma responses between about 20 Hz to 150 Hz (see Figure 4 in chapter 1). Phase-locked FFR is thought to reflect temporal coding by stimulus-synchronisation; and non-phase-locked gamma band responses are associated with rate-based coding (Nourski and Brugge, 2011).

The PL and NPL components of EFRs found in the current data showed a strikingly similar pattern with that found in the Lu et al. (2001) single unit study, when the two components of EFRs were plotted as a function as the modulation rates (see Figure 3 in Chapter 2). The NPL following responses increased quasi-monotonically as the modulation rate increased from about 30 Hz to 80 Hz, whereas the PL following responses declined as the rate increased from about 40 Hz to 80 Hz. Within the gamma band frequency range (~ 30 Hz to 80 Hz), there was a negative correlation between PL and NPL responses. The transition of a PL dominant following pattern to a NPL dominant following profile occurred around 50 Hz, the same point of transition observed between synchronised and desynchronised populations in the Le et al. (2001) study. Although a direct link between the neuronal activity in single unit and macro brain signals recorded in EEG/MEG needs to be made with caution (Buzsaki et al., 2012), neurophysiological studies on monkeys have shown a positive correlation between gamma power and neuronal spike rates (Steinschneider et al., 2008, Ray and Maunsell, 2011).

5.2.3 Functional significance of temporal-to-rate transformation

The finding of a population of primary auditory cortex neurons that changes firing rate as a function of AM rate without synchronising to modulation frequency (Lu et al., 2001, Liang et al., 2002, Wang et al., 2003), is suggestive of of a progressive transformation from synchrony based temporal coding to rate-based (or rate-place based) coding as information ascends in the auditory system (Eggermont, 2014; Joris, 2004), a transformation that is evident by the level of the inferior colliculus (IC). Figure 1 illustrates the progressive reduction of temporal resolution along the ascending auditory pathway as well as increasing rate-based encoding at higher stations of the auditory system.

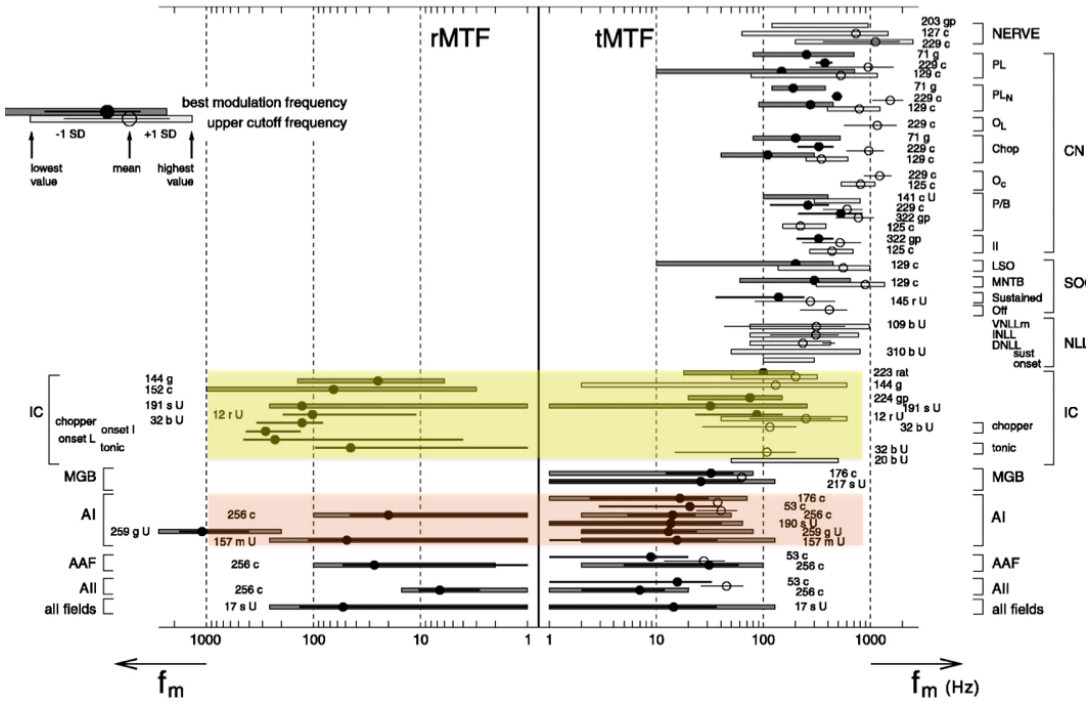

Figure 1 Dual encoding at higher level of the auditory system and the reduction of temporal resolution along the ascending auditory pathway illustrated by an overview of rate-based coding represented by rMTF (left panel) and temporal coding represented by tMTF (right panel) properties at different anatomical levels. Each entry shows means or medians (circles) % SD (lines) and lowest and highest values (bar). Dark bars, thick lines, and solid circles are for rBMFs (left) and tBMFs (right); light bars, lines, and empty circles are for upper tMTF cutoff frequencies (right). For convenience of comparison, the left panel and the right panel are arranged mirror-symmetric. The population measures are taken from published data for one anatomical level, sublevel, or cell class; the numbered reference to the publication is shown next to the data, followed by a letter indicating the species (b, bat; c, cat; g, gerbil; gp, guinea pig; m, marmoset; r, rabbit; s, squirrel monkey), and the letter "U" if unanesthetized. Note that part of the differences between studies reflects differences in the metrics used (in particular upper cutoff, which is often defined as a corner frequency or alternatively as the upper limit of significant phase-locking). The dual encoding in inferior colliculus (IC) and primary auditory cortex (A1) is highlighted in yellow and pink, respectively. Figure courtesy of Joris (2004).

The transformation from largely synchrony based encoding at low levels of the auditory system to largely rate-based encoding at higher levels has important functional implications (Wang et al., 2008). The significance of the rate-based, non-phase-locked neural responses hinges on the fact that these responses represent "processed" information rather than faithful representations of temporal structure in sounds. Considering that higher-level processing requires temporal integration over a time window preceding and following a particular time of

interest, this temporal-to-rate transformation seems necessary at the cortical level of the auditory system, for several reasons. For one, it permits integration of information over time and over different modalities (such as vision, where temporal encoding at the periphery is much slower than audition). This implies that the reduction in the precision of temporal coding along the ascending auditory pathway, and the accompanying temporal-to-rate transformation, are necessary for the integration of rapid auditory information in the cerebral cortex with the intrinsically slower information flow from other sensory modalities.

