Disentangling the effects of stimulus context on auditory responses using electroencephalography

Paul M. Briley, BSc (York)

Thesis submitted to the University of Nottingham

for the degree of Doctor of Philosophy

July 2011

Abstract

A ubiquitous feature of neural responses is their dependence on stimulus context. One prominent contextual effect is the reduction in neural response size with stimulus repetition, known as "adaptation". As adaptation is often stimulus-specific, it has been used in visual neuroimaging studies to probe mechanisms of stimulus representation that would otherwise be hidden due to the limited spatial resolution of the available measurement techniques. However, work on the visual system has suggested that stimulus-specific adaptation may not only *reflect* stimulus representations, but may itself also *modify* representational information. The four studies described in this report examined the effects of stimulus context on auditory cortical responses using electroencephalography (EEG).

The first study used adaptation to examine the neural representation of musical pitch in auditory cortex. Whilst pitch is often treated as a single dimension, namely, the repetition rate of the stimulus waveform, in music, pitch actually has two dimensions: pitch height (the octave in which a note resides) and pitch chroma (the position of the note within an octave). The current study provided evidence for an explicit representation of pitch chroma in an anterolateral region of non-primary auditory cortex.

The second, third and fourth studies examined the auditory "mismatch response" (MMR). The MMR refers to the increase in response size to a stimulus when it is presented infrequently (as a "deviant") compared to when it is presented frequently (as a "standard"). The second study found that the MMR could not be fully accounted for by a passive release from adaptation. Instead, the MMR seemed to reflect a sharpening of the neural representation of the adaptor stimulus with repeated presentation. This suggests that the MMR may be involved in perceptual learning.

The third study examined the time courses of the contextual effects on neural responses. Both short- and longer-term effects were observed, with the effects differing between the different components of the auditory evoked response. Notably, the N1 component was influenced by complex effects that seemed to partially reflect the longer-term probabilities of certain short segments of the stimulus sequence, whereas the P2 was influenced by a strong suppressive effect with a remarkably short time course. The fourth study examined whether the contextual effects on auditory-evoked transient and sustained responses are sensitive to the absolute, or the relative, stimulus probabilities. For the transient N1 response, the most striking finding was that adaptation was broadly tuned for deviant stimuli, but sharply tuned for stimuli that were, in terms of their relative probabilities, standards. In contrast, the sustained response appeared to be influenced by a different effect, which facilitated responses to deviant stimuli.

The current results suggest that contextual effects differ vastly between different deflections of the auditory-evoked responses, that they include effects that are both complex and long-lasting (of the order of ten seconds or longer), and that they involve not only suppressive, but also facilitatory effects.

Acknowledgements

I would like to thank first Dr Katrin Krumbholz, for her remarkable ideas, encouragement and support. For her enthusiasm and patience, for elucidating the method of EEG and for teaching me the skills needed to embark on my scientific journey.

For funding and facilities, I thank the Medical Research Council (MRC). I would also like to thank all the students and staff at the MRC Institute of Hearing Research. In particular, the members of the EEG group (in order of appearance): David Magezi, Barrie Edmonds, Nick Clark, Cris Lanting, Gemma Hutchinson, Heather Gilbert, Jessica de Boer and Mark Fletcher.

For help with data collection, I thank Charlie Caunt (Chapter 1), Diana Omigie (Chapter 3), Helen Young & Alex Utev (Chapter 4).

I thank my parents, Mike and Jackie, for caring and nurturing me, and for not complaining too much when I went AWOL during the writing-up period. I am grateful to my brother, Simon, who I am attempting to discourage from doing a PhD, but who may well dominate the field of Sports Science for years to come. I am indebted to my sister, Amanda, who from a young age (≈ 9) has expressed enthusiasm for Psychology and debated my research. I would also like to thank Patch, Domino, Maisie and Mishka.

For her support, patience, love and humour, I dedicate this thesis to my wife, Sarah. Words cannot express how much you mean to me.

Table of contents

List of abbreviations

Adjar	adjacent response technique (Woldorff, 1993)
AFz	central forehead electrode
Ag/AgCl	silver/silver chloride
ALA	anterolateral area of the superior temporal plane
BESA	brain electrical source analysis software
CF	centre frequency
cP1, cN1, cP2	deflections of the pitch-change response
Cz	vertex electrode
dB	decibel
DC	direct current
dP1, dN1, dP2	deflections of the deviant-minus-standard difference waveform
EEG	electroencephalography
EEGLAB	EEG analysis toolbox (Delorme & Makeig, 2004)
EOR	energy onset response
ERB_N	equivalent rectangular bandwidth for normally-hearing people (Glasberg & Moore, 1990)
ERP	event-related potential
fMRI	functional magnetic resonance imaging
HL	hearing level
Hz	hertz
kHz	kilohertz
kΩ	kiloohms
MEG	magnetoencephalography
MMR	mismatch response

ms	milliseconds
nAm	nanoamperes
P1, N1, P2	transient deflections of the auditory event-related potential
PCR	pitch change response
RI	regular interval
RMS	root mean square
s	seconds
SCD	scalp current density
SD	standard deviation
SOA	stimulus onset asynchrony
SPL	sound pressure level
SR	sustained response
SSI	stimulus specificity index
t	time
TDT	Tucker-Davies Technologies
TE1.0	cytoarchitectonic subdivision of primary auditory cortex (Morosan et al., 2001)
TVD	total variation distance
Δf	frequency separation
$\Delta f0$	fundamental frequency (pitch) separation
µV	microvolts

General introduction

Early recordings from single neurons of animals revealed that sensory-evoked responses do not remain constant throughout a stimulus, but instead decrease over time (Adrian, 1928). Adrian (1928) termed this response suppression, "adaptation". Shortly afterwards, Berger (1929; in Swartz & Goldensohn, 1998) demonstrated that neural activity can be recorded non-invasively from the human cortex, using electrodes placed on the surface of the scalp. With this method, known as "electroencephalography" (EEG), it was found that adaptation does not only occur during a stimulus, but can also influence responses to subsequent stimuli, even when these stimuli are separated by several seconds from the adapting stimulus (Davis et al., 1966; Butler, 1968). The greatest response suppression was found when the subsequent stimulus was identical to, rather than different from, the adaptor. Thus, adaptation is stimulus specific. Subsequent work has found that stimulus-specific adaptation is ubiquitous across the senses (e.g. McLaughlin & Kelly, 1993; Di Lorenzo & Lemon, 2000; Adelman & Herson, 2004; Ulanovsky et al., 2004; Kohn, 2007). However, its roles, and functional properties, remain poorly understood. The current report describes four EEG studies that investigated the effects of stimulus context on auditory cortical responses.

Since the early EEG studies, stimulus-specific adaptation has primarily been used as a tool to investigate the neural representations of visual stimulus attributes. This is because adaptation provides a means to overcome the limited spatial resolution of current neuroimaging methods (for a review, see Grill-Spector et al., 2006). Typically, stimuli are presented in pairs, consisting of an "adaptor" and a "probe". The extent to which the adaptor suppresses the probe response would be expected to reflect the extent to which the neural populations sensitive to the adaptor and probe

overlap. If the adaptor and probe are represented by similar groups of neurons, the probe should elicit a small response. If they are represented by different neural groups, the probe should recruit many "fresh", unadapted neurons, and thus elicit a large response.

The first study of this project, described in **Chapter 1**, sought to use the adaptation paradigm in the auditory domain, to examine the neural representation of musical pitch. Pitch is a perceptual property of periodic sounds, such as vowels and musical notes. It is fundamental to sound perception, conveying meaning and emotion in speech and music, yet it remains unclear how pitch is represented in the human brain. It may be that pitch is represented as a single, linear dimension, ranging from "low" to "high". This would be consistent with the close relationship between pitch and the stimulus repetition rate (Plack & Oxenham, 2005). Alternatively, pitch may be represented as a circular, "chroma" dimension, consistent with the cyclical nature of Western, and many non-Western, musical scales (Burns, 1999). In music, pitch chroma corresponds to the note name (e.g. "C"). Krumbholz et al. (2003) demonstrated that a pitch-specific response could be recorded by presenting a pitch-evoking probe stimulus immediately after a noise adaptor. In the current study, both the adaptor and the probe had pitch. Probe responses were measured as a function of the adaptor-probe pitch separation. If pitch is represented as a linear dimension, the probe response should increase monotonically with the pitch separation. However, if there is an explicit representation of pitch chroma, the function relating the probe response to the pitch separation should be non-monotonic with a dip in response size at octave pitch separations where the adaptor and probe have the same chroma.

Whilst adaptation is typically thought to *reflect* stimulus representations, it may be that adaptation is a more complicated effect that can itself *modify*

representational information. Indeed, Desimone (1996), based on work in the visual system, suggested that the tuning curves of neurons sensitive to a stimulus might sharpen when that stimulus is repeated (see also Wiggs & Martin, 1998). As a result, neurons optimally tuned to the repeated stimulus would continue to respond, whilst neurons tuned to other stimuli would cease responding, creating a sparser and potentially more efficient stimulus representation. In the auditory domain, the response to a stimulus differs depending on whether it is presented frequently or infrequently. This effect is typically studied with the classic auditory oddball paradigm. The difference in the size of the response to infrequent "deviant" stimuli presented amongst frequent "standards" is referred to as the "mismatch negativity" or "mismatch response" (MMR; for a review, see Näätänen et al., 2007). Whilst the MMR could arise due to a change in the neural representations of standards and deviants, the possibility that the MMR arises merely due to a passive release from adaptation cannot currently be ruled out (see Jääskeläinen et al., 2004). This is because standard stimuli are more likely to be preceded by the same, rather than a different, stimulus than deviant stimuli. The second study, described in **Chapter 2**, examined whether the MMR arises merely from the accumulated adaptational effects of the stimuli preceding standards and deviants. The response to a probe tone was measured when it was preceded by one, two or three adaptors, at four adaptor-probe frequency separations. Responses after one adaptor were used to estimate the contribution of adaptation to responses after multiple adaptors, to see if adaptation could fully account for the MMR.

As the study reported in **Chapter 2** used short stimulus sequences, it could not fully examine the time courses of the contextual effects underlying the MMR. The third study, described in **Chapter 3**, presented stimuli in continuous sequences akin to

those used previously to investigate the MMR. It is typically assumed that the MMR reflects the longer-term ("global"), rather than the short-term ("local"), stimulus probabilities (Horváth & Winkler, 2004). However, the classic oddball paradigm presents standards and deviants in Bernoulli sequences, in which the local and global probabilities are identical. Therefore, one cannot easily distinguish whether an effect is sensitive to the local or the global stimulus probabilities, as the two are confounded. In order to examine the memory spans of the contextual effects underlying the MMR, the current study dissociated the local and global stimulus probabilities using two methods. The first method manipulated the local and global probabilities separately by presenting stimuli in Markov sequences. In a Markov sequence, the local probability of a stimulus depends on the identity of the preceding stimulus, whilst the global probability of each stimulus within the sequence as a whole stays constant. The second method examined local and global probabilities separately, by measuring responses in Bernoulli sequences as a function of the local stimulus history.

The mechanism underlying the MMR is thought to take into account the relationships amongst stimuli in a sequence (Sussman et al., 2003). Thus, the MMR should be sensitive to the relative, rather than the absolute, stimulus probabilities. However, in the classic oddball paradigm, there is only one standard stimulus and one deviant stimulus. This means that one cannot alter the relative stimulus probabilities without affecting the absolute probabilities. The final study, reported in **Chapter 4**, compared responses in a traditional oddball paradigm with responses in a modified oddball paradigm, which contained more than two stimuli. The stimuli differed in acoustic frequency, which enabled changes in the stimulus specificity of the contextual effects to be examined. By using longer-duration stimuli than those used in the second and third chapters, the study sought to elicit both transient and sustained

evoked responses. The contextual effects on the sustained response (SR) are particularly poorly understood. Whilst Picton et al. (1978a; 1978b) found only weak contextual effects on the SR, Kretzschmar & Gutschalk (2010), using the oddball paradigm, found much stronger effects.

Overview of key studies

Two studies provided particular impetus to this work, namely Jääskeläinen et al. (2004) and Ulanovsky et al. (2004). Jääskeläinen et al. conducted a number of experiments on humans that provided support for an adaptation-based interpretation of the MMR, whilst Ulanovsky et al. explored the time course of stimulus-specific adaptation (SSA) in single neurons of cat auditory cortex.

Jääskeläinen et al. (2004)

Jääskeläinen et al. challenged two key findings that have been used to argue that the MMR is generated by a group of neurons distinct from those that generate the sensory-driven responses to standards and deviants. In doing so, the authors provided support for their view that the MMR arises due to differential adaptation in populations of neurons sensitive to the standard and deviant stimuli, rather than due to an endogenous, deviance-detection mechanism (the traditional view of the MMR). The two key findings that they addressed were: (1) The MMR appears to arise only when there is a repetitive "standard" stimulus, whilst the sensory-driven N1 response occurs to any perceivable sound (Näätänen, 1992). (2) When modelled with equivalent current dipoles, the source of the MMR appears somewhat anterior to the source of the N1 (Tiitinen et al., 1993).

In their first experiment, Jääskeläinen et al. presented a pure-tone probe stimulus every 3.5 seconds. Each probe was preceded by one, two or four standards. Neural responses were measured with EEG and its magnetic counterpart, "magnetoencephalography" (MEG). They found a significant MMR in all conditions, suggesting that a repetitive standard stimulus was not a pre-requisite for MMR elicitation. In the second experiment, Jääskeläinen et al. presented 1/5-octave wide

noise bursts in oddball sequences. In different conditions, the centre frequencies of the standard and deviant bursts differed by one, two or four octaves. Recordings were again made with EEG and MEG. The authors fitted two dipoles to the N1 response in the MEG data – one to the initial, ascending phase of the N1 and one to the latter, descending phase. The latter dipole localised anterior to the former. Interestingly, as the frequency separation between the standard and deviant decreased, the posterior N1 was suppressed more rapidly than the anterior N1. This led to an anterior shift in the centre of gravity of the neural activity which, when fitted with a single dipole, would lead to an apparent shift in source location. The authors concluded that this shift in the relative contributions of two, sensory-driven neural groups could explain the previously reported location differences between the N1 and the MMR. Thus, they argued that there was no need to posit an additional MMR source.

Jääskeläinen et al.'s study provoked a rebuttal the following year by leading proponents of the traditional MMR interpretation (Näätänen et al., 2005). One of their key criticisms was that, whilst Jääskeläinen et al. demonstrated that adaptation could produce effects that resembled the MMR, they did not demonstrate that adaptation *caused* the MMR. Indeed, it has long been acknowledged that adaptation contributes to the deviant-minus-standard difference waveform (Näätänen, 1990; 1992). Thus, Jääskeläinen et al.'s demonstration of an "MMR" after a single standard is not too surprising – Näätänen et al. (2005) argue that, after a repeated standard, there is a "true MMR" alongside this adaptation-related effect. Likewise, Jääskeläinen et al.'s demonstration of differences in the frequency tuning of adaptation between anterior and posterior N1 sources was dismissed as interesting, but not measuring the true MMR. The second experiment in this thesis addresses Näätänen et al.'s (2005) theory

directly, examining whether there is an additional contribution to the MMR after a repeated, rather than a single, standard.

Ulanovsky et al. (2004)

Single- and multi-unit studies of contextual effects in auditory cortex had primarily focussed on short time scales, presenting pairs of stimuli separated by intervals of up to 200-300 ms (e.g. Calford & Semple, 1995; Brosch & Schreiner, 1997). Typically, in these studies, the response to the second stimulus in each pair was suppressed, an effect termed "forward masking" to relate it to a psychoacoustic phenomenon of the same name. In psychoacoustic forward masking, a stimulus is rendered inaudible by a closely preceding stimulus (Jesteadt et al., 1982; Moore, 1978). Over these short time intervals, postsynaptic inhibition is likely to play a prominent role in response suppression; over longer time intervals, inhibition is minimal (Wehr & Zador, 2005). However, a number of psychoacoustic phenomena (for example, auditory streaming; Carlyon et al., 2001) operate over relatively long time scales, on the order of seconds and minutes. Ulanovsky et al. studied the responses of single neurons in auditory cortex, using stimuli that could probe these longer time scales.

Ulanovsky et al. presented long sequences of 230-ms pure tones with an SOA of 736 ms. In their main, "oddball", experiment, two spectral frequencies, centred close to the neuron's best frequency, were presented as standards and deviants. The frequency separation and probabilities of the stimuli were varied in separate blocks, and each block was presented twice, once with the lower, and once with the higher, frequency as standard. Neural firing rates were typically higher to deviant, than standard, stimuli; an effect that was most pronounced at the largest frequency

separation. Interestingly, the normalised difference in firing rate, a measure of SSA, was shown to correlate across conditions; that is, some neurons generally displayed stronger SSA than other neurons. The magnitude of SSA did not correlate with a host of neural response parameters (e.g. sound-level threshold, best frequency and response area sharpness). However, neurons within each cortical column displayed similar levels of SSA, suggesting the presence of "adaptation columns" (although the authors do not report any map-like organisation of these columns). To characterise the time course of SSA, Ulanovsky et al. averaged responses across neurons, separately for standards and deviants and separately for each stimulus position. Except for one stimulus (the 10% deviant, with the largest standard-deviant frequency separation, for which no adaptation was observed), responses decreased exponentially across trials. The decrease was more rapid for stimuli with higher, than lower, probabilities.