5.2.4 Perceptual relevance of dual temporal representations

The existence of different neuronal mechanisms for encoding slow and fast temporal modulations accords well with how we actually perceive sounds (Moore et al., 2001). Temporal modulations at different rates are related to different perceptual sensations of non-speech sounds (Eggermont, 2001) and different linguistic units of speech (Rosen, 1992, Leong and Goswami, 2013). Temporal modulations that fluctuate slowly or acoustic events that are repeated with a low frequency (below about 25 Hz) are likely to perceived as discretely occurring individual events. Such slow temporal modulations are associated with the perception of rhythm and the syllabic structure of speech (Poeppel, 2003, Poeppel et al., 2008, Giraud and Poeppel, 2012). According to the dual representation hypothesis, the encoding of temporal fluctuations at these low rates is mostly conducted through a stimulus-synchronised coding scheme with a high fidelity representation of timing (Lu et al., 2001, Liang et al., 2002, Wang et al., 2003), which is critical for identifying a discrete acoustic event or locating the onset of a syllable in the speech stream.

On the other hand, temporal modulations above about 40 Hz are perceived as a continuous sound (Miyazaki et al., 2013). This fits well with a rate-based encoding scheme that captures overall temporal rate but discards the individual identities of events. It has been reported that rapid-varying temporal fluctuations in the range 40 - 500 Hz (Rosen, 1992, Joris et al., 2004, Nourski and Brugge, 2011) are important for the perception of periodicity pitch and

discriminating phonemes in speech (Langner, 1992, Crone et al., 2001, Eggermont, 2001). Between about 25 Hz to 70 Hz, intermediate perceptions like roughness and may be explained by the co-existence of dual neuronal encodings (Langner, 1992, Wang et al., 2008). Taken together, these two fundamentally distinct methods of encoding fast and slow times scales therefore provide a parsimonious neurophysiological explanation for how we actually perceive temporal phenomena (Moore et al., 2001).

5.2.5 Implications for speech perception

It has been shown that the primary auditory cortex has the capacity to differentially encode perceptually significant features of speech, including formant transitions in vowels, and the place and manner of articulation in consonants (Mesgarani et al., 2008). This invites us to ask whether the neural temporal representations found in non-speech sounds in animals and surgical patients, as well as healthy adults reported in the present thesis have implications for human perception of speech. A recent study by Mesgarani and colleagues addressed this question using ECoG data recorded from epilepsy patients as they listened to natural, continuous speech (Mesgarani et al., 2014). Response selectivity to distinct phonetic features was identified in the superior temporal gyrus (STG) and the encoding of acoustic properties was mediated by a distributed population response. These researchers also discovered that phonetic features were directly related to tuning for spectro-temporal acoustic cues, some of which were encoded in a nonlinear fashion or by an integration of multiple cues. Taken together, the findings of the Mesgarani et al. (2014) study demonstrated the acoustic-phonetic representation of speech in the human STG. Given the relevance of encoding spectro-temporal features of phonemes in the auditory cortex, this finding underscores the importance of the current endeavor to understand how temporal features are being encoded at a cortical level.

For neuroscientists, solving the puzzle of ‘speech perception’ involves understanding a list of processes, including the perceptual representations of the acoustic signal, the phonetic

tagging of those representations, accessing the phonological lexicons, and then accessing the semantics of those lexicons (Phillips, 1998). Recent advances in brain imaging using electrophysiological recordings have opened up a window for us to further looking into those processes involving in speech perception in healthy humans. One of the processes attracting intensive attention is how the human auditory cortex processes complex sounds, including the temporal processing involved in the perception of speech. The auditory cortex in humans is a pivotal region of the auditory system which is reciprocally connected with the thalamus, midbrain, and lower brainstem, as well as with more distant cortical areas of frontal and parietal lobes, where higher-level linguistic processing takes place (Wang et al., 2008). Understanding where and how temporal sound features are represented and encoded within the human auditory cortical complex continues to represent a challenge for auditory neuroscience (Nourski and Brugge, 2011). The current work on representations of temporal modulations in the human auditory cortex is only one small piece of the puzzle associated with speech perception. Nonetheless, given the important role of temporal envelope information in speech recognition (Drullman et al., 1994b, a, Drullman, 1995, Shannon et al., 1995, Smith et al., 2002), understanding how the temporal information is encoded in the auditory cortex is no doubt a crucial piece for solving the puzzle.

Indeed, recent neurolinguistic models suggest that the encoding of temporal modulations of speech envelope plays an essential role in speech perception (Giraud and Poeppel, 2012, Peelle et al., 2012, Goswami and Leong, 2013). These models contend (Poeppel, 2003, Zeng et al., 2005, 2013) that there are multiple temporal structures embedded in the speech envelope, with each of these different temporal scales corresponding to a specific linguistic unit (such syllables and phonemes). Figure 2 provides an example of one correspondence between temporal structures and linguistic units. Models such as this provide a comprehensive framework for accommodating many existing results from psychophysical studies of temporal processing in humans. For example, damage to the temporal envelope of

speech by smearing (Drullman et al., 1994b) or by compressing speech in time (Ahissar et al., 2001) causes difficulty in comprehending speech stimuli. More straightforwardly, replacing the spectral content of a speech signal with white noise, but preserving the temporal envelope using a speech-noise-chimaera technique, yields speech recognition that is perceived as accurately as natural speech (Smith et al., 2002).

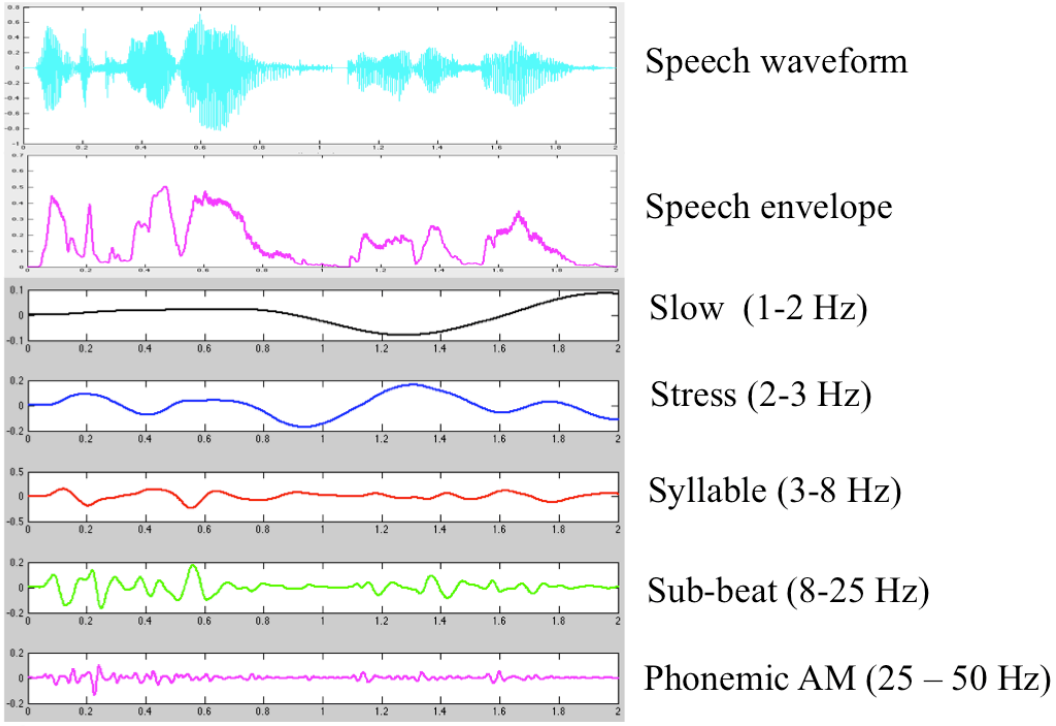

Figure 2. Example of multiple temporal structures of speech envelope. The top panel shows the waveform of a natural speech sentence ('We become blind to what is familiar.'). The second panel is the overall speech envelope extracted from the speech waveform using Hilbert Transform. The bottom five panels present the multiple tiers decomposed from the speech envelope with predefined frequency range for each tier (Smith et al., 2002, Leong, 2012).