To characterise further the time courses of these contextual effects, Ulanovsky et al. constructed a history tree for each stimulus probability, following the approach of Squires et al. (1976), who examined the auditory P300 (a late, attention-related ERP component that occurs when a participant detects an infrequent target stimulus). Each tree plotted the neural firing rate as a function of the immediately-preceding stimulus history, allowing one to compare short-term, "local" effects (within a tree), with longer-term, "global" effects (between trees). Within each tree, responses were larger for stimuli preceded by a different stimulus than for stimuli preceded by an identical stimulus. This "one-back" effect extended further, into an "n-back" effect, with larger responses for sequences that began with a different, rather than an identical, stimulus (at least for three- or four-back, the furthest the authors examined). Importantly, there were still differences between trees when comparing firing rates to identical, short stimulus sequences, suggesting the presence of both short-term and

longer-term effects. Ulanovsky et al. quantified these effects using a linear model with separate (additive) local and global components. The global component was the overall stimulus probability (10%, 30%, 50%, 70% or 90%), whilst the local component reflected the abundance of the stimulus in the immediately-preceding sequence (normalised and weighted to emphasise the most recent stimuli). Fitting the model to the data, the local and global components were both statistically significant. From this, the authors concluded that there were at least two, separate, time constants of adaptation – a short time constant (≈ 1.5 s, based on the model fits) and a longer time constant (> 10 s).

One criticism of Ulanovsky et al.'s model is that the "global" component could merely reflect unaccounted-for local effects. For example, Ulanovsky et al. found that a stimulus preceded by three identical stimuli gave a smaller response when the stimulus had a 70% overall probability than when it had a 30% probability. This could merely be because, in the 70% case, the fourth preceding stimulus was also likely to have been an identical stimulus whereas, in the 30% case, the fourth preceding stimulus was likely to have been different. In Chapter 3, I use stimulus history trees to analyse the time courses of the contextual effects on the P1, N1 and P2 components of the auditory event-related potential. I fit responses with a model that takes into account unmeasured portions of the stimulus sequence by weighting its parameters according to the longer-term stimulus probabilities.

Chapter 1. Electroencephalographic evidence for a representation of pitch chroma in non-primary auditory cortex.

1.A INTRODUCTION

Periodic sounds, such as vowels or musical notes, elicit the percept of "pitch". Pitch is an extremely important perceptual attribute of sound. It conveys emotion in speech, melody in music, and it provides a cue by which the auditory system can distinguish sounds of interest from the acoustic background (Scherer, 1995; Carlyon, 2004). Recent neuroimaging studies have identified a region of non-primary auditory cortex that is more active to sounds with, than without, pitch. This region encompasses the anterolateral part of Heschl's gyrus on the superior temporal plane (referred to as "ALA") and the anterior planum temporale (Gutschalk et al., 2002; Patterson et al., 2002; Penagos et al., 2004; Hall et al., 2006; Hall & Plack, 2007; Hall & Plack, 2009; Puschmann et al., 2009). However, little is known about how pitch is represented in this area.

Pitch is thought to be related to the repetition rate of the stimulus waveform, as stimuli with the same repetition rate are typically judged to have the same pitch value (Plack & Oxenham, 2005). Just as repetition rate can be varied along a linear scale, pitch is often treated as a single, linear dimension, ranging from low to high. For instance, male voices are said to have "lower" pitch than female voices and the wailing of a fire alarm is said to have "higher" pitch than the buzzing of a desk fan. This conception of pitch as a single dimension is captured by the American National Standards Institute definition of pitch as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high" (ANSI, 1994). However, a single, linear dimension does not do justice to the perceptual characteristics of musical pitch. The musical scales of Western, and many non-

Western, cultures are cyclical rather than linear, in that each octave is similar to the next (Burns, 1999). This cyclicality captures the finding that a note and its octave are perceived as particularly similar, which emerges as early as three months of age (Demany & Armand, 1984; Dowling, 1999). Adult listeners' judgements of pitch similarity are best represented as a spiral, consisting of a linear pitch "height" dimension and a circular pitch "chroma" dimension (Ueda & Ohgushi, 1987). Pitch height corresponds to the octave in which a note resides, whilst pitch chroma corresponds to the name of the note within that octave. Thus, middle C and the C above have the same pitch chroma, but differ in pitch height. The aim of the current study was to examine whether auditory cortex contains an explicit representation of pitch chroma, or whether pitch is represented in terms of the stimulus repetition rate.

Primary auditory cortex contains a spatial map of sound frequency, with adjacent neurons tuned to adjacent spectral frequencies (Formisano et al., 2003; Eggermont & Roberts, 2004; Talavage et al., 2004). However, a cortical map of pitch has so far proven elusive. It may be that pitch-sensitive neurons are not arranged according to their preferred pitch value, and are instead randomly intermixed in a patch of cortex, or arranged according to another stimulus feature. If this were the case, it would pose a problem for studying the cortical representation of pitch in humans, as non-invasive imaging techniques have limited spatial resolution and would thus not be able to resolve responses from neurons sensitive to different pitch values. The current study overcame this problem by using a paradigm based on response adaptation (Krumbholz et al., 2003). Such paradigms have been widely used in visual neuroimaging studies to probe representational mechanisms in both primary and non-primary cortex (see Grill-Spector et al., 2006, for a review).

The rationale of the adaptation paradigm is that the responsiveness of neurons temporarily decreases when they are made to fire repeatedly. Therefore, the neural response to the transition between an adaptor stimulus and a probe stimulus can be used as a measure of the stimuli's representational similarity: If an adaptor and probe are represented by similar groups of neurons, the probe response would be expected to be small (Fig. 1.1B). In contrast, if an adaptor and probe are represented by different groups of neurons, the probe will recruit many "fresh", or unadapted, neurons, and thus elicit a large response (Fig. 1.1C). The adaptation paradigm relies on the temporal, rather than the spatial, properties of the neural signal. Therefore, whilst the paradigm has primarily been used with functional magnetic resonance imaging (fMRI), it should be particularly effective when applied with electroencephalography (EEG), which has a temporal resolution on the order of milliseconds.

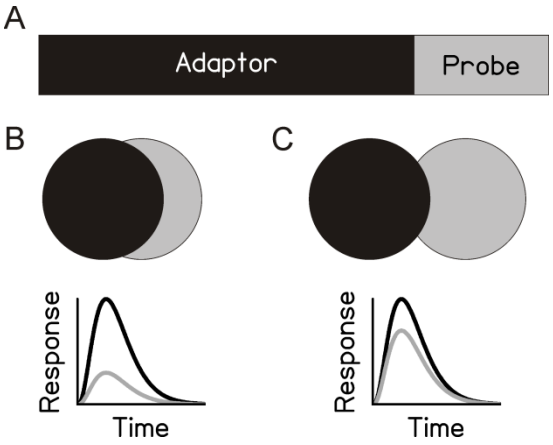

Figure 1.1. A: In the adaptation paradigm, a short "probe" stimulus is preceded by a long "adaptor" stimulus. B: If the adaptor and probe are represented by similar groups of neurons (black and grey circles, respectively), the response to the probe would be expected to be small. C: In contrast, if the adaptor and probe are represented by different groups of neurons, the probe would be expected to elicit a larger response as it will recruit a larger number of "fresh", or unadapted, neurons.

1.B EXPERIMENT 1

1.B.i Methods

Stimuli

On each trial, a 1500-ms adaptor stimulus was directly followed by a 250-ms probe stimulus. There was a silent interval of 1500 ms between trials. Stimuli were complex tonal sounds similar to those produced by musical instruments in echoic concert halls. They were produced by regularising the temporal fine structure of samples of Gaussian random noise. This was achieved using a delay-and-add algorithm, where the original noise sample is added to a copy of itself, after having been delayed by *d* ms. The process is iterated *n* times. This increases the preponderance of time intervals at *d* ms in the stimulus waveform. The resulting "regular interval" (RI) sound evokes a pitch corresponding to the reciprocal of the delay, *d* (Yost, 1996). In a conventional periodic sound with a deterministic waveform, *d* corresponds to the repetition period and its reciprocal corresponds to the repetition rate. The salience of the pitch increases with the number of iterations, *n* (Yost et al., 1996). In the current study, *n* was fixed at 16, which evokes a relatively strong pitch, whilst the repetition rate was varied.

The repetition rate of the probe stimulus on each trial was selected from a 1/3-octave range around 125 Hz. This value is the mean pitch of a human male voice, roughly an octave below middle C on the piano. The repetition rate of the adaptor stimulus was set relative to that of the probe. Adaptor and probe stimuli were gated on and off with overlapping 10-ms cosine ramps so that the energy envelope remained constant across the transition. The combined waveform was then bandpass-filtered from 0.8 to 3.2 kHz. The stimuli were generated afresh for each trial, using new samples of Gaussian random noise, and were presented at about 70 dB SPL. During

the experiment, a continuous lowpass noise with a 0.8-kHz upper cutoff was presented at 40 dB SPL per auditory filter bandwidth (as defined by the equivalent rectangular bandwidth for normally-hearing people, ERB_N; Glasberg & Moore, 1990), to mask any audible distortion products created by the non-linear nature of cochlear processing.

As well as increasing the dominance of time intervals at *d* ms in the stimulus waveform, the delay-and-add process also creates ripples in the stimulus spectrum, with peaks at harmonics (integer multiples) of the stimulus repetition rate (Yost, 1996; Yost et al., 1996). This could be a problem, because a note and its octave have more harmonics in common than a note and its half-octave (Fig. 1.2 A-B). The greater harmonic overlap means that the octave adaptor might suppress a larger number of the neurons sensitive to the probe's individual spectral harmonics than the half-octave adaptor. This would lead to a smaller response from the neural population. However, harmonic overlap is only a problem if the peripheral auditory system is able to resolve the harmonics of the adaptor, so that different harmonics excite different neural groups (Fig. 1.2 D). The spectral effects cannot occur if the peripheral auditory system is unable to resolve the stimulus harmonics, so that different harmonics excite the same neural groups (Fig. 1.2 C). In this case, the only systematic difference between the adaptor and probe is the difference in their temporal structure. The passband of the RI sounds and the repetition rate of the probe stimulus were chosen so that the probe stimulus would be partially resolved by the peripheral auditory system, defined as 2-3.25 harmonics per 10-dB auditory-filter bandwidth ($1.8 \cdot ERB_N$; Shackleton & Carlyon, 1994). Therefore, adaptor pitches below the probe pitch were unresolved, as the stimulus harmonics were closer together (Fig. 1.2C), whereas

adaptor pitches above the probe pitch were resolved, as the stimulus harmonics were further apart (Fig. 1.2D; see also Fig. 1.3).

Stimuli were generated in Matlab (The Mathworks Inc., Natick). They were digital-to-analogue converted with a 24.4 kHz sampling rate and 24-bit amplitude resolution using a TDT System 3 (Tucker Davis Technologies, Florida) and delivered diotically to participants through headphones (K 240 DF; AKG, Vienna).

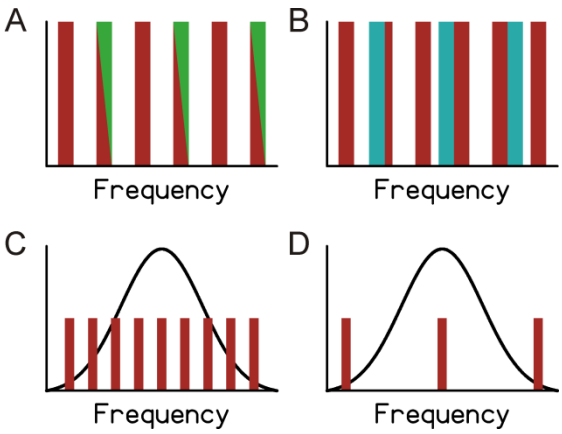

Figure 1.2. A note and its octave (A, red and green bars, respectively) share half of their spectral harmonics. In contrast, a note and its half-octave (B, red and blue bars, respectively) have fewer harmonics in common. C: However, the spectral detail of a stimulus is unavailable to the auditory system when the spacing of the harmonics is closer than the auditory-filter bandwidth (black solid curve). D: The spectral detail is preserved when the harmonics are spaced further apart than the auditory-filter bandwidth.

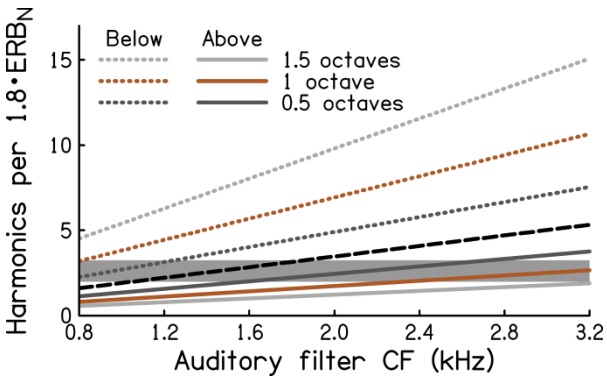

Figure 1.3. The number of spectral harmonics per auditory-filter bandwidth, as a function of the auditory-filter centre frequency (CF). The parameter is the pitch of the adaptor stimulus (the probe stimulus is plotted as a dashed black line). The grey bar represents the limits of resolvability, as defined by Shackleton & Carlyon (1994; 2-3.25 harmonics per auditory filter). Adaptor pitches below the probe were predominantly unresolved over the stimulus bandwidth (0.8 to 3.2 kHz), whilst adaptor pitches above the probe were predominantly resolved.

Procedure

Musical intervals are often expressed in cents, a logarithmic measure in which 1200 represents a pitch separation of an octave. One hundred cents is one semitone on the equally-tempered musical scale, which is the pitch interval between two adjacent keys on the piano. On each trial the adaptor pitch was 0.5, 1 or 1.5 octaves above or below the probe pitch (600, 1200 or 1800 cents). Note that in the 1-octave condition, the adaptor and probe had identical pitch chroma, whereas in the 0.5- and 1.5-octave conditions, the adaptor and probe were separated in pitch chroma by half an octave (the maximum possible pitch-chroma separation). Stimuli were randomly presented in four blocks of twenty minutes each.

Electroencephalography and data pre-processing

Auditory evoked potentials were recorded from Ag/AgCl electrodes. Electrodes were placed at 32 quasi-equidistant scalp locations, according to the 10% system (Easycap, Herrsching). Skin-electrode impedances were kept below 5 kΩ. During recording, channels were referenced to an electrode at Cz, with an electrode at AFz serving as ground. Signals were amplified by a BrainAmp DC amplifier (Brainproducts GmbH, Munich) and analogue bandpass-filtered online from 0.1-250 Hz. The signals were sampled at 500 Hz and stored for offline analysis using the Brain Vision Recorder software (Brainproducts GmbH, Munich). Participants watched a silent, subtitled DVD of their choice.

Recordings were processed offline with the EEGLAB toolbox (Delorme & Makeig, 2004), which runs under Matlab (The Mathworks Inc., Natick). Recordings were lowpass-filtered at 35 Hz, down-sampled to 250 Hz per channel by averaging together two consecutive sample points and then re-referenced to average reference.

Data were segmented into epochs ranging from 100 ms before the start of the adaptor stimulus to 500 ms after the end of the probe stimulus (total duration of 2600 ms), and responses were baseline-corrected to the 100 ms pre-stimulus interval. Epochs that contained unusually large potentials (joint probability limits of ± 3 SD) were automatically rejected and the remaining epochs were submitted to an independent component analysis for each run and each participant separately, using the Infomax extended ICA algorithm (Bell & Sejnowski, 1995; Lee et al., 1999). Components representing eye blinks, lateral eye movements and heartbeats were removed by manual inspection. Epochs were then averaged to obtain event-related potentials (ERPs) for each participant and condition.

ERPs were imported into BESA (MEGSIS Software GmbH, Munich) and source waveforms from each hemisphere were derived from a two-dipole model, with one dipole at the centre of mass of primary region TE1.0 in each hemisphere (Morosan et al., 2001). The orientations of the dipoles were fitted within a window encompassing the P1, N1 and P2 peaks of the mean probe response (0 to 300 ms after stimulus onset), averaged across participants and conditions. Preliminary statistical analyses found no significant interactions involving the effect of hemisphere ($p > 0.05$), therefore source waveforms were averaged across hemispheres, and further analyses were performed on these averages.

Data analysis

Adaptor stimuli elicited typical energy-onset responses (EORs), with peaks corresponding to the P1, N1 and P2 components of the auditory ERP (Fig. 1.4; Näätänen and Picton, 1987). Responses to probe stimuli (referred to as "pitch-change responses"; PCRs) were similar in morphology to the adaptor EORs (Fig. 1.4).