In line with the multiple temporal structures in speech is the role of brain oscillations at different frequency bands in processing speech. The correspondence between low frequency syllabic structure and theta oscillations, and between high frequency phonemic structure and gamma oscillations, have been considered to reflect the underlying neural mechanisms (Giraud and Poeppel, 2012, Peelle and Davis, 2012). Regardless of frequency bands, the processing mechanisms that different frequency band oscillations undergo is a manner of phase-locking. This leads to the conclusion that deficits in temporal processing are purely a matter of "time"—or the corresponding frequency (Tallal, 2004, Goswami, 2011, Goswami and Leong, 2013). For

example, the application of such models of speech perception to dyslexia draws on the inference that abnormal temporal sampling of speech results in inaccurate representations of speech features and impaired speech processing. It remains unclear, however, what sampling rate is specifically affected and how this might impair phonological representations in dyslexia. A number of authors have suggested that the perceptual and phonological problems associated with dyslexia are specifically associated with impaired oscillatory processing in the slow theta range (Abrams et al., 2009, Goswami, 2011). In the temporal sampling framework proposed by Goswami and colleagues (2011), impaired phase-locking of theta rhythms results in impaired syllable parsing and difficulties in perceiving the components of syllables. On the other hand, others have emphasized problems in temporal sampling at higher frequency gamma band rates, resulting in abnormal representations of phonemes (Lehongre et al., 2011). Moreover, children with specific language impairment have been reported to exhibit a deficit in processing rapidly presented acoustic information, in the range of tens of milliseconds, which covers the phonemic level of speech perception and production (Tallal et al., 1993, Tallal, 2004). These conflicting results are difficult to reconcile in current theoretical models that focus on phase-locked neuronal responses.

More importantly, the biggest challenge these models are facing is the current view about the limited capacity of auditory cortex in encoding rapid changing temporal information, which is crucial for the perception of phonemes (Poeppel et al., 2008, Giraud and Poeppel, 2012). On other hand, a strictly phase-locking encoding scheme is not just neurophysiologically unavailable (i.e. limited resolution of temporal code in the cortex), but also linguistically unnecessary. For example, in phonetics, consonants are not acoustically independent from a following vowel and there is no single point at which consonant-vowel (CV) syllables can be broken apart into discrete phonological units as one is processing the continuous acoustic signal (see Figure 3 for an example). CV sequences are inextricably realized as single, dynamic acoustic entities. Therefore, at the level of phonetic structure, the exact timing of the

modulation onset is not clear-cut. The transformation of a temporal-based code to a rate-based code is therefore likely to be relevant for effective linguistic perception, especially in view of the fact that a rate-based code provides a more faithful representation of the acoustics.

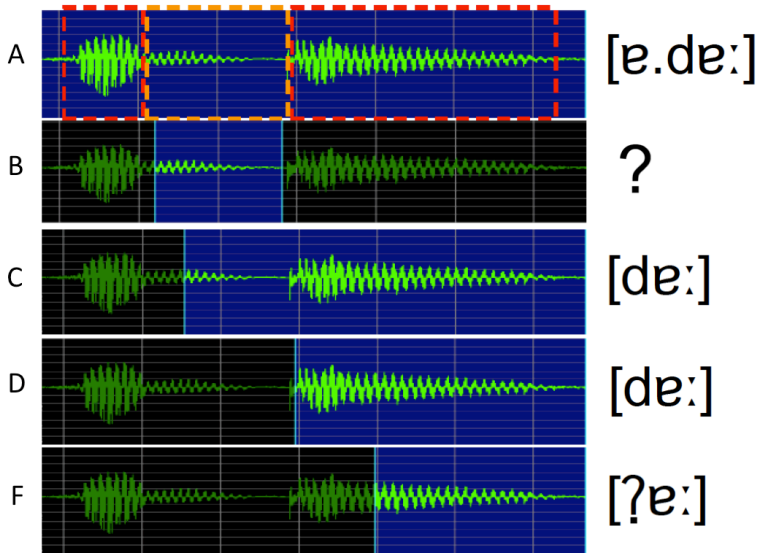

Figure 3. Limitations of speech segmentation. (A) Acoustic waveform of a VCV syllable. (B) The highlighted part shows the acoustic waveform of consonant 'd'. But a 'd' is not perceived. (C) and (D) The highlighted part shows the acoustic waveform of a perceptual prominent CV syllable. (F) Listen to the highlight part forms a perception between the CV syllable and the single vowel in isolation (Delattre et al., 1955, Liberman et al., 1967).

The finding that there exist dual synchronisation-based and non-synchronisation-based representations of temporal modulations therefore offers a potential role for the non-phase-locked encoding mechanism in speech perception. While the range of modulation rates does not comprehensively cover all the linguistically relevant temporal features, it includes low AM frequencies (from 1 to 30 Hz) that are essential for the perception of manner of articulation and for the syllabic rate, as well as for the high AM frequencies (above 50 Hz) that are essential for the perception of voicing and prosody (Rosen, 1992, Giraud et al., 2000, Nourski and Brugge, 2011). Whereas the PL plays a more prominent role in the analysis of speech in the lower frequency ranges, such as in the theta frequency band, the NPL comes to play a more prominent role in the higher frequency range, such as in the gamma frequency band. These dual representations therefore permit different features of the speech stream to be processed in

parallel at different time scales. It is worth noting in this regard that the dual temporal-based and rate-based representations are consistent with the theta-gamma nesting model in which the amplitude of high frequency gamma oscillations is modulated by the phase of low frequency theta oscillations (Giraud and Poeppel, 2012).

The different language deficits that have been suggested to result from difficulties in processing information at different frequency ranges might actually reflect deficits in a different type of encoding, instead of be purely impairments of processing at specific frequency bands. This proposal is consistent with the functional implications that follow from the dual PL and NPL representations, placing significance on the fact that the rate-based NPL neural responses represent processed instead of preserved temporal information. The prominence of NPL cortical processing indicates that the processing of continuous acoustic signals is conducted on a "segment-by-segment" basis rather than "moment-by-moment", which is exactly how speech is parsed in higher-level processing. This feature of higher-level processing follows from the observation that it requires temporal integration over a time window preceding and following a particular time of interest. To fulfil the segment-based partitioning of speech, it is apparently necessary to engage a temporal-to-rate transformation at the cortical level of the auditory system. In other words, the reduction in the temporal limit on stimulus-synchronized encoding in the auditory cortex is a prerequisite for successful segmentation of speech and therefore crucial for later phonetic mapping.

5.3 Cortical processing of AMs in the developing brain

Chapter 4 reported the envelope following responses in normal hearing preschool-aged children, as measured using a pediatric MEG system. The data revealed EFR phase-locking in children's auditory cortex at lower AM rates but this dropped off quickly at rates above about 15 Hz. Above about 25 Hz, phase-locking was no longer detected in the present experimental settings. These results are consistent with behavioral evidence of limited temporal processing

capabilities in children, as obtained in psychophysiological studies (Irwin et al., 1985, Wightman et al., 1989, Trehub et al., 1995, Banai et al., 2011, Moore et al., 2011, Fox et al., 2012, Buss et al., 2014).

5.3.1 Underdeveloped temporal processing

In the current data, the immature phase-locking profile of EFRs in children exhibited a bias towards slow modulations. Comparing our results to those from previous studies in children, we found that, in the absence of acoustic stimulation, the slow oscillatory activity of children's auditory cortex demonstrated higher power than that of adults. Figure 2 in Chapter 4 summarizes the power spectrogram of these children's quasi-resting brain oscillations, which is consistent with previous EEG findings (Marshall et al., 2002; Saby and Marshall, 2012; Yordanova, 2008).

Previous electrophysiological studies of children's temporal processing has largely focused on the high frequency ASSR, e.g., 80 Hz (Levi et al., 1995). This might not be as informative as possible about the cortical processing in children, for two reasons: First, the ASSRs at high frequencies found in infants and young children are most likely generated in the subcortical structures such as the brainstem (Moore, 2002). Second, as compared with high frequency brain oscillations, the slow waves in the ongoing brain signals were more prominent in children. The bias of slow modulations in children, as uncovered in the present work, implies that auditory processing in children is more likely to be accomplished by slow oscillations rather than by oscillations at higher frequencies. This is also consistent with the observation that high-frequency gamma-band power increases with age (Tierney et al., 2013).

Previous EEG studies of the ASSR/EFR in children have reported a response that is at least partly and perhaps largely generated by subcortical sources (Purcell et al., 2004). MEG is considered to be relatively insensitive to such subcortical sources, due to the greater dependence of MEG on distance from the measuring sensors, as compared to EEG (Baillet et

al., 2001; Johnson et al., 2010; Nakasatp et al., 1994). While it is in principle possible to detect brainstem signals with MEG, this usually requires extraordinary efforts (e.g., thousands of trials) to achieve the signal-to-ratio (SNR) that is required to extract subcortical responses such as wave V of the auditory brainstem response (ABR; Parkkonen et al., 2009). The present MEG results therefore represent responses that originate mainly or entirely from the auditory cortex, with little or no contribution from subcortical regions. The ability to focus our measurements specifically on cortical activity is important because the subcortical and cortical components of the auditory system have quite different temporal encoding capabilities in the mature brain. Single unit measurements show a progressive decrease in the upper limit of synchronizing via phase-locking to the periodicity of a sound at higher levels of the auditory system, from about 4 kHz at the level of the auditory nerve to about 50 Hz at the level of primary auditory cortex (Joris et al., 2004). A further complication is that subcortical and cortical levels of the auditory system have quite different maturational trajectories (Moore, 2002; Moore and Linthicum, 2007).

Our results indicate that the maturing auditory cortex in children has a limited capacity to encode and represent sound periodicities relative to the adult brain, with a sharp drop-off in phase-locking at rates above 15 Hz, with no measurable following response to AMs faster than about 25 Hz. This indicates that sound envelope encoding, like other auditory cortical responses (Kraus et al., 1993; Moore, 2002; Moore and Linthicum, 2007; Ponton et al., 2000; Sharma et al., 1997; Sussman et al., 2008; Wunderlich and Cone-Wesson, 2006) has a prolonged maturational trajectory, with earlier development of responses to slow envelope periodicities and much later development of responses to high frequency modulations.

5.3.2 Implications for language acquisition

The capacity to process temporal modulations, especially modulation rates below about 20 Hz, is essential for speech perception and is presumably an important component of language acquisition. The phonological bootstrapping hypothesis for language acquisition posits that

phrasal prosodic cues are crucial for infants and young children to segment continuous speech in order to form the kinds of pre-lexical representations that are essential for early word learning (Christophe and Dupoux, 1996, Ramus et al., 1999, Christophe et al., 2003, Leong et al., 2014). For infants who are not equipped with lexical knowledge, a non-lexical method of identifying word boundaries is an essential component for the acquisition of the words of the local language (Christophe and Dupoux, 1996, Ramus et al., 1999, Christophe et al., 2003). Prosody or rhythm, which plays an important role in speech processing and phonological bootstrapping, is therefore an essential ingredient of language acquisition (Mehler and Christophe, 1995). Our finding of a strong preference in children's brains for slow modulations is consistent with previous findings showing that infants and children are dependent on slow prosodic cues in segmenting speech. Analyses of speech temporal structures have shown that infant-directed speech is primarily stress-dominant, whereas adult-directed speech is primarily syllable-dominant (Leong et al., 2014). The stress-based dominance of infant-directed speech could reflect that the infant brain is more likely to tune to slower speech rhythms or cues in early language learning. In addition to the slow modulation preference, EFRs in children also showed a right hemisphere laterality in the present study. This is consistent with the previous finding that prosodic processing is preferentially processed in the right hemisphere in children (Friederici and Alter, 2004, Gandour et al., 2004, Homae et al., 2006, Wartenburger et al., 2007).

5.3.3 Implications for disorders in language development

Recent neurolinguistic models contend that the neural mechanisms for encoding temporal modulations of the speech envelope play an essential role in speech perception (Giraud and Poeppel, 2012, Peelle and Davis, 2012). Application of such models of speech perception to dyslexia draws on the logic that abnormal temporal sampling of speech will result in inaccurate representations of speech features and impaired speech processing. However it remains unclear what sampling rate is specifically affected and how this might impair phonological

representations in dyslexia. A number of authors have suggested that the perceptual and phonological problems associated with dyslexia are specifically associated with impaired oscillatory processing in the slow theta range (Abrams et al., 2009, Goswami, 2011). In the temporal sampling framework proposed by Goswami and colleagues (Goswami, 2011), impaired phase-locking of theta rhythms results in impaired syllable parsing and difficulties in perceiving the components of syllables. On the other hand, others have emphasized problems in temporal sampling at higher frequency rates (e.g., the gamma band), resulting in abnormal representations of phonemes (Lehongre et al., 2011). However, the temporal sampling models described above are presently based entirely on neurophysiological data obtained from adults. The present study contributes to this issue by providing evidence that children show a quite different cortical phase locking profile than adults do: the maturing human auditory cortex emphasized low AM rates and, in contrast to the adult reference group, showed no measureable phase-locking at rates higher than about 25 Hz. Temporal sampling models of language perception in general must accommodate the fact that children's temporal processing capabilities are sharply limited, as compared to those of adults.