Indeed, a previous study, in which a probe RI sound followed a random-noise adaptor, suggested that the transition response was a subcomponent of the EOR (Seither-Preisler et al., 2004). In recognition of this, the peaks of the PCR will be referred to as the change P1 (cP1), change N1 (cN1) and change P2 (cP2). The sizes of both the cN1 and cP2 were modulated by the pitch separation between the adaptor and the probe. As the components have opposite polarities, and their time courses likely overlap (Näätänen and Picton, 1987; Makeig et al., 1997), it was expected that they would partially cancel each other out. As a result, if a manipulation suppresses the cP2 more than the cN1, the cN1 could appear to increase, solely because it is less cancelled by the cP2. To demonstrate this cancelling effect, cP2 values were flipped in sign and correlated with the corresponding cN1 values. As expected, peak amplitudes were negatively correlated ($r(60) = -0.379$, $p < 0.005$); as one peak decreased, the other tended to increase. To overcome this problem, the amplitude of the PCR was quantified as the peak-to-peak amplitude from the cN1 to the cP2. The latency of the PCR was taken as the latency of the cN1 peak.

Subjects

Ten subjects participated in Experiment 1 (mean age: 22, 5 females). All subjects had pure-tone hearing thresholds at or below 20 dB HL at octave frequencies between 250 and 4000 Hz and had no history of audiological or neurological disease. Studies were approved by the Ethics Committee of the University of Nottingham School of Psychology.

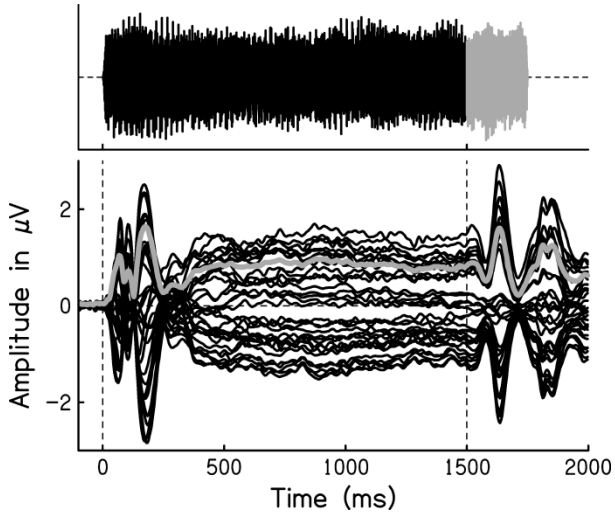

Figure 1.4. Top: The stimulus waveform, with the adaptor plotted in black and the probe plotted in grey. Bottom: The mean event-related potential. Each black line represents a recording channel. The root-mean-square (RMS) amplitude is plotted in grey. The onsets of the adaptor and probe stimuli elicited transient neural responses (the EOR and the PCR, respectively).

1.B.ii Results

Effect of pitch chroma

The cortical response to a change in musical pitch (the "pitch-change response"; PCR) was measured using pairs of complex tonal stimuli, referred to as "RI sounds". If pitch is represented as a single, linear dimension, the amplitude of the PCR should increase monotonically with increasing pitch separation. Contrary to this hypothesis, PCR amplitudes were smaller for the octave pitch separation than either the 0.5- or the 1.5-octave separations (Fig. 1.5). This effect occurred for both resolved and unresolved adaptor stimuli. With unresolved adaptors, the effect could not arise from neural sensitivity to individual spectral components of the stimuli. Instead, the result suggests that the PCR is generated in an area of auditory cortex that contains an explicit representation of pitch chroma.

To quantify the effect of pitch chroma, each octave PCR ("same chroma" condition) was compared to the mean of the 0.5- and 1.5-octave PCRs either side ("different chroma" condition). The 0.5- and 1.5-octave separations have a mean pitch-height interval that is identical to the octave separation, but their mean pitch-chroma interval is different. PCR amplitudes were entered into a linear mixed model, with fixed factors of chroma ("same" or "different") and spectral resolution ("resolved" or "unresolved"), and a random effect of subjects. The main effect of spectral resolution was significant [$F(1,28) = 13.441$, $p < 0.005$], reflecting a larger PCR amplitude for resolved than unresolved adaptors. Importantly, the main effect of chroma was highly significant [$F(1,28) = 29.865$, $p < 0.001$]. There was no significant interaction between the effects of chroma and spectral resolution [$F(1,27) = 0.026$, $p = 0.874$], indicating that the chroma effects for resolved and unresolved adaptors were statistically identical.

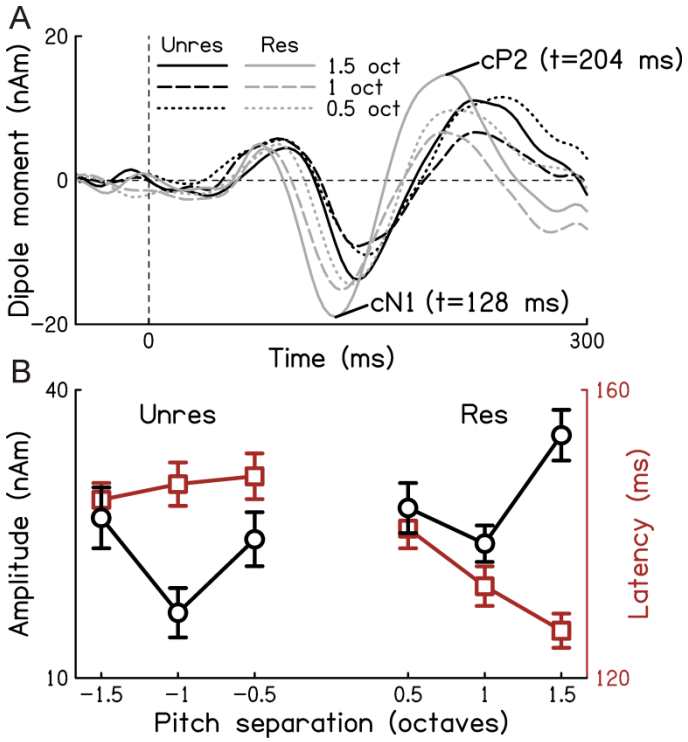

Figure 1.5. A: Source waveforms for the probe stimuli in Experiment 1, averaged across subjects. Note that the response occurred earlier, and with greater magnitude, in the resolved conditions than in the unresolved conditions. B: Mean PCR amplitudes (black) and latencies (red) for each pitch separation. Error bars represent one standard error. PCR amplitudes were smaller at the octave, than the 0.5- and 1.5-octave, pitch separations. In the resolved conditions only, the PCR amplitude was greater for the 1.5-, than the 0.5-, octave separation. The PCR latency was relatively constant in the unresolved conditions, but decreased with increasing pitch separation in the resolved conditions.

Effect of pitch height

The effect of pitch height was quantified by comparing the PCRs for the 0.5- and 1.5-octave pitch separations. These conditions differed in pitch height, but had the same pitch chroma, as they were an octave apart. Interestingly, there was a large effect of pitch height for resolved adaptors, but only a small effect for unresolved adaptors. This was confirmed statistically with a linear mixed model, with fixed factors of height (0.5 and 1.5 octaves) and spectral resolution, and a random effect of subjects. There was a significant main effect of spectral resolution, reflecting larger PCR amplitudes for resolved than unresolved adaptors [$F(1,27) = 15.412$, $p < 0.005$], as well as a main effect of pitch height [$F(1,27) = 12.131$, $p < 0.005$]. Importantly, the interaction of spectral resolution and pitch height was also significant [$F(1,27) = 6.329$, $p < 0.05$]. Paired comparisons found a significant pitch-height effect for resolved adaptors only ($p < 0.001$; unresolved adaptors: $p = 0.5$). This suggests that the height effect primarily arises from neurons sensitive to the individual spectral components of the adaptor, rather than from neurons explicitly sensitive to pitch height.

PCR latency

For unresolved adaptors, the latency of the PCR was remarkably long, around 150 ms after the onset of the probe stimulus. This is 40 ms longer than the latency of the EOR to adaptor stimuli (around 110 ms). The latency was also relatively unaffected by pitch separation. In contrast, the PCR latency for resolved adaptors decreased with increasing pitch separation, going from 136 ms for the 0.5-octave separation to 124 ms for the 1.5-octave separation. These findings were confirmed statistically using a linear mixed model with a fixed factor of spectral resolution, a

covariate of pitch separation (0.5, 1 or 1.5 octaves) and a random effect of subjects. In essence, the statistical model fits a regression line through the three pitch separations for resolved and unresolved adaptors separately, and compares the slopes of the regression lines between the two spectral-resolution conditions. The analysis revealed a significant main effect of pitch height [$F(1,47) = 14.831$, $p < 0.001$], but not of spectral resolution [$F(1,47) = 0.231$, $p = 0.633$]. However, the interaction of spectral resolution and pitch height was significant [$F(1,47) = 5.684$, $p < 0.05$], indicating that the slope of the function relating pitch separation to PCR latency was steeper for resolved adaptors than for unresolved adaptors. Subsequently, slopes were fitted to the PCR latencies for individual participants, and one-sample *t*-tests were used to compare slopes against zero for the resolved and unresolved conditions separately. The slope was significant for resolved adaptors [$t(9) = 7.965$, $p < 0.001$], but not for unresolved adaptors [$t(9) = 0.967$, $p = 0.359$]. This means that the PCR latency reduces as the detectable spectral change increases. Therefore, the latency changes may result from the same process that leads to the apparent effect of pitch height on PCR amplitudes.

Summary

PCR amplitudes were smaller when the adaptor and probe had the same pitch chroma than when they had different pitch chroma. As this effect occurred for both resolved and unresolved adaptors, it implies an explicit representation of pitch chroma in auditory cortex. In contrast, increases in pitch height only led to increases in PCR amplitude when adaptor stimuli were resolved. The effect of pitch height co-occurred with reductions in the PCR latency.

1.C EXPERIMENT 2

1.C.i Methods

Experiment 2 was mostly identical to Experiment 1, except that stimuli were pure tones. As pure tones consist of only a single frequency component, they are fully resolved by the auditory system. On each trial, the adaptor frequency was either above or below the probe frequency. This was to facilitate comparison with the results of Experiment 1, in which unresolved adaptors always had a pitch height below that of the probe whilst resolved adaptors always had a pitch height above the probe. The frequency of the pure-tone probe was selected from 1/3-octave range around 500 Hz on each trial. The adaptor frequency was set relative to this. Adaptor and probe stimuli were gated on and off with overlapping 10-ms cosine ramps so that the energy envelope remained constant across the transition. Stimuli were presented at 70 dB SPL. To equate the tones for loudness, masking noise, filtered such that equal energy (20 dB SPL) would fall into each peripheral auditory filter (ERB_N), was played continuously. Nine subjects participated in this experiment (mean age: 24.3, 5 females). Recording and data analysis parameters were the same as those used in Experiment 1.

1.C.ii Results

In contrast to the results of Experiment 1, PCR amplitudes did not display a dip at the octave. Rather, the size of the PCR increased linearly with increasing pitch separation (Fig. 1.6). A linear mixed model with fixed factors of chroma and direction-of-change ("low to high" or "high to low") found no significant main effect of chroma [$F(1,25) = 0.011$, $p = 0.917$] nor any interaction of chroma and direction [$F(1,24) = 0.071$, $p = 0.792$]. There was a significant main effect of direction [$F(1,25)$

= 46.731, $p < 0.001$], in that adaptor frequencies above the probe produced a larger PCR than adaptor frequencies below the probe.

A similar analysis of the height effect, found that the PCR was significantly larger for the 1.5-octave frequency separation than the 0.5-octave separation [$F(1,25) = 14.813$, $p < 0.005$]. As before, the main effect of direction was significant [$F(1,25) = 29.936$, $p < 0.001$], whilst the interaction of height and direction was non-significant [$F(1,24) = 0.067$, $p = 0.799$]. A further linear mixed model compared the PCR latencies, with direction as a fixed factor and frequency separation as a covariate. PCR latencies were independent of frequency separation. Neither the main effects nor the interaction were significant [direction of change: $F(1,42) = 1.589$, $p = 0.214$; frequency separation: $F(2,42) = 0.323$, $p < 0.726$; interaction: $F(1,42) = 0.037$, $p = 0.848$].

The latency of the pure-tone PCR was around 100 ms, the same as the latency of the EOR to adaptor stimuli. This is much earlier than the PCR for RI sounds (around 150 ms for unresolved adaptors and 130 ms for resolved adaptors). A linear mixed model was used to compare PCR latencies between the pure tone, resolved RI sound and unresolved RI sound conditions. Responses were averaged across pitch separations. The effect of condition was significant [$F(2,16.942) = 198.331$, $p < 0.001$], as were all pair-wise comparisons ($p < 0.001$). A similar analysis on the PCR amplitudes also found a significant condition effect [$F(2,15.244) = 48.090$, $p < 0.001$] and significant differences between each of the three conditions ($p < 0.05$). The PCR was largest for pure tones and smallest for unresolved RI sounds.

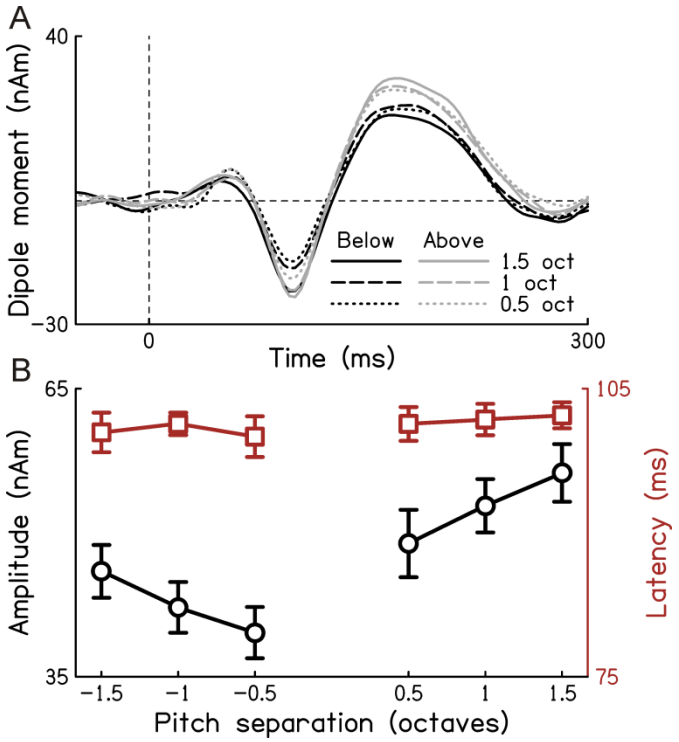

Figure 1.6. A: Mean source waveforms for the probe stimuli in Experiment 2. B: PCR amplitudes (black) and latencies (red) for each pitch separation. Error bars represent one standard error. PCR amplitudes increased with increasing pitch separation, whilst PCR latencies stayed relatively constant.

Summary

The PCR for pure tones had a shorter latency, and larger amplitude, than the PCR for RI sounds. Further, the pure-tone PCR did not display any effects of pitch chroma. Coupled with the results of Experiment 1, these findings suggest that the pure-tone PCR is dominated by the response of an early, spectrally-sensitive region of auditory cortex. In contrast, the PCR for unresolved RI sounds may be dominated by the response of a later, chroma-sensitive region. The relative contributions of the early and late areas presumably depend on the degree of spectral resolution, with the PCRs for the more resolved adaptors receiving greater, and the PCRs for the less resolved adaptors receiving lesser, contribution from the early source. This would explain the latency changes, and the apparent effect of pitch height, found for "resolved" adaptors in Experiment 1.

1.D EXPERIMENT 3

The PCR for unresolved RI sounds had a longer latency than the PCR for pure tones. Potentially, this could be explained in two ways: either, both PCRs are generated in the same auditory region but processing takes longer for unresolved stimuli, or the PCRs are generated in different auditory regions, at different stages of the processing hierarchy. For instance, the PCRs for pure tones and unresolved RI sounds might be generated in primary, and non-primary, auditory cortex, respectively. The responses to the resolved RI sounds might reflect a mixture of both sources. To test this hypothesis, Experiment 3 performed source analyses on the PCRs. Unresolved RI sounds, resolved RI sounds and pure tones were presented at the largest pitch separation (1.5 octaves). Neural responses were recorded from a larger set of electrodes to improve the accuracy of source localisation.

1.D.i Methods

Procedure

Stimuli were presented in six blocks. In the first two blocks, RI sounds were presented with an adaptor pitch 1.5 octaves below the probe pitch ("unresolved RI"). In the second two blocks, RI sounds were presented with an adaptor pitch 1.5 octaves above the probe pitch ("resolved RI"). In the final two blocks, pure tones were presented with an adaptor frequency 1.5 octaves above or below the probe frequency. To compensate for the smaller response sizes for unresolved RI sounds observed in Experiment 1, unresolved-RI blocks lasted 22 minutes whereas resolved-RI and pure-tone blocks lasted 14 minutes. Blocks were presented in a random order for each participant.

Electroencephalography and data pre-processing

Auditory evoked potentials were recorded from 63 scalp locations, using an extended 10-20 arrangement that provided greater coverage of the inferior scalp than the 10% system or the standard 10-20 montage ("Infracerebral" electrode cap; Easycap, Herrsching). This electrode arrangement provides data that are better suited for the localisation of auditory sources.

Recordings were processed using BESA software (MEGIS Software GmbH, Munich). For each participant, recordings from all six runs were concatenated and searched for potentials resembling eye blinks. These potentials were averaged and their first spatial principal component (which accounted for at least 99% of the variance) was used to define an artefact topography. This process was repeated for the lateral eye movements. Subsequently, data were segmented into epochs ranging from 100 ms before the start of the adaptor stimulus to 500 ms after the end of the probe stimulus, as for the other experiments. Epochs containing potentials ±120 μV after correction for eye blinks and lateral eye movements were discarded. Epochs were then averaged to obtain ERPs for each participant and condition.

Data analysis

Experiment 3 compared the loci of the neural generators underlying the PCR for the unresolved RI, resolved RI and pure-tone conditions. For each participant and condition, two equivalent current dipoles were fitted to a 40-ms time window around the peak of the PCR. Neither the locations nor the orientations of the dipoles were constrained. Each fit was obtained by minimising the residual error between the observed scalp potential distribution and the modelled scalp potential distribution, using the BESA software (MEGIS Software GmbH, Munich). Dipoles that localised

at the boundaries of the BESA head model were excluded from subsequent analyses. Only two dipoles were excluded, one left-hemisphere dipole from the resolved RI condition and one left-hemisphere dipole from the unresolved RI condition. All other source models explained over 98% of the variance in the scalp potential distributions.

Confidence intervals were computed for each condition and hemisphere using a between-subjects bootstrap procedure (Efron, 1993). In this procedure, the dipole co-ordinates for each participant were entered into a pool, from which 1000 re-samples were drawn. Each re-sample was obtained by randomly sampling with replacement from the original sample, until the re-sample was the same size as the original. The dipole co-ordinates within each re-sample were then averaged. This yielded a distribution, for each parameter, consisting of 1000 parameter estimates. A bootstrap distribution approximates the sampling distribution of a parameter (Efron, 1993). Therefore, the standard deviation of the bootstrap distribution is equivalent to the standard error of the mean. As each bootstrap distribution was normal in shape and was centred close to the original mean value, 95%-*t*-confidence intervals were derived.

In a similar manner, within-subjects permutation tests were used to compare dipole locations and orientations between conditions (Efron, 1993). In a permutation test, the sampling distribution of a statistic is computed from the data under the null hypothesis that there is no difference between conditions and thus the condition labels (e.g. “unresolved RI sounds” and “pure tones”) are interchangeable. Each data point in a sampling distribution (n = 1000) was obtained by randomly permuting the condition labels for each participant to create two new re-samples, and calculating the difference between the mean values of the re-samples. If the actual difference between a dipole parameter from the two conditions fell at the extreme of the

parameter's sampling distribution, the null hypothesis was rejected. Sampling distributions were computed for the Euclidean distance between dipole locations. When the Euclidean distance was significant, distributions were also computed for the three Cartesian co-ordinates separately.

The orientation of a dipole reflects the net direction of current flow in the group of active neurons (Scherg, 1990). The dipole will be oriented perpendicular to the cortical surface, as the EEG signal predominantly reflects the activity of pyramidal neurons, which are oriented in this manner (Nunez & Srinivasan, 2006; the pyramidal neurons are spatially aligned, so their electric fields sum and can be recorded at the scalp surface). Dipole orientations cannot be compared using Cartesian co-ordinates, as they are not points in three-dimensional space but direction vectors with two degrees of freedom. Instead, differences were assessed by computing the central angle between two orientation vectors, along with a permutation distribution for the central angle. When the size of the central angle was significant, dipole orientations were projected into three planes: sagittal, coronal and transversal. Permutation tests were performed on the angles of rotation in each of these planes (see Tiitinen et al., 1993 and Crottaz-Herbette & Ragot, 2000, where orientations are compared in the sagittal plane). As angles are discontinuous (1 degree is next to 359 degrees), they were compared using circular statistics (Berens, 2009). This involves transforming angles to unit vectors, performing calculations on the vectors, then transforming back to angular form.

Subjects

Eight subjects participated in this experiment (mean age 22.9, 5 female). Participants for Experiment 3 were pre-screened to ensure their EEG responses had an

adequate signal-to-noise ratio for source analysis. In the screening test, participants were presented with a continuous sequence of 100-ms pure tones (frequency 1000 Hz ± 1/6 octave), with an onset asynchrony of 1.5 seconds. Neural responses were recorded from an electrode at the vertex, with a linked-mastoid reference. Only participants with an N1 amplitude 7 μV or above were included in the current study.

1.D.ii Results

In agreement with the findings from the earlier experiments, the PCRs for pure tones occurred earlier, and had greater amplitude, than the PCRs for unresolved RI sounds (Fig. 1.7A). PCRs for resolved RI sounds fell between those for pure tones and unresolved RI sounds in both amplitude and latency. The results were assessed statistically with two linear mixed models, with fixed factors of hemisphere and condition (unresolved RI, resolved RI and pure tone), and a random effect of subject. For both amplitude and latency, the main effect of condition was significant [amplitude: $F(2,37) = 78.603$, $p < 0.001$; latency: $F(2,37) = 201.089$, $p < 0.001$], whilst the main effect of hemisphere was not [amplitude: $F(1,37) = 0.035$, $p = 0.852$; latency: $F(1,37) = 0.330$, $p = 0.569$]. The interaction between condition and hemisphere was also non-significant [amplitude: $F(2,35) = 0.962$, $p = 0.392$; latency: $F(2,35) = 0.879$, $p = 0.424$]. All three conditions were significantly different from each other [amplitude: $p < 0.005$; latency: $p < 0.001$).

Figure 1.7B shows voltage maps for the PCRs from one participant. Maps displayed a vertex negativity and temporal positivity, consistent with bilateral generators in auditory cortex. However, in the left hemisphere, the positive pole of the unresolved RI map was superior to that of the pure-tone map. Both hemispheres also showed small differences in the position of the negative pole, with the pole for the

unresolved map being anterior to that of the pure-tone map. This suggests a difference in the locations and/or orientations of the underlying generators for the pure-tone and unresolved RI conditions in this particular participant. This observation is supported by maps of scalp current density (SCD, Fig. 1.7C), which plot the change in current density across the scalp calculated with the surface Laplacian. The surface Laplacian acts as a spatial filter, reducing the influence of lower spatial frequencies associated with current smearing at the skull-scalp boundary and emphasising well-localised sources in the cortical convexity (Nunez & Srinivasan, 2006). As a result, it provides an estimate of the potential at the cortical surface and helps to disambiguate overlapping topographies arising from simultaneously active neural generators, such as complementary sources in the left and right hemispheres. In both hemispheres, the negative pole of the SCD map was further anterior for the unresolved condition than the pure-tone condition. In the right hemisphere, the positive pole was also anterior.

To test whether the above observations hold for all participants, two unconstrained equivalent current dipoles were fitted to the PCR for each participant and condition. Each dipole models the neural activity of a circumscribed region of cortex. The dipole is located at the centre of mass of the active region and is oriented in the direction of the net intracellular current flow (Scherg, 1990). Confidence intervals for the mean dipole locations (Fig. 1.7D) were produced using a bootstrapping procedure. Dipole locations and orientations were compared statistically using within-subjects permutation tests (see Methods). In both hemispheres, the Euclidean distance between the unresolved and pure-tone dipole locations was significant (left: $p < 0.05$; right: $p < 0.05$). The unresolved dipole was significantly lateral to the pure-tone dipole in the left hemisphere ($p < 0.001$) and significantly anterior to the pure-tone dipole in the right hemisphere ($p < 0.05$). The

central angle between the dipole orientations was also significant in both hemispheres (left: $p < 0.001$; right: $p < 0.05$). Therefore, dipole orientations were projected into sagittal, coronal and transversal planes. In both hemispheres, unresolved dipoles were significantly less vertical in the sagittal plane (left: $p < 0.05$; right: $p < 0.05$) and significantly less radial in the transversal plane (left: $p < 0.05$; right: $p < 0.05$) than pure-tone dipoles.

Whilst significant differences in dipole location were only found in the lateral-medial dimension in the left hemisphere and in the anterior-posterior dimension in the right, the differences in dipole orientation imply that the source of the PCR differed in both dimensions in both hemispheres. Dipoles are oriented perpendicular to the cortical surface (see Methods; Nunez & Srinivasan, 2006). Therefore, because the superior temporal plane becomes more horizontal from its medial to its lateral extent, lateral dipoles will be less radially oriented in the transversal plane (and more vertically oriented in the coronal plane) than medial dipoles. Likewise, as the superior temporal plane initially becomes steeper going posterior to anterior, anterior dipoles will be less vertical in the sagittal plane than posterior dipoles. Note that EEG is more sensitive to differences in dipole orientation than dipole location, because orientation changes lead to larger differences in the distribution of electrical potential at the scalp (Nunez & Srinivasan, 2006). The resolved dipoles fell between the pure-tone and unresolved dipoles in location and orientation (not shown), although these differences did not reach significance ($p > 0.05$).

Summary

PCR dipoles for unresolved stimuli localised lateral (left hemisphere) and anterior (right hemisphere) to those for pure tones. These differences in location were

accompanied by orientation differences in both the sagittal and transversal planes. The results support the hypothesis that the longer PCR latency for unresolved RI sounds than pure tones arises because the unresolved PCR is generated in an auditory region that is at a later stage of the processing hierarchy. The earlier area appears to be sensitive to the stimulus spectrum, whilst the later area is explicitly sensitive to pitch chroma (Experiments 1 and 2). The spectrally-sensitive area might correspond to primary auditory cortex, which is on the medial portion of Heschl's gyrus, whilst the chroma-sensitive area might lie on an anterolateral part of non-primary auditory cortex.

As pure tones give rise to a percept of pitch chroma, it is likely that they too activate the pitch-sensitive region. However, the strong response of the spectrally-sensitive region may have dominated the pure-tone PCR, overriding any chroma effects in the much weaker, non-primary response. Consistent with this interpretation, Hall et al. (2002) found strong responses to pure tones in early, primary auditory areas, but considerably weaker responses in later, non-primary areas. Note that the spectrally-sensitive region would contribute little to the PCR for unresolved stimuli, because there was little detectable spectral difference between the adaptor and probe.

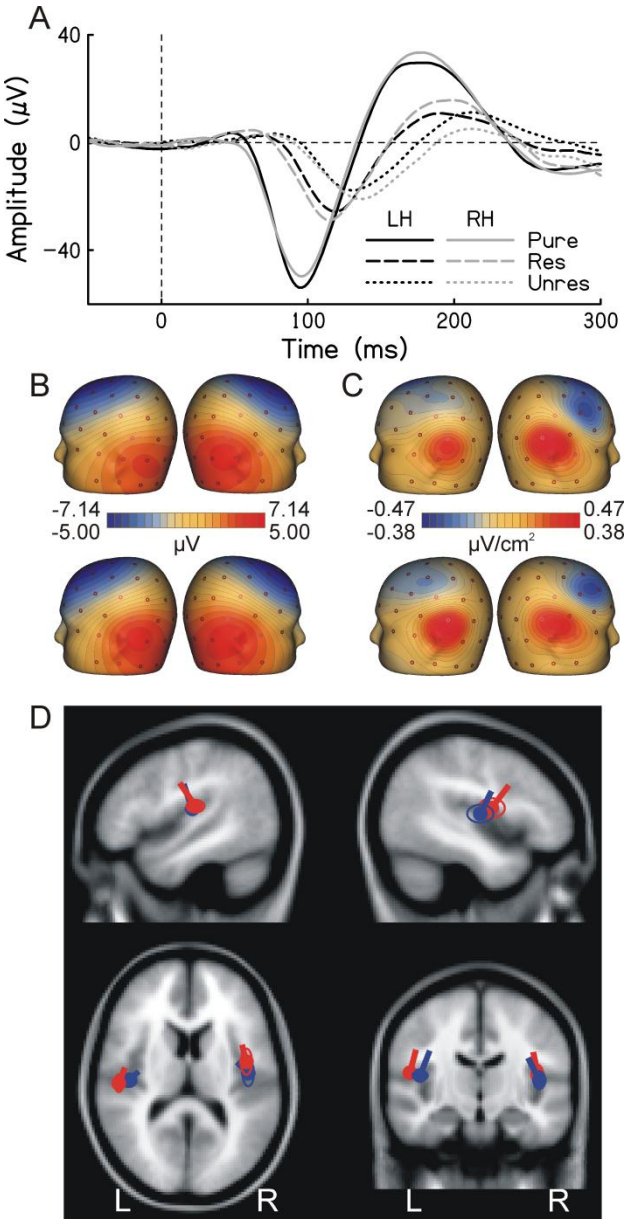

Figure 1.7. A: Mean source waveforms for the probe stimuli in Experiment 3. The response occurred earliest, and with greatest magnitude, in the pure-tone condition, whilst it was smallest, and occurred latest, in the unresolved-RI condition. B-C: Maps of field potential (B) and current source density (C) for the PCR to pure tones (top) and unresolved RI sounds (bottom), for one representative participant. D: Mean dipole source locations and orientations for the PCR to pure tones (blue) and unresolved RI sounds (red). Data are plotted in the sagittal plane (top), the transversal plane (bottom left) and the coronal plane (bottom right). Ellipses show 95%

confidence intervals for each dipole location. The unresolved-RI dipole localised anterior (right hemisphere) and lateral (left hemisphere) to the pure-tone dipole.

1.E EXPERIMENT 4

In Experiment 1, the PCR for RI sounds was smaller for the octave pitch separation than the 0.5- or 1.5-octave separations. This suggests that groups of neurons in auditory cortex are tuned for pitch chroma. However, the octave interval is consistently described as sounding pleasant ("consonant"), whilst the half-octave interval is described as sounding unpleasant ("dissonant"; see Schellenberg & Trehub, 1994 and McDermott et al., 2010). It may be that PCR amplitudes reflect a pitch interval's dissonance, rather than its chroma separation. To test this, Experiment 4 presented RI sounds with a larger set of pitch separations. Some of the pitch separations were perceptually consonant (700, 900, 1200 and 1600 cents), whereas others were perceptually dissonant (600, 1300 and 1800 cents). This experiment also acted as a replication of the findings of Experiment 1, as it used a different group of participants.

1.E.i Methods

Procedure

In Experiment 4, a larger number of pitch intervals were tested with RI sounds. Some of these intervals were perceptually consonant (e.g. 700 cents, the "perfect fifth") whilst others were perceptually dissonant (e.g. 1300 cents, the "minor second"). Stimuli were randomly presented in four blocks of 24 minutes each. In all other respects, stimulus parameters were identical to those used in Experiment 1.

Electroencephalography and data pre-processing

The recording parameters were identical to those used in Experiments 1 and 2.

Data analysis

Experiment 4 examined whether the PCR reflects the difference in chroma between the adaptor and probe pitches, or the perceptual dissonance of the pitches. To test this, mean PCR amplitudes were obtained for each pitch separation by averaging across participants. A number of models were fitted to the data, of the form:

$$p_{\Delta f0} = a \cdot |\sin \pi \cdot \Delta f0| + m \cdot \Delta f0 + b$$

where $\Delta f0$ is the pitch separation in octaves. The model contains three additive components, each with one free parameter. The first component models the circle of pitch chroma using a sine wave that will reach a minimum every octave. The free parameter a scales the sine wave vertically. The second component models the linear rise of pitch height. The free parameter m is the gradient of the slope. The third component contains the intercept parameter, b. The full model was fitted to the data and compared with reduced versions of the model to identify its significant components. For example, to test the significance of the pitch-chroma component, the full model was compared to the model with the pitch-chroma component omitted. Resolved and unresolved conditions were fitted separately. Models were compared using *F*- tests, which take into account differences in model complexity. *F*-ratios were obtained using the following formula:

$$F = \frac{(SS_1 - SS_2)/(df_1 - df_2)}{SS_2/df_2}$$

where index 1 refers to the simpler model (the model with fewer parameters) and index 2 refers to the more complex model. *SS* stands for the sum of the squared differences between the actual data points and the modelled data points (the error sum of squares). *df* stands for the model degrees of freedom (number of data points minus number of free parameters). The degrees of freedom associated with this *F*-ratio has a

numerator of $(df_1 - df_2)$ and a denominator of df_2. If the PCR reflects perceptual consonance, rather than pitch chroma, the proposed model should fit the data poorly; the data should deviate from the sinusoidal function, with peaks and troughs corresponding to perceptually dissonant and perceptually consonant intervals, respectively. To further test whether PCR amplitudes reflect perceptual consonance, ratings of perceptual consonance for pitch intervals within an octave were obtained from McDermott et al. (2010; averaged across instrument type). It was assumed that the consonance of pitch intervals larger than an octave would be similar to the consonance of the same intervals within an octave (for example, a 1.5-octave interval would be rated similarly to a 0.5-octave interval). These ratings were sorted from least consonant to most consonant and the rank orders were entered into a Spearman's rank correlation along with the PCR amplitudes.

Subjects

Twelve subjects participated in this experiment (mean age: 22.1, 6 females). None of these participated in Experiment1.