5.4 Limitations

5.4.1 Energy leaking issue

David and colleagues (David et al., 2006) demonstrated that the subtraction-based method to separate evoked (PL) and induced (NPL) power could be problematic due to the effect of latency jitter using computational simulation. Normally, the latency jitter is small and negligible when only low frequencies are considered. However, small latency jitter can cause evoked responses appear as induced power at higher frequencies. The un-detected "evoked" energy would leak into the "induced" power spectrogram if a simple subtraction method was used to compute the induce response. Moreover, this energy leaking would become larger for higher frequencies given the same latency jitter. Since our NPL ERSP showed an increase of

magnitude as the frequencies increased from about 30 Hz to 80 Hz, and the PL and NPL components are negative correlated. This fits well into the prediction of energy leaking hypothesis. Therefore, a critical question here is whether the NPL following responses can be partially or even completely attribute to the energy leaking from PL power spectrogram to the NPL power spectrogram? This question is technically challenging due to the low precision in estimating the latency for each frequency in an AM sweep paradigm. Therefore, it is difficult to precisely quantify the amount of energy leaking at each frequency caused by the current time-frequency analysis algorithm.

That said, the absence of NPL in the brainstem when PL could be clearly observed, at least suggests that the energy leaking alone is unlikely to explain the NPL profile found in the auditory cortex (Chapter 3). More importantly, if the NPL can be attributed to the energy leaking from PL, that means the actual (detected + undetected) phase-locked following responses would have a profile similar to the total ERSP, which shows a sustained following responses through the higher frequencies tested (40 through to 80 Hz). This sustained phase-locked following responses would mean a high temporal resolution up to about 80 Hz for cortical processing of temporal modulations, which is very unlikely according the converging evidence from electrophysiological studies in humans (Ross et al., 2000, Picton et al., 2003, Tlumak et al., 2011, Nodarse et al., 2012, Miyazaki et al., 2013) as well as single unit recordings in animals (Lu et al., 2001, Joris et al., 2004).

We emphasize that this mis-estimation problem *does not* apply to the total ERSP power computations used here (Steinschneider et al., 1998; Brugge et al., 2013). Therefore, *regardless* of whether our ERSP results are attributable to energy leaking, or to the dual contributions of PL and NPL responses, they do demonstrate that our current understanding of the auditory ASSR/EFR is incomplete because it is based analytic methods that emphasize phase-locking and essentially discard brain activity that is not tightly phase-locked to the periodicity of the acoustic stimulus.

5.4.2 Source modelling

In the current studies, a bilateral equivalent current dipole (ECD) model was used for reconstructing the source waveforms, on which time-frequency analyses were performed. To locate the bilateral dipoles, dipole fitting was conducted based on the identified M100 response on single subject data. The use of a M100 as a locator for source waveform reconstruction for examining EFRs lies on the assumption that both AEF and EFR are generated in the core auditory region and spatially overlapped with each other (Herdman et al., 2003). We would point out that several studies have indicated that the neural sources of the ASSRs to AM sounds, especially at the low frequencies, might originate from many cortical networks, including both primary and non-primary auditory cortex (Liegeois-Chauvel et al., 2004, Gourevitch et al., 2008). Moreover, Lütkenhner and Mosher (Lütkenhöner and Mosher, 2006) showed that the equivalent current dipoles of the low-rate ASSR were located 5–10 mm more medially than the ECD of M100, which is consistent with ASSRs in the 40 Hz range reported in several MEG studies (Draganova et al., 2002, Herdman et al., 2003, Ross et al., 2005). Therefore, using the M100 as a locator for estimating the EFRs sources in the current study might cause a lower than optimal signal-to-noise for exacting EFRs. Nonetheless, given the spatial resolution of MEG, it is difficult to separate those neural sources very precisely, and the dipole location of a neural response is usually interpreted as the center of all its sources (Wang et al., 2012). Hence, despite the potential issues, it is reasonable to adopt such a method for source estimation.

On the other hand, using the same ECD for the EFRs to all the AM rates might not be the optimal model given the ASSRs/EFRs to different modulation rates could be different. However, a fMRI study by Giraud and colleagues (Giraud et al., 2000) found distinct regions were activated during AM sounds, but at the cortical level, both high and low AM frequencies elicited different responses at the same cortical neural substrates. Recent MEG study by Wang

and colleagues (Wang et al., 2012) also reported that the ASSR ECD locations evoked by AM with different modulation rates between 3.5 Hz to 31.5 Hz were not distinguishable.

5.5 Future research questions

There are several ways in which the current thesis work on the temporal encoding of the sound envelope can be extended.

5.5.1 Can the NPL component of EFRs be observed in isolation?

The rationale for this question lies under this question is that if NPL reflects the rate-based coding of temporal modulations by the non-synchronized neuron populations and this rate-based coding become more prominent in higher rates while temporal codes become absent above about 100 Hz (Wang et al., 2003, Wang, 2007), then using a AM sounds with modulation rates between 100 Hz to 200 Hz might allow us to observe the NPL following responses in the absence of PL component. This would not just provide further support for the dual temporal representations of AMs, but also evidence for the temporal-to-rate transformation in the core regions. In line with this, using a fixed rate AM sounds instead of a AM sweep would allow us to perform source localization directly on the ASSRs without using M100 as a locator. More accurate source estimation might then help to identify both the core and belt regions for EFRs. The PL and NPL components in the belt regions should show that NPL component dominated the processing of AMs. This can be directly predicted from the observation of mix-mode neurons in the core region, which indicates the temporal-to-rate transformation has not completed and the further transformation needs to take place in the belt region (Wang et al., 2003, Yin et al., 2011).

5.5.2 Are dual representations relevant to speech perception?

Amplitude modulations deliver temporal cues for many linguistic features crucial for speech perception (Rosen, 1992, Drullman et al., 1994a, Luo and Poeppel, 2007, Doelling et al.,

2014). On the other hand, the primary auditory cortex has been found encoding perceptually significant features, including formant patterns in vowels and place and manner of articulation in consonants (Mesgarani et al., 2008). Therefore the important question needs to address in future studies is whether these neural temporal representations found in non speech sounds can account for human perception of speech.

5.5.3 When do the dual representations converge on the adult form?

In the current thesis, we found that children did not show measureable ERSP PL and NPL components in their following responses. But this finding needs to be interpreted with caution. Although the ITPC measure revealed a phase-locked following profile in children, the ERSP measure showed no clear sign of envelope tracking. The absence of following pattern in ERSP is mostly likely due to the poor signal-to-noise ratio (SNR) in child's data. Future studies need to endeavor to increase the SNR using a different paradigm, such as ASSRs to fixed rate AM sounds. Also, the current study only examined children at preschool age, a pool of children across a span of ages – from preschooler to teenagers would be important to address the question whether the dual representations can be observed in the developing brain. If yes, when do these representations emerge and reach adult form.