1.E.ii Results

For both resolved and unresolved adaptors, the PCR was smaller for the octave pitch separation than the 0.5- and 1.5-octave separations (Fig. 1.8). This replicates the results of Experiment 1 in a new group of participants. If the PCR reflects perceptual dissonance, the function relating PCR amplitude to pitch separation should show peaks and dips, corresponding to dissonant and consonant intervals, respectively. This, however, was not observed. Instead, there was a dip near the octave and the amplitude increased monotonically as the pitch-chroma interval

increased. This is consistent with the hypothesis that the PCR reflects pitch-chroma representation in auditory cortex. For example, next to the 0.5-octave (600 cents) pitch separation is the 700-cents separation, known as the "perfect fifth". Whilst the half-octave ("tritone") is highly dissonant, the perfect fifth is highly consonant, yet both intervals displayed similar PCR amplitudes. A linear mixed model with fixed factors of consonance ("consonant" or "dissonant") and spectral resolution found only a main effect of spectral resolution [$F(1,154) = 23.834$, $p < 0.001$]. The main effect of consonance [$F(1,154) = 1.448$, $p = 0.231$] and the interaction between consonance and spectral resolution [$F(1,153) = 0.931$, $p = 0.336$] were both non-significant. Further, a correlation of the mean PCR amplitude for each pitch separation with ratings of perceptual consonance obtained from McDermott et al. (2010) was non-significant [rho(14) = -0.299, $p = 0.298$].

For unresolved adaptors, the function relating PCR amplitude to pitch separation resembled a half-cycle of a sinusoid, consistent with the circular nature of pitch chroma (Fig.1.8). For resolved adaptors, this effect appeared superimposed on a linear increase with pitch height. The mean data, averaged across participants, were fitted with a model that included a pitch-chroma and a pitch-height component, plus a constant intercept (red lines in Fig. 1.8; see Methods). The model fit the data well, explaining over 75% of the variance (resolved: $r^2 = 0.927$; unresolved: $r^2 = 0.785$). The pitch-chroma component was significant for both resolved and unresolved adaptors [resolved: $F(1,4) = 27.560$, $p < 0.01$; unresolved: $F(1,4) = 13.895$, $p < 0.05$]. In contrast, the pitch-height component was significant for resolved adaptors only [$F(1,4) = 29.055$, $p < 0.01$], and not for unresolved adaptors [$F(1,4) = 1.693$, $p = 0.263$]. The model was also fitted to individual participants' data and values for the pitch-height term were entered into a one-sample *t*-test. The pitch-height term was

significantly different from zero for resolved adaptors only [$t(11) = 6.293$, $p < 0.001$; unresolved: $t(11) = 0.857$, $p = 0.410$].

Summary

Experiment 4 demonstrated that the amplitude of the PCR is sensitive to pitch chroma rather than consonance or dissonance. PCR amplitudes for unresolved adaptors were fit well with a model that contained a pitch-chroma component, plus a constant intercept. For resolved adaptors, the model fit was significantly improved by including a pitch-height component.

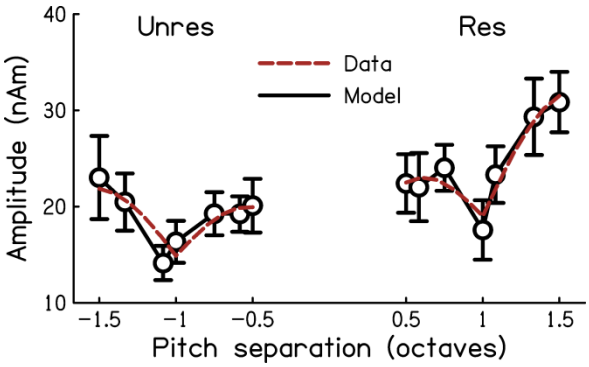

Figure 1.8. Mean PCR amplitudes for each pitch separation (black). Error bars represent one standard error. Fits of a simple function, incorporating a sine term to model the circular nature of pitch chroma and a linear slope to model the increase in pitch-height difference between the adaptor and probe, are shown by the dashed red lines.

1.F DISCUSSION

By using the adaptation paradigm, this study took advantage of the high temporal resolution afforded by EEG to examine the cortical representation of pitch. Neural responses to complex stimuli, which consist of multiple spectral components, were explicitly sensitive to pitch chroma. This finding did not arise from sensitivity to individual spectral frequencies, nor did it arise from sensitivity to musical consonance or dissonance. Latency differences suggested that the chroma-sensitive response was generated at a later stage of the processing hierarchy than the response to a spectral change. Further, the chroma-sensitive response localised to a region anterior and lateral to the spectrally-sensitive response. Tentatively, this suggests that the chroma-sensitive response came from the anterolateral pitch area identified previously with fMRI (Patterson et al., 2002; Penagos et al., 2004; Hall et al., 2006; Puschmann et al., 2009).

Functional role of pitch chroma

In a spiral representation of pitch, the dimensions of chroma and height are inter-dependent; one cannot alter pitch chroma without increasing or decreasing pitch height. However, the current study found that neural responses can be sensitive to pitch chroma without showing sensitivity to pitch height. This suggests that, at some stage in the processing pathway, height and chroma are represented independently. The result is consistent with that of Warren et al. (2003), who found that stimuli with variations in pitch height or pitch chroma activated partially distinct areas of auditory cortex (although they did not test the difference for statistical significance). The chroma-sensitive region lay slightly anterior to primary auditory cortex. This region might represent the start of a pathway involved in identifying sounds (Ahveninen et

al., 2006) and tracking them over time (Warren et al., 2003). In contrast, the height-sensitive region lay posterior to primary auditory cortex. It might represent the start of a pathway involved in processing sound location and segregating sounds into perceptual streams (Griffiths & Warren, 2002). The idea of pitch being processed in two distinct pathways could reconcile conflicting findings about the locus of pitch-sensitive neurons in the cortex; areas anterior and posterior to primary auditory cortex have been suggested in different studies (Hall & Plack, 2009). A role for pitch chroma in conveying sound identity is supported by the finding that native speakers of tonal languages (such as Mandarin or Cantonese) are more sensitive to pitch chroma than are speakers of non-tonal languages (such as English or French; Deutsch, 2006). In non-tonal languages, pitch conveys prosody and emotion, whereas in tonal languages, pitch also conveys the identity (meaning) of sounds. If pitch chroma conveys identity in music, one should be able to identify a tune in which notes are shifted by one or more octaves (preserving pitch chroma but changing pitch height). Deutsch (1969; 1972) found perfect identification when all notes were shifted similarly, but performance was at chance when successive notes were selected from different octaves. Whilst the latter result seems inconsistent with a role for pitch chroma in conveying tune identity, it might be expected if pitch height plays a role in sound segregation (Warren et al., 2003). This is because consecutive notes with a large pitch-height separation might be separated into different perceptual streams, which would make them more difficult to follow (Micheyl & Oxenham, 2010).

Formation of a chroma representation

For periodic sounds, such as vowels and musical notes, pitch is closely related to the repetition rate of the stimulus waveform (Yost, 1996; Yost et al., 1996).

However, this "temporal" cue occurs alongside a "spectral" cue, as the frequency spectrum of periodic sounds contains peaks at harmonics of the repetition rate. A great deal of debate has focussed on whether the auditory system uses spectral or temporal cues for pitch processing (de Cheveigné, 2005). The current study found sensitivity to pitch chroma, even when the spectral detail of a stimulus was unavailable to the auditory system (i.e. when stimuli were unresolved). This implies that pitch-chroma processing does not depend on spectral information. However, it does not rule out the possibility that spectral information could contribute to pitch-chroma processing when it is available (i.e. when stimuli are resolved).

The detailed temporal structure of a stimulus is largely preserved in the firing patterns of neurons in the auditory nerve. The action potentials ("spikes") of these neurons are "phase-locked" to local peaks in the stimulus waveform, up to rates of at least 4 kHz (Winter, 2005; Wang et al., 2008b). In temporal models of pitch perception, it is thought that neurons at later stages of the auditory pathway measure the time intervals between spikes and select those intervals that occur most often (Meddis & Hewitt, 1991a; 1991b; Yost et al., 1996). These intervals will correspond to the stimulus period and integer multiples of the period. The auditory system is assumed to use the shortest interval (corresponding to the stimulus period), which is the reciprocal of the stimulus repetition rate, to derive a measure of pitch. Therefore, in this model, pitch would be represented as a single dimension (repetition rate). Ohgushi (1983) proposed that the longer prominent time intervals (corresponding to integer multiples of the stimulus period) are also processed further and that octave similarity arises implicitly because many of these longer intervals are the same for notes separated by an octave (Ohgushi, 1983). However, this model could not account for neural sensitivity to pitch-chroma separations other than the octave. Patterson

(1986) proposed an alternative model, in which pitch chroma is extracted explicitly from the incoming spike trains, by a system that detects relationships amongst inter-spike intervals. He demonstrated that certain relationships are present in all notes that share the same pitch chroma.

The upper limit of phase locking decreases after the auditory nerve. Therefore, after the stimulus period has been extracted from the incoming spike train, it is presumably represented with a "place" code, in which different neurons signal different stimulus periods by their mean firing rates (Winter, 2005; Wang et al., 2008b). Potentially, pitch chroma could be derived from the place representation of the stimulus period, at a later stage of processing. Alternatively, if prominent inter-spike intervals other than the stimulus period are also represented with a place code, a pitch-chroma mechanism could detect relationships amongst these time intervals (cf. Patterson, 1986) by integrating responses from select groups of neurons.
Neuroimaging studies have suggested that pitch processing involves multiple stages of the auditory pathway, including brainstem structures, but that a complete pitch representation is not formed until non-primary auditory cortex (Griffiths et al., 1998; 2001; Patterson et al., 2002). It may be that the conversion of prominent time intervals to a place code, and the extraction of pitch height, occurs at an earlier processing site, whilst a complete representation of pitch chroma is formed in non-primary auditory cortex.

Remaining issues

The lack of a pitch height effect for unresolved stimuli is puzzling. There is an apparent effect for resolved stimuli, but it is unclear whether this is a genuine pitch effect, or whether it arises solely from adaptation in spectrally-sensitive neurons (see

Methods). If it is a genuine pitch effect, it would suggest that pitch-height processing requires spectral information. This would provide partial support for the theory that resolved and unresolved stimuli are processed by different neural mechanisms (Carlyon, 1998). If it is not a genuine pitch effect, it could be that pitch height is represented in a form that does not affect the overall size of the EEG signal. For example, some form of temporal representation could be used, such as the difference in response latency between neurons (Chase & Young, 2007). Alternatively, pitch height could be represented sub-cortically and thus be invisible to the cortical EEG measurements used in the current study.

A second issue is whether the auditory system contains a spatial map of pitch. This has been suggested for the cortex (Schulze & Langner, 1997; Schulze et al., 2002) and the inferior colliculus (Langner et al., 2002). However, these findings have been difficult to replicate. This may be because the studies did not fully dissociate the effects of spectral frequency from the effects of pitch (McAlpine, 2004). Unless controlled for, the frequency component at the repetition rate of the stimulus (termed the "fundamental frequency") could give rise to apparent pitch effects in neurons sensitive to individual spectral components. Even when this frequency is removed from a complex stimulus, it could be re-created in the cochlea as a "distortion product", due to the non-linear nature of cochlear processing (Pressnitzer & Patterson, 2001; McAlpine, 2004). Therefore, efforts must be made to mask the fundamental frequency. The current study achieved this by presenting a continuous masking noise over the full range of repetition rates. In addition, RI sounds were used, which have random phases and are thus unlikely to produce strong distortion products (Pressnitzer & Patterson, 2001). Whilst the current study found neural sensitivity for pitch chroma, this does not necessarily mean that neurons are spatially arranged according to their

preferred pitch-chroma value. Further, if there is a spatial map of pitch chroma, this map would not be expected to resemble the linear progression of the tonotopic map for spectral frequency (Eggermont & Roberts, 2004), or the retinotopic map for visual space. Rather, the topology of any pitch map might be similar to that of the pinwheel map of image orientation in visual cortex, with adjacent orientations arranged around a central point (Bonhoeffer & Grinvald, 1991), because, like orientation, pitch chroma is a cyclical perceptual dimension. A technique with a higher spatial resolution is needed to demonstrate a map with such a topology (e.g. Haynes & Rees, 2005).

Potential applications

Around 4% of the general population have life-long difficulties in appreciating and producing music (Kalmus & Fry, 1980), a condition referred to as "congenital amusia" (for reviews, see Peretz & Hyde, 2003; Pearce, 2005). These people often find music unpleasant, and seek to avoid social situations that involve music. It is currently unclear whether their difficulties stem from a problem with fine-grained pitch perception (Peretz et al., 2002) or a problem with a more general cognitive mechanism (e.g. spatial processing; Douglas & Bilkey, 2007). The paradigm used in the current study could provide a powerful tool for dissociating these explanations. This is because it provides a direct measure of the neural representation of pitch, unconfounded by attention or task requirements. If amusia reflects a pitch-processing deficit, one might expect differences in pitch representation between amusics and a control group (such as broader neural tuning for pitch in amusics). In contrast to amusia, a much smaller number of people possess "absolute pitch" (or "perfect pitch"), a remarkable ability to recognise or produce a note without any form of reference (for a review, see Levitin & Rogers, 2005). This ability primarily concerns

pitch chroma, as people with absolute pitch can still misjudge a note's octave (Bachem, 1937). It may thus be that the neural representation of pitch chroma is altered in people with absolute pitch. The current paradigm could be used to test this hypothesis.

Finally, it is unknown whether the neural representation of pitch chroma is innate or whether it develops with musical exposure. Support for the innate hypothesis would come from the demonstration that very young infants can perceive pitch chroma. A study by Demany & Armand (1984), which found that three-month-old infants treat a note and its octave as similar, is consistent with the innate hypothesis, although the result could be due to the consonance of the octave interval, rather than pitch chroma per se (see Zentner & Kagan, 1998). A recent EEG study by He and Trainor (2009) found that a pitch change elicited a mismatch negativity (MMN) component in 4-month-old infants but not in 3-month-old infants. However, the MMN is thought to reflect a mechanism involved in detecting changes in regular aspects of the acoustic input (Winkler et al., 1996; Schröger, 1997). He and Trainor's (2009) results could reflect the development of the regularity mechanism, rather than the development of a neural pitch representation. In contrast, the current paradigm provides a direct measure of pitch representation, which could be used to study pitch processing in very young infants.

Chapter 2. Adaptation sharpens frequency tuning in human auditory cortex.

2.A INTRODUCTION

A ubiquitous feature of neural responses is their dependence on stimulus context. The neural response to a probe stimulus is typically suppressed by a preceding stimulus, an effect known as "repetition suppression", or "adaptation" (see Grill-Spector et al., 2006, for a review). Adaptation is usually measured with discrete sequences consisting of a single adaptor stimulus followed by a probe (Fig. 2.1B). This paradigm has been widely used in research on the visual system to make inferences about how stimulus attributes are represented in the cortex (Grill-Spector et al., 2006). This is because suppression of the probe response is thought to be greatest when the adaptor and probe activate similar neural groups. Adaptation is also prevalent in the auditory system (Ulanovsky et al., 2004). Indeed, the study reported in Chapter 1 was able to use the adaptation paradigm to make inferences about how musical pitch is represented in the auditory cortex. However, it is generally thought that the auditory system possesses an additional contextual mechanism that might serve a special early-warning role by responding to changes in the acoustic environment. This may be important because the visual system can be remarkably poor at detecting changes outside the current focus of attention (the phenomenon of "change blindness"; Simons & Chabris, 1999).

The additional mechanism is thought to extract regular aspects of the auditory input and then elicit a response when a stimulus deviates from these regularities (Winkler et al., 1996; Schröger, 1997). The response is called the "mismatch negativity", or "mismatch response" (MMR). The current report will use the latter term, as the response is not always monophasic and surface negative. For example, bi- and tri-phasic MMRs were found for a number of stimulus types in a report by

Näätänen et al. (2004; Fig. 2), and Novitski et al. (2007) found predominantly positive MMRs in infants. The MMR is usually measured with the oddball paradigm (Fig. 2.1A), in which infrequent "deviant" stimuli are presented in a sequence of frequent "standards". Deviants typically elicit a larger neural response than standards. This difference is the MMR. The MMR is thought to be distinct from adaptation in two ways: (1) It is thought to reflect an active, rather than a passive, mechanism. (2) It is thought to be separate from the sensory-driven response. However, Jääskeläinen et al. (2004) recently challenged these assumptions by demonstrating a significant MMR for probe responses in the adaptation paradigm. As a single adaptor cannot setup stimulus regularity, Jääskeläinen et al. (2004) argued that the MMR reflects a passive release from adaptation, rather than an active deviance-detection mechanism unique to the auditory system. They proposed that the MMR arises in the oddball paradigm because responses to infrequently-presented stimuli are less adapted than responses to frequently-presented stimuli. However, whilst adaptation may explain the MMR to probes preceded by a *single* adaptor, it remains unclear whether adaptation is sufficient to explain the MMR to probes preceded by *multiple* adaptors, as in the oddball paradigm. It may be that multiple adaptors invoke an additional deviance-related process.

The current study addressed this question by generating a quantitative estimation of the contribution of adaptation to the MMR for probes preceded by multiple adaptors based on the MMR for probes preceded by a single adaptor. In this way, the study sought to bridge the gap between the adaptation paradigm and the oddball paradigm. The experiment was designed such that it would allow us to compare the stimulus specificity of the adaptational effect with the stimulus

specificity of any deviance-related effect, and also to investigate whether this effect reflects endogenous or sensory-driven activity.

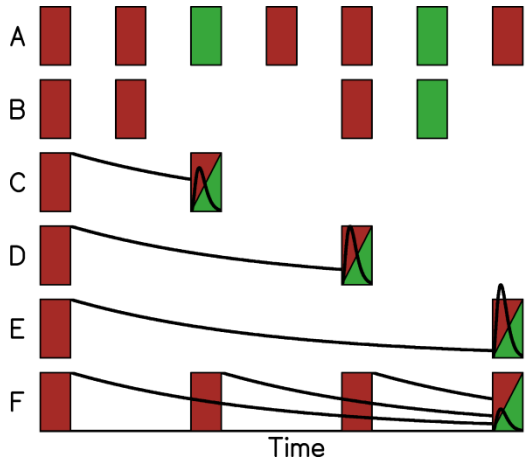

Figure 2.1. A: In the oddball paradigm, infrequent "deviant" stimuli (green) are presented amongst frequent "standard" stimuli (red). B: In the adaptation paradigm, stimuli are presented in pairs, an "adaptor" stimulus followed by a "probe" stimulus. The adaptor and probe may be identical (red-red), or may differ (red-green). C-F: The time course of adaptation can be measured by varying the SOA between the adaptor and probe in the adaptation paradigm. In a simple adaptation model, the combined effects of multiple adaptors on the probe response (F) would be assumed to correspond to the added effects of each individual adaptor separately (C-E).

2.B METHODS

In the simplest model of adaptation, the combined effects of multiple adaptors on the response to a probe stimulus would be assumed to correspond to the sum of the separate effects of each individual adaptor. For example, the probe response after three adaptors (Fig. 2.1F) could be predicted by presenting the first, second and third adaptors separately (Figs 2.1C-E) and then combining the observed adaptation effects. To derive a more general equation for adaptation after multiple adaptors, one must first describe the time course of adaptation. The first part of this study measured the time course of adaptation by varying the stimulus-onset asynchrony (SOA) between pairs of pure tones (adaptor and probe), for four different frequency separations. These measurements were used to predict probe responses in the second part of the study, in which either two or three adaptors were presented before the probes.

Stimuli

Pure tones of 110-ms duration (including 10-ms cosine-squared onset and offset ramps) were generated using a TDT System 3 (Tucker Davis Technologies, Florida) with 24-bit amplitude resolution and a 24.4-kHz sampling rate. Stimuli were presented diotically via headphones (K 240 DF; AKG, Vienna) at 60 dB SPL. The experiment was controlled by a Matlab (The MathWorks Inc., Natick) program, which passed stimulus conditions to the TDT system via the RP2 ActiveX interface. To equate the tones for loudness, masker noise, filtered such that equal energy (20 dB SPL) would fall into each peripheral auditory filter bandwidth (ERB_N; Glasberg & Moore, 1990), was played continuously.

Adaptation measured for a single adaptor

Probe tones were preceded by a single adaptor, with a stimulus onset asynchrony (SOA) of 125, 250, 500 or 1000 ms and an inter-trial onset asynchrony of 5 seconds. On each trial, the frequency of the probe tone was chosen randomly from a 1/3-octave range around 1000 Hz. The adaptor frequency was either the same as, or 1/6-octave, 1/2-octave or 1.5-octaves above, the probe frequency.

The 500- and 1000-ms SOAs were presented randomly within the same session, whilst the 125 and 250 ms SOAs were presented in separate sessions. In the 125- and 250-ms sessions, probe tones were omitted on half the trials. This was because neural responses to the adaptor and probe partially overlap at these short SOAs. The "probe absent" trials measured undistorted adaptor responses; these were subsequently subtracted from the "probe present" trials to obtain undistorted probe responses. This was done separately for each frequency separation. This approach is similar to the Adjar procedure (Woldorff, 1993) except that it uses a direct measurement of the overlapping response rather than an estimate calculated from the data. Each session was 80-minutes long, split into four runs of 20 minutes each. During recording, participants watched a subtitled DVD of their choice.

Mismatch response measured for multiple adaptors

On each trial, either two or three adaptors preceded the probe. The SOA between stimuli was always 500 ms. Trials were randomly presented within a single 80-minute session, split into four runs of 20 minutes each. In all other respects, stimulus parameters were the same as in the one-adaptor sessions. When the adaptor and probe had the same frequency, the probe was said to be "standard", when they had different frequencies the probe was said to be "deviant".

Electroencephalography and data pre-processing

The recording parameters, and the data pre-processing steps, were mostly identical to those used in Experiments 1, 2 and 4 of Chapter 1. Briefly, neural responses were recorded from 32 Ag/AgCl electrodes, arranged according to the 10% system. Signals were amplified and band-pass filtered (0.1-250 Hz), then sampled at 500 Hz and stored for offline analysis. Using EEGLAB software (Delorme & Makeig, 2004), recordings were lowpass filtered at 35 Hz, down-sampled to 250 Hz per channel and re-referenced to average reference. The data were segmented into epochs ranging from 100 ms before the first stimulus of a trial to 500 ms after the last stimulus, and responses were baseline-corrected to the 100 ms pre-stimulus interval. Epochs containing unusually large potentials (joint probability limits of ± 3 SD) were automatically rejected and stereotyped artefacts were removed by manual inspection following a statistical decomposition of the data into maximally independent components. The remaining epochs were averaged to obtain an ERP for each participant and condition. Event-related potentials were imported into BESA (MEGSIS Software GmbH, Munich) and source waveforms were derived from a two-dipole model, with one dipole at the centre of mass of primary region TE1.0 in each hemisphere (Morosan et al., 2001). The orientations of the two dipoles were fitted to a window encompassing the P1, N1 and P2 peaks of the mean probe response (0 to 250 ms after stimulus onset), averaged across participants and conditions. Subsequently, source waveforms were averaged across hemispheres.

Data analysis

Probes elicited typical onset responses with peaks around 56, 108 and 188 ms, corresponding to the P1, N1 and P2 (Näätänen and Picton, 1987). As in Chapter 1,

both the N1 and the P2 were modulated by the stimulus parameters - in this case, the adaptor-probe frequency separation and the SOA (Fig. 2). As noted in Chapter 1, if a manipulation suppresses the P2 more than it suppresses the N1, the N1 could appear to increase, solely because it is less cancelled by the P2. Indeed, studies that have claimed to observe enhancement of the N1 response at very short SOAs appear to be confounded by a substantial decrease in the size of the P2, leading also to a wider and larger N1 peak (for example, see Budd & Michie, 1994; Wang et al., 2008a). To demonstrate the cancelling effect, P2 values from each event-related potential were flipped in sign and correlated with the corresponding N1 values. As expected, peak amplitudes were negatively correlated [$r(180) = -0.371$, $p < 0.001$]. Therefore, as for Chapter 1, and in line with earlier work that found meaningful relationships between the auditory potentials and the frequency separation or SOA between stimuli (for example, Davis et al., 1966; Butler, 1968; Nelson & Lassman, 1968), responses were quantified as the peak-to-peak amplitude from the N1 to the P2. Subsequently, probe responses were expressed as a proportion of their unadapted amplitude and converted to percent adaptation, using the formula:

$$a_{\Delta f,t} = \left(1 - \frac{probe_{\Delta f,t}}{adaptor_{\Delta f=0,t}}\right) \times 100$$

where Δf is the adaptor-probe frequency separation and t is the adaptor-probe SOA. Note that the response to the adaptor, when it had the same frequency as the probe, is used as a measure of the unadapted response. In sequences with more than one adaptor, the first adaptor was used for the measure of the unadapted response, whilst in the one-adaptor conditions with the 250-ms or 125-ms SOAs, the undistorted adaptors from the "probe-absent" trials were used.

Bootstrap confidence intervals and permutation tests

A sustained response (SR) was observed during the two- and three-adaptor conditions, and during the one-adaptor condition at the two longest SOAs, where the adaptor and probe responses did not overlap. To compare the neural locus of this SR with that of the prominent N1 sensory-driven response, equivalent current dipoles were fitted to the grand average N1 and SR voltage maps using BESA. Confidence intervals were computed for the dipole locations using a bootstrap procedure, which involved fitting dipoles to 1000 between-subjects re-samples of the N1 and SR voltage maps (Efron, 1993). This yielded a bootstrap distribution for each dipole parameter, from which 95%-*t*-confidence intervals were obtained. This procedure is similar to that used in Chapter 1. However, in Chapter 1, dipole co-ordinates were initially obtained for each participant, and the bootstrap distribution was formed from re-samples of these co-ordinates. The advantage of the current method is that it allows dipole fitting to be performed on voltage maps with high signal-to-noise ratios, despite fewer stimulus presentations in the current study, because each map will be the average across a number of participants. Higher signal-to-noise ratios lead to more accurate source localisation (van Hoey et al., 2000; Wang & Gotman, 2001). Within-subjects permutation tests were used to compare dipole locations and orientations between the N1 and SR. Again, the procedure was similar to that used in Chapter 1 except that *voltage maps* were permuted within participants, rather than dipole co-ordinates. Each point (n = 1000) in a permutation distribution was computed by averaging across voltage maps and calculating a difference between the dipoles that were fitted to these maps.

Subjects

Sixteen subjects (mean age of 24 years, range 18-40 years; five males) participated in this study. Eleven (for all conditions except one-adaptor, 250 ms SOA) or twelve participated in each of the experimental sessions. All subjects had hearing thresholds at or below 20 dB HL at octave frequencies between 0.25 and 8 kHz and had no history of audiological or neurological disease. This study was approved by the Ethics Committee of the University of Nottingham School of Psychology.

2.C RESULTS

Adaptation measured for a single adaptor

For four frequency separations, the time course of adaptation was measured by presenting pairs of stimuli (adaptor and probe) with different SOAs. Neural responses to the probes were similar in morphology to the adaptor responses, but were reduced in amplitude. This reduction, quantified as percent adaptation, was greater at shorter SOAs and smaller frequency separations (Fig. 2.2). On a linear ordinate, adaptation decayed more rapidly for small frequency separations than for larger frequency separations (Fig. 2.3A). As a result, the spread of the four functions (the effect of frequency separation) appeared larger at shorter, than at longer, SOAs. The reduction in the size of adaptation when the adaptor differs from the probe, as compared to when the adaptor is identical to the probe, is often taken as a measure of the stimulus specificity of the stimulated cortex (Grill-Spector & Malach, 2001). Under this assumption, it would appear that the frequency specificity of the auditory system changes in the period following an adaptor. However, this assumption ignores the shape of the adaptation decay function; for all frequency separations, adaptation decayed roughly exponentially with SOA. An exponential process is expected based on studies that attribute adaptation to the depletion of neurotransmitter pools at stimulated synapses ("synaptic depression"; Chung et al., 2002; Freeman et al., 2002; Wehr & Zador, 2005; Asari & Zador, 2009). Therefore, when adaptation is plotted on a logarithmic ordinate, the decay functions appear linear (Fig. 2.3B). Critically, the functions for different frequency separations now also appear parallel. This indicates that the frequency specificity of the auditory system actually stays fixed after a single adaptor and that, regardless of frequency separation, the decay of adaptation can be described by a single exponential time constant. This has important implications for

studies attempting to use the adaptation paradigm to investigate stimulus specificity. If the time course of adaptation is not taken into account, experiments might underestimate or overestimate stimulus specificity.

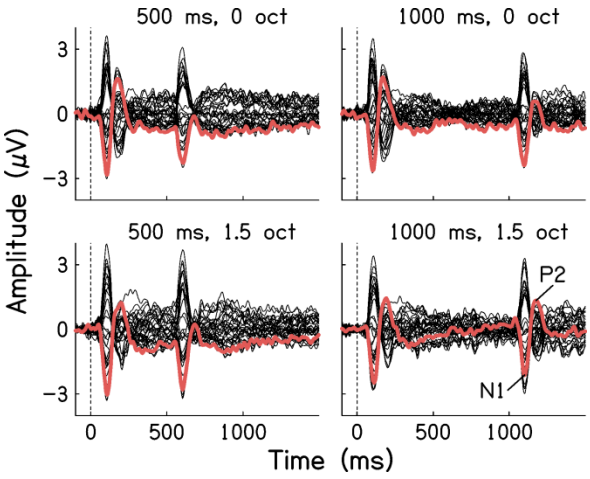

Figure 2.2. Mean event-related potentials in four conditions. Each black line is the response from one recording electrode; the vertex electrode (Cz) is highlighted in red. On the time axis, zero marks the start of the adaptor stimulus. The adaptor elicited a triphasic neural response, which was followed by a similar, but smaller, response to the probe.

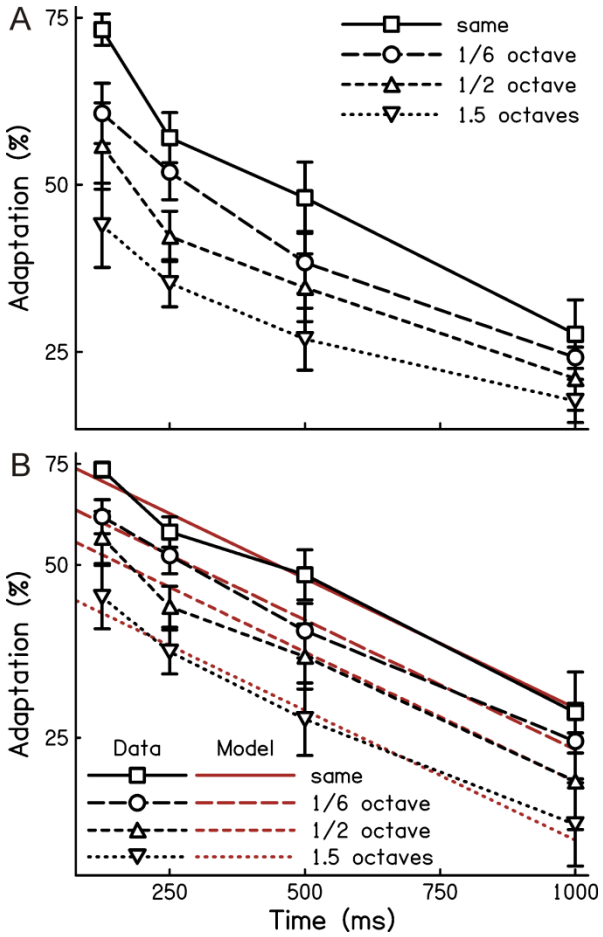

Figure 2.3. The decay of adaptation over time, plotted on a linear (A) and a logarithmic (B) ordinate. The parameter is the frequency separation between the adaptor and probe. Error bars represent one standard error. The data can be described by an exponential decay function, with a single time constant, but a different intercept for each frequency separation (red lines in panel B).

These observations were confirmed statistically using a linear mixed model, performed on the log-transformed data, with frequency separation entered as a fixed factor and SOA entered as a covariate. In essence, the statistical model fits a regression line through the four SOA points for each frequency separation and compares these regression lines across the four frequency separations. Importantly, the interaction of frequency separation and SOA was non-significant [$F(3,154.677) = 0.267$, $p = 0.849$], indicating that the regression lines can be regarded as parallel. Both the main effects of frequency separation [$F(3,157.670) = 10.603$, $p < .001$] and SOA [$F(1,169.495) = 112.111$, $p < .001$] were significant. In pairwise comparisons between the frequency separations, all differences were significant ($p < .05$) except that between the 0- and 1/6-octave conditions and that between the 1/6- and 1/2-octave conditions. This confirms that the four regression lines are parallel with different intercepts.

Therefore, the time course of adaptation was modelled using the exponential function:

$$a_{\Delta f,t} = e^{mt + b_{\Delta f}}$$

where t is the time in seconds after the onset of the adaptor and Δf is the frequency separation between the adaptor and the probe. The time constant, m, was obtained by fitting a straight line through the four SOA points, taken from the log-transformed data averaged across participants and frequency separations. Intercepts were then fitted separately for each frequency separation. The intercept, $e^{b_{\Delta f}}$, was 79.75%, 67.46%, 59.29% and 46.91% for the 0-octaves, 1/6-octave, 1/2-octave, and 1.5-octaves separations respectively. The time constant, m, was -1.04 s^{-1}, corresponding to a half-life of 665 ms. This is similar to the time constant of synaptic depression observed in neurons of rodent primary auditory cortex (Wehr & Zador, 2005; Asari &

Zador, 2009). The model fit the data well (Fig. 2.3B, red lines), explaining over 96% of the variance for all frequency separations.

The mismatch response measured for multiple adaptors

The time course of adaptation, measured with single adaptors, was used to estimate the contribution of adaptation to probe responses after two and three adaptors. Estimates were made for each of the four frequency conditions: same frequency, when the probe was standard, and the 1/6-, 1/2- and 1.5-octave separations, when the probe was deviant. In the simplest model of adaptation, the combined effects of multiple adaptors on the probe response would be assumed to correspond to the product of the separate effects of each adaptor presented individually. Thus, for 3 adaptors with onset-onset intervals of 500 ms, the size of the probe response, $p_{f\Delta}$, can be predicted using the formula:

$$p_{f\Delta} = 100 \times \left(1 - \frac{a_{\Delta f,t=1.5}}{100}\right) \times \left(1 - \frac{a_{\Delta f,t=1.0}}{100}\right) \times \left(1 - \frac{a_{\Delta f,t=0.5}}{100}\right)$$

where $a_{\Delta f,t}$ is percent adaptation, derived from the previous equation.