5.6 Conclusions

Most neurons in auditory cortex are limited in their capability to precisely synchronise to temporal modulation at rates faster than about 50 Hz. Hence, a central question in auditory neurophysiology concerns how the full range of perceptually relevant modulation rates might be encoded in the cerebral cortex. This thesis shows with that the human auditory cortex transitions between a phase-locked (PL) mode of responding to modulation rates below about 50 Hz, and a non phase-locked (NPL) mode at higher rates. Precisely such dual response modes are predictable from the behaviors of single neurons in auditory cortices of non-human

primates. Our data point to a common mechanistic explanation for the single neuron and MEG results and support the hypothesis that two distinct types of neuronal encoding mechanisms are employed by the auditory cortex to represent a wide range of temporal modulation rates. This dual encoding model allows slow and fast modulations in sounds to be processed in parallel, an important requirement of theoretical frameworks that posit a multiplexed sampling of different time scales in speech.

5.7 References

Abrams DA, Nicol T, Zecker S, Kraus N (2009) Abnormal cortical processing of the syllable rate of speech in poor readers. J Neurosci 29:7686-7693.

Ahissar E, Nagarajan S, Ahissar M, Protopapas A, Mahncke H, Merzenich MM (2001) Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc Natl Acad Sci* 98:13367-13372.

Banai K, Sabin AT, Wright BA (2011) Separable developmental trajectories for the abilities to detect auditory amplitude and frequency modulation. *Hear Res* 280:219-227.

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA (2009) Coding of repetitive transients by auditory cortex on Heschl's gyrus. J Neurophysiol **102**:2358-2374.

Buss E, Hall IIIJW, Porter H, Grose JH (2014) Gap Detection in School-Age Children and Adults: Effects of Inherent Envelope Modulation and the Availability of Cues Across Frequency. *J Speech Lang Hear Res* 57:1098-1107.

Buzsaki G, Anastassiou CA, Koch C (2012) The origin of extracellular fields and currents - EEG, ECoG, LFP and spikes. Nat Rev Neurosci 13:407-420.

Christophe A, Dupoux E (1996) Bootstrapping lexical acquisitionn: The role of prosodic structure *Linguist Rev* 13:383-412.

Christophe A, Nespor M, Guasti MT, Ooyen BV (2003) Prosodic structure and syntactic acquisition: the case of the head-direction parameter. Dev Sci 6:211-220.

Crone NE, Boatman D, Gordon B, Hao L (2001) Induced electrocorticographic gamma activity during auditory perception. Clin Neurophysiol **112**: 565-582.

David O, Kilner JM, Friston KJ (2006) Mechanisms of evoked and induced responses in MEG/EEG. *Neuroimage* **31**:1580-1591.

Delattre PC, Liberman AM, Cooper FS (1955) Acoustic loci and transitional cues for consonants. *J Acoust Soc Am* 27:769-774.

Doelling KB, Arnal LH, Ghitza O, Poeppel D (2014) Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *NeuroImage* 85 Pt 2:761-768.

Draganova R, Ross B, Borgmann C, Pantev C (2002) Auditory cortical response patterns to multiple rhythms of AM sound. Ear Hear 23:254-265.

Drullman R (1995) Temporal envelope and fine structure cues for speech intelligibility. *J Acoust Soc Am* 97:585-592.

Drullman R, Festen JM, Plomp R (1994a) Effect of reducing slow temporal modulations on speech reception. *J Acoust Soc Am* 95:2670-2680.

Drullman R, Festen JM, Plomp R (1994b) Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 95:1053-1064.

Eggermont JJ (2001) Between sound and perception: reviewing the search for a neural code. *Hear Res* **157**:1-42.

Eggermont JJ, Wang X (2011) Temporal coding in auditory cortex In: The Auditory Cortex (Winer, J. A. and Schreiner, C. E., eds), pp 309-328: Springer US.

Fox AM, Reid CL, Anderson M, Richardson C, Bishop DV (2012) Maturation of rapid auditory temporal processing and subsequent nonword repetition performance in children. Dev Sci 15:204-211.

Friederici AD, Alter K (2004) Lateralization of auditory language functions: A dynamic dual pathway model. *Brain Lang* 89:267-276.

Gandour J, Tong Y, Wong D, Talavage T, Dzemidzic M, Xu Y, Li X, Lowe M (2004) Hemispheric roles in the perception of speech prosody. *NeuroImage* 23:344-357.

Giraud A-L, Lorenzi C, Ashburner J, Wable J, Johnsrude I, Frackowiak R, Kleinschmidt A (2000) Representation of the temporal envelope of sounds in the human brain. J Neurophysiol 84:1588-1598.

Giraud AL, Poeppel D (2012) Cortical oscillations and speech processing: emerging computational principles and operations. Nat Neurosci **15**:511-517.

Goswami U (2011) A temporal sampling framework for developmental dyslexia. Trends Cogn Sci 15:3-10.

Goswami U, Leong V (2013) Speech rhythm and temporal structure: Converging perspectives? *Lab Phonol* 4:67-92.

Gourevitch B, Le Bouquin Jeannes R, Faucon G, Liegeois-Chauvel C (2008) Temporal envelope processing in the human auditory cortex: response and interconnections of auditory cortical areas. *Hear Res* 237:1-18.

Herdman AT, Wollbrink A, Chau W, Ishii R, Ross B, Pantev C (2003) Determination of activation areas in the human auditory cortex by means of synthetic aperture magnetometry. *NeuroImage* 20:995-1005.

Homae F, Watanabe H, Nakano T, Asakawa K, Taga G (2006) The right hemisphere of sleeping infant perceives sentential prosody. Neuroscience research 54:276-280.

Irwin RJ, Ball AKR, Kay N, Stillman JA, Rosser J (1985) The Development of Auditory Temporal Acuity in Children. *Child Dev* 56:614-620.

John MS, Picton TW (2000) Human auditory steady-state responses to amplitude-modulated tones: phase and latency measurements. *Hear Res* **141**:57-79.

Joris P, Schreiner C, Rees A (2004) Neural processing of amplitude-modulated sounds. *Physiol Rev* **84**:541-577.

Langner G (1992) Periodicity coding in the auditory system. *Hear Res* 60:115–142.

Lehongre K, Ramus F, Villiermet N, Schwartz D, Giraud AL (2011) Altered low-gamma sampling in auditory cortex accounts for the three main facets of dyslexia. Neuron **72**:1080-1090.

Leong V (2012) Prosodic Rhythm in the Speech Amplitude Envelope : Amplitude Modulation Phase Hierarchies (AMPHs) and AMPH Models. vol. Doctor of Philosophy.