When the probe had the same frequency as the adaptors (i.e. when it was standard), the predicted response sizes matched the observed response sizes practically perfectly, with deviations of only 0.25% for two and 0.17% for three adaptors (Fig. 2.4). Remarkably, this demonstrates that the effects of multiple adaptors on the probe response simply accumulated when the stimulus frequency remained constant. Despite the first adaptor reducing the response to the second adaptor, it did not affect the second adaptor's ability to suppress the probe response. Thus, there seems to be a dissociation between the response elicited by a stimulus and the amount of adaptation caused by that stimulus. A similar dissociation has been

observed previously in neurons of the primary visual cortex (Carandini et al., 2002; Freeman et al., 2002) and inferior colliculus (Brimijoin & O'Neill, 2010): stimuli that are poor at eliciting neural responses themselves can still have considerable suppressive effects on responses to subsequent stimuli. This has also been found psychophysically. The amount of perceptual masking caused by an auditory stimulus remains the same even when the stimulus is rendered inaudible by masking from a preceding sound (Plack et al., 2006).

When the probe was a different frequency from the adaptors (i.e. when it was deviant), the predicted response sizes substantially underestimated the actual response sizes (Fig. 2.4). This means that the effects of multiple adaptors on the probe response did not simply accumulate when there was a change in stimulus frequency. Instead, there was an additional contribution to deviant responses, which depended on the preceding adaptor stimuli being presented in a sequence. This finding was confirmed statistically by subtracting the predicted response sizes from each participant's observed response sizes. This gives prediction errors, which were submitted to a linear mixed-model analysis. A significant main effect of frequency separation was found [$F(3,73) = 11.863$, $p < 0.001$]; the prediction errors for all frequency separations were significantly greater than the prediction error for the same-frequency condition ($p < 0.005$). The number of adaptors (two or three) did not have any significant effect on the prediction error [$F(1,73) = 0.233$, $p = 0.631$] and the interaction of the number of adaptors and the frequency separation was also not significant [$F(3,70) = 0.174$, $p = 0.914$].

For a given condition, if the actual response sizes deviate systematically from the predicted response sizes, the prediction errors will be significantly different from zero. As there was no main effect of the number of adaptors on prediction error, or an

interaction between the number of adaptors and the frequency separation, prediction errors were averaged across the number of adaptors (two and three). Prediction errors for the four frequency separations were then compared against zero using one-sample *t*-tests. The prediction error was not significantly different from zero for the same-frequency condition [$t(10) = 0.024$, $p = 1$], approached significance for the 1/6-octave condition [$t(10) = 2.354$, $p = 0.152$] and was significant for both the 1/2-octave and 1.5-octaves conditions [$t(10) = 4.865$, $p < 0.005$; $t(10) = 4.828$, $p < 0.005$]. The above *p*-values have been adjusted for multiple comparisons using the Šidák method (Abdi, 2007).

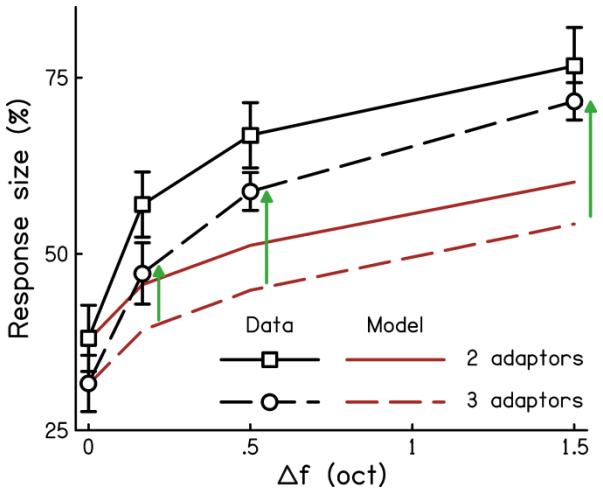

Figure 2.4. Probe response sizes after two or three adaptor stimuli, plotted as a function of frequency separation. Error bars represent one standard error. Predictions based on measurements with a single adaptor are shown in red. When the adaptor and probe differed in frequency, the predictions underestimated the size of the probe response. The prediction error increased with increasing frequency separation (green arrows).

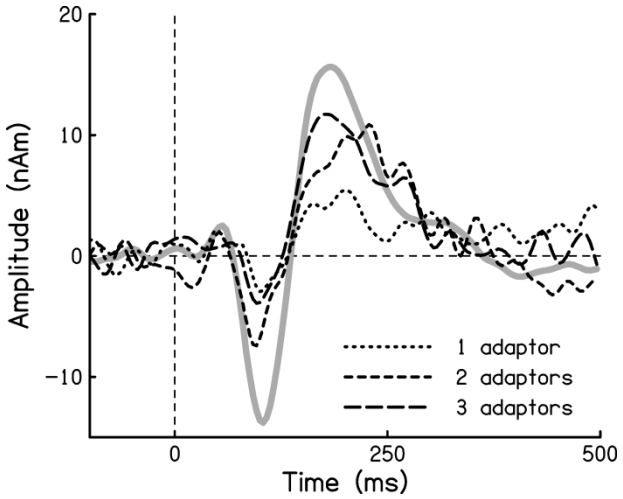

Figure 2.5. Difference waveforms obtained by subtracting the response to the standard probe from the mean responses to deviant probes (averaged across frequency separations). The parameter is the number of adaptors preceding the probe. For comparison, the adaptor response (taken from "probe absent" trials in the 125- and 250-ms SOA conditions) is plotted in grey, at half its original amplitude. Note that the difference waveforms contain peaks and polarities coincident with those of the adaptor response.

The nature of the additional response to deviant stimuli

Neural responses to deviant probes after multiple adaptors were significantly larger than predicted based on responses to probes after single adaptors. In principle, this additional response to deviant stimuli could arise either from an increase in the sensory-driven response or from superposition of an additional, endogenous, response. To dissociate these two possible explanations, difference waveforms were calculated by subtracting standard probe waveforms from deviant probe waveforms. Difference waveforms were computed for the one adaptor (500 ms SOA), two adaptor and three adaptor conditions. If the additional response to deviant stimuli is due to a superimposed endogenous response, then this endogenous response should emerge in the difference waveform. If the additional response is due to an increase in the sensory-driven response, then the difference waveform should reflect the sensory-driven response. Specifically, the difference waveform should be similar in morphology to the adaptor response.

Often, difference waveforms such as these are used to display the MMR, but only the negative portion of the waveform is analysed further. However, the difference waveforms measured in the current study appeared triphasic, with peaks and polarities consistent with the original sensory-driven response (Fig. 2.5). This was the case for difference waveforms after one, two and three adaptors, though the waveform after a single adaptor was smaller than that after multiple adaptors. To confirm that the MMR is multiphasic, rather than monophasic and negative, peak amplitudes were extracted from the difference waveforms using 40-ms windows centred on the N1 and P2 peaks in the adaptor response. These peaks will be termed the dN1 and dP2 respectively. The adaptor response was taken from the "probe absent" trials in the 250-ms and 125-ms SOA conditions, when the adaptor had the

same nominal frequency as the probe. Collapsing across deviant frequency separations, the dN1 was significantly different from zero after two and three adaptors [one adaptor, $t(10) = 1.969$, $p = 0.383$; two adaptors, $t(10) = 5.367$, $p < 0.005$; three adaptors, $t(10) = 3.300$, $p < 0.05$]. The dP2 was significantly different after one, two and three adaptors [one adaptor, $t(10) = 5.146$, $p < 0.005$; two adaptors, $t(10) = 6.900$, $p < 0.001$; three adaptors, $t(10) = 5.691$, $p < 0.005$]. The above *p*-values are Šidák-adjusted for multiple comparisons.

Amplitudes for the dN1 and dP2 were entered into linear mixed models to test for differences between conditions, with fixed factors of frequency separation (1/6 octave, 1/2 octave and 1.5 octaves) and number of adaptors. The number of adaptors had a significant effect on the size of the dN1 [$F(2,87.226) = 4.047$, $p < 0.05$], as did the frequency separation [$F(2,80.171) = 3.729$, $p < 0.05$]. The dN1 was larger after two adaptors than after one adaptor or three adaptors ($p < 0.05$). It was also larger for the 1.5-octave frequency separation than the 1/6-octave frequency separation ($p < 0.01$). The interaction of the number of adaptors and the frequency separation was not significant [$F(4,76.164) = 0.120$, $p = 0.975$]. Similarly, the number of adaptors had a significant effect on the size of the dP2 [$F(2,90.546) = 5.112$, $p < 0.01$], as did the frequency separation [$F(2,83.177) = 4.357$, $p < 0.05$]. The dP2 was larger after three adaptors than after one adaptor ($p < 0.005$). The dP2 after three adaptors was also larger than the dP2 after two adaptors, although this effect did not quite reach significance ($p = 0.058$). The dP2 was larger for the 1.5-octave frequency separation than the 1/6-octave ($p < 0.01$) or 1/2-octave separations ($p < 0.05$). The interaction of the number of adaptors and the frequency separation was not significant [$F(4,79.212) = 0.514$, $p = 0.726$].

In summary, the additional response to deviant probes after multiple adaptors appears to reflect a modulation of the sensory-driven response rather than the superposition of an additional, monophasic negativity. MMRs were at least biphasic, with both peaks affected by the number of adaptors and the frequency separation between the adaptors and the probe. Note that biphasic MMRs have also been frequently observed in previous studies (see, for example, Fig. 2 in Näätänen et al., 2004).

Sustained response

Most MMR studies present stimuli in continuous sequences. This means that each response must be baseline-corrected separately and any ongoing activity that persists throughout a sequence will be lost. As the current study used discrete sequences with silent intervals in between, responses could be baseline-corrected to a period before the start of each sequence. Surprisingly, doing this revealed a sustained response (SR) that lasted throughout a stimulus sequence (Fig. 2.6A). It began shortly after the onset response to the first adaptor in a sequence and ended shortly after the probe response. Onset responses to stimuli after the first adaptor rode on the SR. The SR was found in all conditions except those where the onset response to the adaptor partially overlapped the onset response to the probe (one adaptor, 125-ms or 250-ms SOA).

For each participant and condition, response waveforms were baseline corrected to the 100-ms period before the first adaptor and averaged across frequency separations. The SR was measured by taking the mean amplitude in a 200-ms time window before the onset of the probe; the N1 was measured by taking the mean amplitude in a 40-ms time window centred on the first adaptor. To confirm the

presence of a SR, amplitudes at the vertex electrode (Cz) were compared against zero using one-sample *t*-tests; all were found to be significant [one adaptor, 500-ms SOA: $t(10) = 4.052$, $p < 0.01$; one adaptor, 1000-ms SOA: $t(10) = 6.220$, $p < 0.001$; two adaptors: $t(10) = 3.986$, $p < 0.05$; three adaptors: $t(10) = 4.754$, $p < 0.005$; Šidák-adjusted *p*-values].

To compare the amplitude of the SR across conditions, Cz amplitudes were submitted to two linear mixed models. The first examined the effect of SOA, comparing the one-adaptor 500-ms and 1000-ms SOA conditions; SOA was found to be non-significant [$F(1,20) = 0.917$, $p = 0.350$]. The second examined the effect of number of adaptors, comparing the one-adaptor 500-ms SOA condition, the two-adaptor condition and the three-adaptor condition; number of adaptors was also non-significant [$F(2,20.629) = 0.716$, $p = 0.501$].

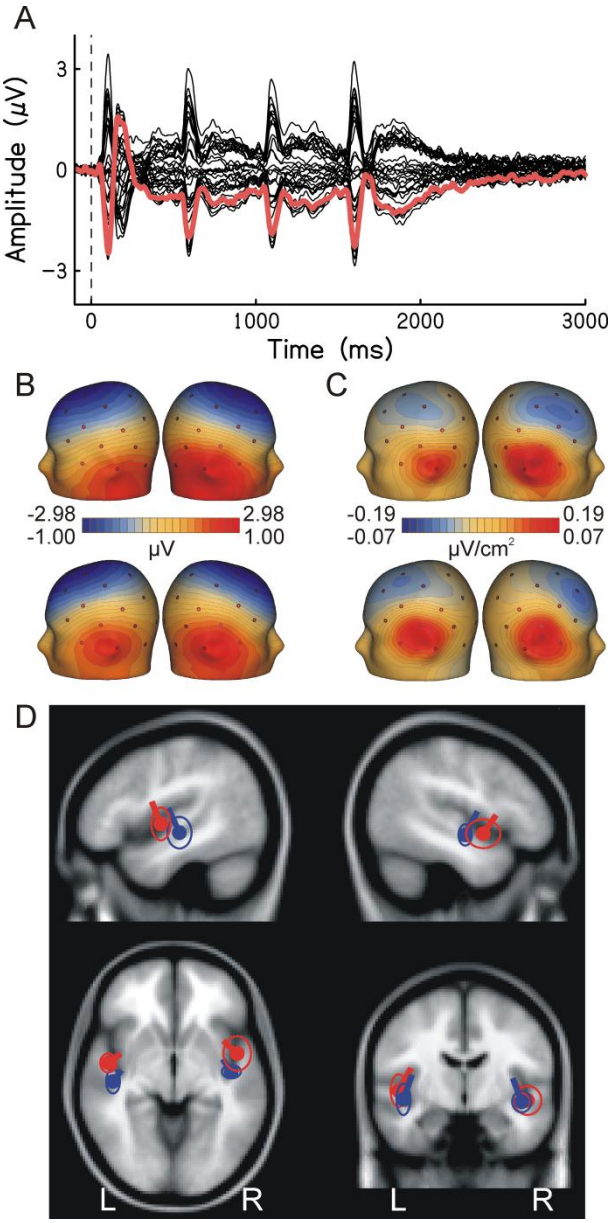

Figure 2.6. A: Mean event-related potentials in the three-adaptor condition (averaged across frequency separations). Each black line is the response from one recording electrode; the vertex electrode (Cz) is plotted in red. The SR was manifest as a sustained displacement from baseline voltage that began after the response to the first stimulus and finished shortly after the response to the probe stimulus. B-C: Field potential (B) and current source density (C) maps for the N1 (top) and SR (bottom). D: Dipole source locations and orientations for the N1 (blue) and SR (red). Data are plotted in the sagittal plane (top), the transversal plane (bottom left) and the coronal

plane (bottom right). Ellipses show 95% confidence intervals for each dipole location. The SR dipoles localised anterior to the N1 dipoles in both hemispheres. Orientation differences were also apparent in the right hemisphere.

Thus, throughout each stimulus sequence there was a SR, which appeared insensitive to the number of adaptors in the sequence or the SOA between stimuli. This SR may be a correlate of the monitoring process that initiates a call for attention when a repetitive stimulus is changed (Näätänen, 1992). To examine whether the neural locus of the SR differed from that of the sensory-driven response, voltage maps of the SR and N1 were produced for each participant, averaged across conditions. Like the N1 field map, the SR map displayed a vertex negativity and temporal positivity (Fig. 2.6B). However, the negative pole of the SR map was anterior to that of the N1 map, and the positive pole of the SR map was both anterior and superior to that of the N1 map. This suggests a difference in the location and/or orientation of the underlying generators of the N1 and SR. This is supported by maps of scalp current density (SCD; Fig. 2.6C), which are better able to separate responses from the left and right hemispheres (Nunez & Srinivasan, 2006). Both the N1 and SR maps showed two pairs of sinks and sources, one in each hemisphere. In line with the voltage maps, the sinks and sources appeared anterior in the SR maps than the N1 maps, suggesting an anterior shift in the source location. However, in the right hemisphere, the anterior shift in the negative pole was more pronounced than the shift in the positive pole, suggesting a rotation of the SR dipole also.

To examine whether the N1 and SR were generated in significantly different regions of cortex, grand-average field maps were calculated by averaging over participants. Two unconstrained equivalent current dipoles were then fit to the SR and N1 maps separately (Fig. 2.6D); a between-subjects bootstrapping procedure was used to produce confidence intervals for the dipole locations and within-subjects permutation tests were used to test for significant differences between the dipole locations and orientations (see Methods).

In each hemisphere, the Euclidean distance between the N1 and SR dipole was significant ($p < 0.001$). The SR dipole was significantly anterior to the N1 dipole in each hemisphere (left: $p < 0.005$; right: $p < .0001$). Differences in the medial/lateral (left: $p = 0.202$; right: $p = 0.189$) and superior/inferior (left: $p = 0.077$; right: $p = 0.816$) dimensions were not significant. To examine differences in orientation, the central angle between the N1 and SR dipoles was computed. The size of the central angle was significant in the right hemisphere ($p < 0.05$) but not the left hemisphere ($p = 0.088$). The orientations of the right hemisphere dipoles were then projected into the sagittal, coronal and transversal planes and their angles compared using circular statistics (Berens, 2009). In the sagittal plane, the SR orientation was significantly less vertical than the N1 orientation ($p < 0.05$). In the coronal plane, the SR orientation was significantly more radial than the N1 orientation ($p < 0.01$). The latter effect was also apparent in the left hemisphere.

Adaptation sharpens frequency tuning after multiple adaptors

The additional response to deviant probes after multiple adaptors appears to reflect a modulation of the sensory-driven response to the probes. Whilst adaptation from multiple adaptors combined in an independent manner when the probes were standard, responses to deviant probes were influenced by effects that depended on the adaptors being presented in a sequence. Therefore, the adaptation model, which was derived from probe responses after only a single adaptor, under-predicted the response sizes to deviant probes. The size of the prediction error increased with increasing frequency separation between the adaptor and probe (compare the green arrows in Fig. 2.4); the prediction error for the 1.5-octave separation was significantly larger than the prediction error for the 1/6-octave separation (linear-mixed model; $p < 0.05$).