Leong V, Goswami U (2013) Assessment of rhythmic entrainment at multiple timescales in dyslexia: Evidence for disruption to syllable timing. *Hear Res* 1-21.

Leong V, Kalashnikova M, Burnham D, Goswami U (2014) Infant-directed speech enhances temporal rhythmic structure in the envelope. . *Int Speech Commun Assoc* 2563-2567.

Levi EC, Folsom RC, Dobie RA (1995) Coherence analysis of envelope-following responses (EFRs) and frequency-following responses (FFRs) in infants and adults. *Hear Res* 89:21-27.

Liang L, Lu T, Wang X (2002) Neural representations of sinusoidal amplitude and frequency modulations in the primary auditory cortex of awake primates. J Neurophysiol 87:2237-2261.

Liberman AM, Cooper FS, Shankweiler DP, Studdert-Kennedy M (1967) Perception of the speech code *Psychological Review* 74:431-461.

Liegeois-Chauvel C, Lorenzi C, Trebuchon A, Regis J, Chauvel P (2004) Temporal envelope processing in the human left and right auditory cortices. Cereb Cortex 14:731-740.

Lu T, Liang L, Wang X (2001) Temporal and rate representations of time-varying signals in the auditory cortex of awake primates. Nat Neurosci 4:1131-1138.

Luo H, Poeppel D (2007) Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. Neuron 54:1001-1010.

Lütkenhöner B, Mosher J (2006) Source analysis of auditory evoked potentials and fields. Philadelphia: Lippincott Williams and Wilkins.

Makeig S (1993) Auditory event-related dynamics of the EEG spectrum and effects of exposure to tones. *Electroencephalogr Clin Neurophysiol* **86**:283-293.
Mehler J, Christophe A (1995) Maturation and learning of language in the first year of life. In: The Cognitive Neurosciences, pp 943-954 Cambridge, MA, US: The MIT Press.
Mesgarani N, Cheung C, Johnson K, Chang EF (2014) Phonetic feature encoding in human superior temporal gyrus. Science 343:1006-1010.
Mesgarani N, David SV, Fritz JB, Shamma SA (2008) Phoneme representation and classification in primary auditory cortex. *J Acoust Soc Am* 123:899-909.
Miyazaki T, Thompson J, Fujioka T, Ross B (2013) Sound envelope encoding in the auditory cortex revealed by neuromagnetic responses in the theta to gamma frequency bands. Brain Res 1506:64-75.
Moore DR (2002) Auditory development and the role of experience. *Br Med Bull* 63:171-181.
Moore DR, Cowan JA, Riley A, Edmondson-Jones AM, Ferguson MA (2011) Development of Auditory Processing in 6- to 11-Yr-Old Children. Ear Hear 32:269–285.
Moore DR, Schnupp JWH, King AJ (2001) Coding the temporal structure of sounds in auditory cortex. Nat Neurosci **4**:1055-1056.
Nodarse EM, Abalo MaCPr, Fortuny AT, Hernández MV, Castellanos AL (2012) Maturational Changes in the Human Envelope-following Responses. *Acta Otorrinolaringol Esp* 63:258-264.
Nourski KV, Brugge JF (2011) Representation of temporal sound features in the human auditory cortex. Rev Neurosci **22**:187-203.
Nourski KV, Brugge JF, Reale RA, Kovach CK, Oya H, Kawasaki H, Jenison RL, Howard III MA (2013) Coding of repetitive transients by auditory cortex on posterolateral superior temporal gyrus in humans: an intracranial electrophysiology study. J Neurophysiol **109**:1283-1295.
Peelle JE, Davis MH (2012) Neural oscillations carry speech rhythm through to comprehension. Front Psychol 3.
Peelle JE, Gross J, Davis MH (2012) Phase-Locked Responses to Speech in Human Auditory Cortex are Enhanced During Comprehension. Cereb Cortex.
Pfurtscheller G, Silva FHLd (1999) Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin Neurophysiol 110:1842-1857.
Phillips D (1998) Sensory representations, the auditory cortex, and speech perception *Semin Hear* **19**:319-332.
Picton TW, John MS, Dimitrijevic A, Purcell D (2003) Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int J Audiol* **42**:177-219.
Poeppel D (2003) The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun* 41:245-255.
Poeppel D, Idsardi WJ, van Wassenhove V (2008) Speech perception at the interface of neurobiology and linguistics. *Philos Trans R Soc Lond* **363**:1071-1086.
Purcell DW, John SM, Schneider BA, Picton TW (2004) Human temporal auditory acuity as assessed by envelope following responses. *J Acoust Soc Am* **116**:3581-3593.
Ramus F, Nespor M, Mehler J (1999) Correlates of linguistic rhythm in the speech signal. Cognition 73:265-292.
Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. PLoS Biol 9:e1000610.
Rosen S (1992) Temporal Information in Speech: Acoustic, Auditory and Linguistic Aspects. Philos Trans R Soc Lond B Biol Sci 336:367-373.

Ross B, Borgmann C, Draganova R, Roberts LE, Pantev C (2000) A high-precision magnetoencephalographic study of human auditory steady-state responses to amplitude-modulated tones *J Acoust Soc Am* 108:679-691.

Ross B, Herdman A, Pantev C (2005) Right hemispheric laterality of human 40 Hz auditory steady-state responses. Cereb Cortex 15:2029-2039.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M (1995) Speech recognition with primarily temporal cues. Science 270:303-304.

Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. Nature 416.

Steinschneider M, Fishman YI, Arezzo JC (2008) Spectrotemporal analysis of evoked and induced electroencephalographic responses in primary auditory cortex (A1) of the awake monkey. Cereb Cortex 18:610-625.

Tallal P (2004) Improving language and literacy is a matter of time. Nat Neurosci 5:721-728.

Tallal P, Miller S, Fitch RH (1993) Neurobiological basis of speech: a case for the preeminence of temporal processing. Ann N Y Acad Sci 682:27-47.

Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn Sci* 3.

Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J (1996) Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. J Neurosci 16:4240-4249.

Tierney A, Strait DL, O'Connell S, Kraus N (2013) Developmental changes in resting gamma power from age three to adulthood. Clin Neurophysiol 124:1040-1042.

Tlumak AI, Durrant JD, Delgado RE, Boston JR (2011) Steady-state analysis of auditory evoked potentials over a wide range of stimulus repetition rates: profile in adults. *Int J Audiol* 50:448-458.

Trehub SE, Schneider BA, Henderson JL (1995) Gap detection in infants, children, and adults. *J Acoust Soc Am* 98:2532-2541.

Viemeister NF (1979) Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am* 66:1364-1380.

Wang X (2007) Neural coding strategies in auditory cortex. *Hear Res* **229**:81-93.

Wang X, Lu T, Bendor D, Bartlett E (2008) Neural coding of temporal information in auditory thalamus and cortex. *Neurosci* **154**:294-303.

Wang X, Lu T, Liang L (2003) Cortical processing of temporal modulations. *Speech Commun* **41**:107-121.

Wang Y, Ding N, Ahmar N, Xiang J, Poeppel D, Simon JZ (2012) Sensitivity to temporal modulation rate and spectral bandwidth in the human auditory system: MEG evidence. J Neurophysiol 107:2033-2041.

Wartenburger I, Steinbrink J, Telkemeyer S, Friedrich M, Friederici AD, Obrig H (2007) The processing of prosody: Evidence of interhemispheric specialization at the age of four. *NeuroImage* 34:416-425.

Wightman F, Allen P, Dolan T, Kistler D, Jamieson D (1989) Temporal Resolution in Children. *Child Dev* 60:611-624.

Yin P, Johnson JS, O'Connor KN, Sutter ML (2011) Coding of amplitude modulation in primary auditory cortex. J Neurophysiol 105:582-600.

Zeng F-G, Nie K, Stickney GS, Kong Y-Y, Vongphoe M, Bhargave A, Wei C, Cao K (2005) Speech recognition with amplitude and frequency modulations. *Proc Natl Acad Sci* 102:2293-2298.

APPENDIX

1. Ethics Approval

MACQUARIE UNIVERSITY

Blake Johnson <blake.johnson@mq.edu.au>

Approved- Ethics application- Johnson (Ref No: 5201300054)

Ethics Secretariat <ethics.secretariat@mq.edu.au> Thu, Mar 28, 2013 at 8:37 AM
To: Dr Blake Johnson <blake.johnson@mq.edu.au>
Cc: Prof Stephen Crain <stephen.crain@mq.edu.au>, Dr Graciela Tesan <graciela.tesan@mq.edu.au>, Dr Jon Brock <jon.brock@mq.edu.au>, Dr Paul Sowman <paul.sowman@mq.edu.au>, Mr Fabrice Bardy <fabrice.bardy@students.mq.edu.au>, Mr Mehdi Parviz <mehdi.parviz@students.mq.edu.au>, Miss Wei He <wei.he5@students.mq.edu.au>, Mrs Joann Tang <huizhen.tang@students.mq.edu.au>

Dear Dr Johnson

Re: "MEG, EEG and fMRI Studies of adult cognition" (Ethics Ref: 5201300054)

Thank you for your recent correspondence. Your response has addressed the issues raised by the Human Research Ethics Committee and you may now commence your research.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf.

The following personnel are authorised to conduct this research:

Dr Blake Johnson
Dr Graciela Tesan
Dr Jon Brock
Dr Paul Sowman
Miss Wei He
Mr Fabrice Bardy
Mr Mehdi Parviz
Mrs Joann Tang
Prof Stephen Crain

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 28 March 2014
Progress Report 2 Due: 28 March 2015
Progress Report 3 Due: 28 March 2016
Progress Report 4 Due: 28 March 2017
Final Report Due: 28 March 2018

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

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

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

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

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy/

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

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

Please retain a copy of this email as this is your official notification of final ethics approval.

Yours sincerely
Dr Karolyn White
Director of Research Ethics
Chair, Human Research Ethics Committee

Ethics application Reference: 5201300054

From: **Ethics Secretariat** <ethics.secretariat@mq.edu.au>
To: Dr Blake Johnson <blake.johnson@mq.edu.au>
Cc: Prof Stephen Crain <stephen.crain@mq.edu.au>,
Dr Graciela Tesan <graciela.tesan@mq.edu.au>,
Dr Jon Brock <jon.brock@mq.edu.au>,
Dr Paul Sowman <paul.sowman@mq.edu.au>,
Mr Fabrice Bardy <fabrice.bardy@students.mq.edu.au>,
Mr Mehdi Parviz <mehdi.parviz@students.mq.edu.au>,
Miss Wei He <wei.he5@students.mq.edu.au>,
Mrs Joann Tang <huizhen.tang@students.mq.edu.au>
Date: Thu, Mar 28, 2013 at 8:37 AM
Subject: Approved- Ethics application- Johnson (Ref No: 5201300054)
Mailed-by: students.mq.edu.au

Dear Dr Johnson

Re: "MEG, EEG and fMRI Studies of adult cognition" (Ethics Ref: 5201300054)

Thank you for your recent correspondence. Your response has addressed the issues raised by the Human Research Ethics Committee and you may now commence your research.

This research meets the requirements of the National Statement on Ethical Conduct in Human Research (2007). The National Statement is available at the following web site:

http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72.pdf.

The following personnel are authorised to conduct this research:

Dr Blake Johnson
Dr Graciela Tesan
Dr Jon Brock
Dr Paul Sowman
Miss Wei He
Mr Fabrice Bardy
Mr Mehdi Parviz
Mrs Joann Tang
Prof Stephen Crain

NB. STUDENTS: IT IS YOUR RESPONSIBILITY TO KEEP A COPY OF THIS APPROVAL EMAIL TO SUBMIT WITH YOUR THESIS.

Please note the following standard requirements of approval:

1. The approval of this project is conditional upon your continuing compliance with the National Statement on Ethical Conduct in Human Research (2007).

2. Approval will be for a period of five (5) years subject to the provision of annual reports.

Progress Report 1 Due: 28 March 2014
Progress Report 2 Due: 28 March 2015
Progress Report 3 Due: 28 March 2016

Progress Report 4 Due: 28 March 2017
Final Report Due: 28 March 2018

NB. If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report for the project.

Progress reports and Final Reports are available at the following website:

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

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. All amendments to the project must be reviewed and approved by the Committee before implementation. Please complete and submit a Request for Amendment Form available at the following website:

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

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that affect the continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at the following websites:

http://www.mq.edu.au/policy/

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

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide the Macquarie University's Research Grants Management Assistant with a copy of this email as soon as possible. Internal and External funding agencies will not be informed that you have final approval for your project and funds will not be released until the Research Grants Management Assistant has received a copy of this email.

Please retain a copy of this email as this is your official notification of final ethics approval.

Yours sincerely
Dr Karolyn White
Director of Research Ethics
Chair, Human Research Ethics Committee

2. Publication

MACQUARIE University

Joann Tang <joann.tang@mq.edu.au>

Your Submission

Clinical Neurophysiology <clinph@elsevier.com> Thu, Jul 16, 2015 at 6:17 AM
To: joann.tang@mq.edu.au, eagtang2007@gmail.com

Ref.: Ms. No. CLINPH-D-15-8412R1
Sound envelope processing in the developing human brain: A MEG study
Clinical Neurophysiology

Dear Dr. Tang,

I am pleased to accept your revised paper for publication in "Clinical Neurophysiology".

Comments from a Reviewer can be found below.

Thank you for submitting the manuscript to "Clinical Neurophysiology".

Best wishes,

David Burke, MD, DSc
Editor-in-Chief
Clinical Neurophysiology

Comments from the Editors and Reviewers:

Reviewer #1: Authors responded properly in detail to all my comments.
The manuscript was improved significantly.
I have nothing more to comment.