The implication of this sequential effect can be observed when adaptation is plotted as a function of the adaptor-probe frequency separation (Fig. 2.7). To account for the exponential decay of adaptation over time, each function was normalised by its value when the adaptor and probe had the same frequency:

$$a_{norm} = \frac{a_{\Delta f}}{a_{\Delta f = 0}}$$

This is equivalent to a linear shift along a logarithmic ordinate. With this transformation, the frequency-tuning functions for different SOAs after a single adaptor overlapped, as expected. This was confirmed statistically by comparing the area under each of the tuning curves (a sharper tuning curve will have a smaller area). This method avoids any assumptions about the shape of the tuning curves. When the four SOA conditions were submitted to a linear mixed model, the effect of SOA was not significant [$F(3,23.077) = 0.119$, $p = 0.948$].

The two- and three-adaptor conditions did not differ significantly from each other [$F(1,7.562) = 0.347$, $p = 0.573$]. However, both appeared considerably sharper than the functions for the one-adaptor conditions. Therefore, a further linear mixed model compared the tuning curve areas between the single- and multiple-adaptor conditions. The effect was significant [$F(1,47.023) = 6.794$, $p < 0.05$], indicating that the tuning curve for the multiple-adaptor conditions was sharper than the tuning curve for the single-adaptor conditions. This means that the frequency specificity of adaptation increased after multiple adaptors. In other words, adaptation was relatively less effective at suppressing the response to deviant probes after multiple than single adaptors. In the one-adaptor conditions, the amount of adaptation for the largest frequency separation was around 60% of the amount for no frequency separation (60%, 62%, 56% and 64% for the 125 ms, 250 ms, 500 ms and 1000 ms SOAs

respectively). After two or three adaptors, this amount was around 40% (38% and 42% for two and three adaptors respectively).

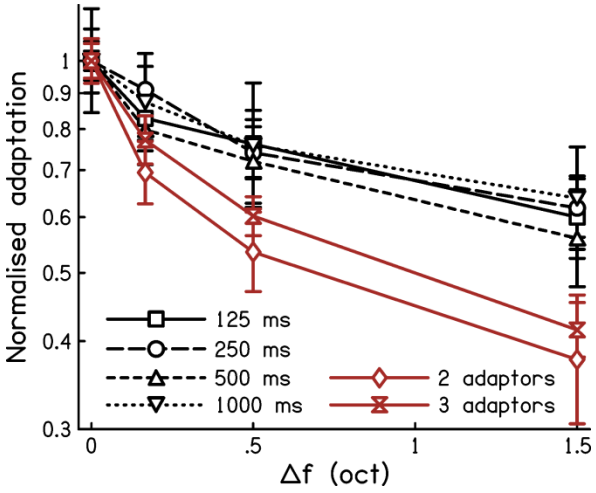

Figure 2.7. Adaptation plotted as a function of frequency separation, for each SOA and number of adaptors. Adaptation was normalised by its value when the adaptor and probe had the same frequency. Whilst adaptation appeared relatively broadly tuned when measured after a single adaptor (black), its tuning was considerably sharper after two or three adaptors (red).

2.D DISCUSSION

The MMR after multiple adaptors could not be fully explained by a passive release from adaptation. Instead, responses to deviant stimuli were much larger than predicted based on responses after a single adaptor. As this misprediction was specific to deviants, it provides evidence for a true mismatch response. Further, it points to a qualitative difference between the adaptation paradigm and the oddball paradigm: multiple adaptors elicit additional effects that cannot be predicted using single adaptors. However, contrary to the traditional view of the MMR, the true MMR reflected a modulation of the sensory-driven response to deviant stimuli, rather than a superimposed, endogenous negativity.

Functional role of the MMR

The MMR is traditionally thought to reflect the action of an early-warning mechanism (Escera & Corral, 2007). According to this view, an MMR is elicited when an unexpected stimulus is presented, to alert brain areas involved in orienting attention. In this model, the role of adaptation may be to fade out repetitive background sounds, decreasing the auditory system's sensitivity to repeated stimuli. Importantly, the current results suggest that stimulus repetition also leads to a sharpening of the neural representation of the adaptor. This could help the auditory system to discriminate the adaptor stimulus from any, non-repetitive stimuli, as it will increase the relative difference in their response sizes. Alternatively, rather than reflecting an early-warning mechanism, the true MMR could reflect neural mechanisms involved in an automatic form of perceptual learning known as "priming". Priming is manifest behaviourally as faster or more accurate processing for a previously-presented ("primed") stimulus than a novel stimulus (Tulving &

Schacter, 1990). The predominant neural correlate of priming is a smaller response to primed than novel stimuli (Ungerleider, 1995; Desimone, 1996; Wiggs & Martin, 1998), an effect that resembles the auditory MMR. Desimone (1996) proposed a sharpening theory of priming that is largely consistent with our current findings. In this theory, stimulus repetition leads to an increase in the stimulus specificity of neurons sensitive to the prime. This would reduce the number of neurons responding to the prime; neurons that initially responded well would continue to respond, whilst neurons that initially responded poorly would stop responding. The result would be a sparser, and potentially more efficient, neural representation of the prime (i.e. the "adaptor" in the terminology of the current study). Our findings suggest that the sharpening mechanism works alongside a suppressive mechanism, such as synaptic depression. This is because stimulus repetition was a pre-requisite for a change in stimulus specificity (the sharpening was observed only after two or three adaptors), even though large amounts of adaptation could be induced after only a single stimulus by reducing the SOA between the adaptor and probe. It is not yet known whether the two mechanisms are independent; it seems likely that suppression is a pre-requisite for sharpening, at least when stimuli are unattended (see Kauramäki et al., 2007, for a demonstration of sharpening due to attention).

The short-term changes in neural representations that we observed may be a pre-requisite for longer-term cortical plasticity. A recent study by Gutnisky et al. (2009) found that passive exposure to an oriented visual grating led to subsequent improvements in participants' ability to discriminate the exposed orientation from similar orientations. These improvements carried over to the following day, and accumulated over a period of more than a week. Long-term learning is often associated with an expansion of the responsive cortical area (Ungerleider, 1995).

However, Karni et al. (1995) found that this expansion was preceded by a reduction in the size of the responsive area (i.e. sharpening). Future research should investigate the interaction between the mechanisms of adaptation, sharpening and plasticity-related expansion, to see whether one is a pre-requisite for the others, or whether these phenomena occur independently.

Sustained activity during stimulus sequences may reflect attentional monitoring

To our surprise, we observed a sustained response that began shortly after the onset of a sequence and ended after the probe stimulus. It remained on at a relatively constant amplitude throughout the sequence, even during the silent intervals between stimuli. The generators of the SR were located anterior to the generators of the sensory-driven response. Previous MMR studies that presented stimuli in long, oddball sequences would have been unable to detect this SR, as they had to baseline correct the response to each stimulus separately. We suggest that the SR reflects a mechanism in frontal cortex involved in monitoring the incoming acoustic input and either alerting attention to unexpected changes or sharpening the response to recurring stimuli. It may be that changes of a sufficient magnitude to re-orient attention would lead to a modulation of the SR. This could explain the results of studies that have performed source analysis on the MMR for large stimulus changes. These studies tend to find an MMR subcomponent with a generator in the inferior frontal cortex (Alho, 1995; Rinne et al., 2000; Opitz et al., 2002). As they looked at neural activity to individual stimuli, rather than across a stimulus sequence, transient modulations in the sustained response would appear as transients in the event-related potentials. The inferior frontal cortex might be part of a multi-modal network involved in orienting attention to sudden changes in sensory input (Downar et al., 2000; Corbetta &

Shulman, 2002). Changes in visual, auditory or tactile input lead to activation in inferior frontal gyrus, as well as other areas including the temporoparietal junction (Downar et al., 2000). An alternative hypothesis is that the sustained response reflects feedback activity from frontal areas to the auditory cortex. Garrido et al. (2009) proposed that the MMR arises from the interaction of bottom-up sensory input and top-down stimulus predictions. In this scheme, the sustained response could mediate changes in stimulus specificity in auditory cortex.

A paradigm for accurately quantifying the true MMR

As the MMR occurs whether or not participants are attending to the stimuli, it has received great interest as a potential tool for studying sensory processing in groups of participants who would be difficult to test behaviourally (Näätänen & Escera, 2000; Näätänen, 2003). However, despite its apparent appeal, clinical studies of the MMR have yielded mixed results (Kane et al., 2000; Pekkonen, 2000; Bishop, 2007). This may be because current MMR paradigms confound the effects of passive release from adaptation with "true" MMR effects, which the current results suggest to be due to changes in the stimulus specificity of the auditory cortex.

In the classic oddball paradigm, the MMR is calculated by subtracting the response to a standard stimulus from the response to a deviant. As standard responses will be more adapted than deviant responses, part of the resulting difference waveform will reflect passive release from adaptation. One attempt to account for passive release from adaptation has been the MMR "control paradigm" (Schröger & Wolff, 1996; Jacobsen & Schröger, 2001). In this paradigm, a stimulus (A) is presented in two ways: as a deviant in an oddball block and as one of a number of equiprobable stimuli in a control block. Importantly, the other equiprobable stimuli

differ from stimulus A by *at least* as much as the standard differs from it in the oddball block. The MMR is calculated by subtracting the response to stimulus A in the equiprobable block from the response to stimulus A in the oddball block. The rationale of this subtraction is that, in both blocks, the response to stimulus A will reflect a passive release from adaptation, but only in the oddball block will stimulus A also reflect a true MMR. Surprisingly, when this paradigm is used for the MMR to a frequency change, the negative peak of the MMR appears to shift to a much later latency, placing it in the time window of the P2, as compared to its normal time window near the N1 (compare, for example, Fig. 1 of Jacobsen & Schröger, 2001, with Fig. 2 of Näätänen et al., 2004). The reason for this may be that the control paradigm overcompensates for passive release from adaptation; i.e. the response to stimulus A is less adapted in the equiprobable block than it is in the oddball block. In the N1 time window, this gives rise to a positivity (termed the "N1 effect" by Jacobsen & Schröger, 2001), as the less-adapted N1 in the equiprobable block is subtracted from the more-adapted N1 in the oddball block. For the same reason, as the P2 also adapts, the subtraction would give rise to a negativity in the time window of the P2 (the more-adapted P2 in the oddball block minus the less-adapted P2 in the equiprobable block). This negativity could be misinterpreted as a true MMR.

The advantage of the paradigm used in the current study is that it accurately quantifies the effect of passive release from adaptation, providing an unconfounded measurement of the "true" MMR. For use in the clinic, this paradigm would need to be shortened. The essential components would be measurements of (1) the response to a deviant probe after two identical adaptors, (2) the response to the probe after the first adaptor alone, (3) the response to the probe after the second adaptor alone.

Conclusions

We demonstrated two contributions to the MMR measured after multiple adaptors. One is passive release from adaptation, which can be measured with a single adaptor. The second appears to reflect a change in the frequency specificity of adaptation with stimulus repetition. This represents a true MMR component. The true MMR may contribute to perceptual learning in the auditory system by improving the discriminability of repeated stimuli. Finally, our paradigm provides a tool to accurately quantify the true MMR, which may help to improve the MMR's clinical utility.

Chapter 3. Short- and longer-term effects of stimulus context on sensory processing in human auditory cortex.

3.A INTRODUCTION

The brain is a dynamic system, in which neural responses are highly sensitive to stimulus context. When stimuli are repeated, responses typically decrease in a stimulus-specific manner, an effect referred to as "stimulus-specific adaptation" (Ulanovsky et al., 2004; Grill-Spector et al., 2006). To measure the time course of adaptation, previous studies have varied the time interval between a probe stimulus and a preceding adaptor stimulus. Using this paradigm, Wehr & Zador (2005) found that adaptation decays exponentially, with a time constant of about one second, in neurons of rat primary auditory cortex. The study reported in Chapter 2 obtained a similar time constant recording human auditory responses using EEG. However, the adaptation paradigm is unlike the situations encountered in natural environments. In natural environments, stimuli are stochastic, with some occurring frequently (such as the rustling of leaves in the wind) and others occurring infrequently (such as the snapping of a twig under a predator's foot). These are situations in which stimulus context plays a critical role, as it helps to establish which sounds have behavioural importance and which have not. It is currently unknown whether the contextual effects have the same time course in these more complex situations as they do in the adaptation paradigm.

A number of studies have used the oddball paradigm to examine contextual effects in a complex listening environment. In the oddball paradigm, infrequent "deviant" stimuli are presented amongst frequent "standards". Deviants tend to elicit a larger neural response than standards, a difference referred to as the "mismatch negativity", or "mismatch response" (MMR; see Näätänen et al., 2007, for a review).

The MMR is thought to arise partly due to adaptation and partly due to an additional contextual effect that is sensitive to regularities in the auditory environment (Winkler et al., 1996; Schröger, 1997). The regularity mechanism could serve an early-warning function, enhancing the response to stimuli that occur relatively infrequently (Escera & Corral, 2007). Alternatively, it could play a role in perceptual learning, enhancing the neural representations of frequent stimuli (Chapter 2).

It is generally assumed that the regularity mechanism is sensitive to the longer-term ("global") stimulus probabilities (Horváth & Winkler, 2004). Indeed, a number of studies have found that the amplitude of the MMR increases as the global probability of the deviant decreases (Näätänen et al., 1983). However, whilst the oddball paradigm may reveal additional contextual effects to the adaptation paradigm, it is more difficult to assess the time courses of these effects. This is because the traditional oddball paradigm presents stimuli in Bernoulli sequences, in which the local and global stimulus probabilities are identical. If a stimulus is presented 30% of the time in a block, it will also be presented with 30% probability on any given trial. Thus, a contextual effect that has a short memory span, and is thus sensitive to the local stimulus probabilities, would behave similarly to an effect that has a long memory span, and is thus sensitive to the global probabilities.

The memory span of a contextual effect can be defined as the time interval, or the number of stimuli, past which the stimulus history no longer influences responses. To examine the memory spans of the contextual effects in the oddball paradigm, the local and global stimulus probabilities must be dissociated. The current study achieved this dissociation using two methods. In the first method, the local and global probabilities were manipulated separately by presenting stimuli in a Markov sequence. In Markov sequences, the local probability of a stimulus depends on the

identity of the preceding stimulus, while the global probability of each stimulus within the sequence as a whole stays constant. In the second method, local and global probabilities were examined separately, by measuring response sizes in a Bernoulli sequence as a function of the local stimulus history, using the global stimulus probability as a parameter. With these methods, the time courses of the contextual effects can be compared with those obtained previously using the adaptation paradigm.

3.B EXPERIMENT 1

3.B.i Methods

Stimuli

As in Chapter 2, pure tones of 110-ms duration (including 10-ms cosine-squared onset and offset ramps) were presented diotically via headphones at 60 dB SPL. To equate the tones for loudness, masking noise, filtered such that equal energy (20 dB SPL) would fall into each peripheral auditory filter, was played continuously.

Procedure

Standard and deviant stimuli were presented with global probabilities of 70% and 30%, respectively, in three oddball conditions. The three conditions differed in the local stimulus probabilities. One condition was a Bernoulli sequence and the other two were Markov sequences, in which the local probabilities were set to favour repeated occurrence of the same stimulus over stimulus change (for an introduction to Markov sequences, see Norris, 1998). This led to a greater clustering of identical stimuli, and thus longer sequences without stimulus change (Fig. 3.1). For example, in the "strong clustering" condition, a standard stimulus was followed by another standard 90% of the time, whilst a deviant was followed by a standard only 23% of the time. Over a sufficient number of stimuli, standards still occurred 70% of the time and deviants 30% of the time.

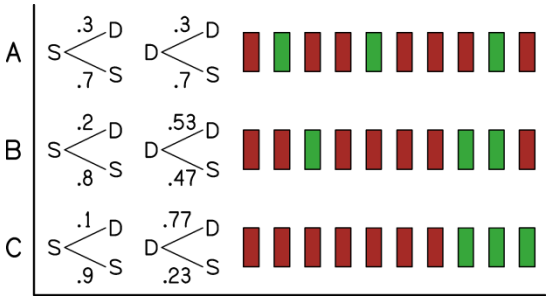

Figure 3.1. Transition probabilities, and example stimulus sequences (red: standard, green: deviant) for the “no clustering” (A), “weak clustering” (B) and “strong clustering” (C) conditions of Experiment 1.

Each of the Markov sequences can be described with a transition matrix, in which the vertical dimension represents the current stimulus (row one for a standard and row two for a deviant) and the horizontal dimension represents the subsequent stimulus. The transition matrix specifies the local stimulus probabilities. To find the longer-term stimulus probabilities, the transition matrix can be raised to a power using matrix multiplication. For example, the stimulus probabilities after ten stimuli in the strong-clustering condition would be given by:

$$\begin{bmatrix} 0.9 & 0.1 \\ 0.2333 & 0.7667 \end{bmatrix}^{10} = \begin{bmatrix} 0.7052 & 0.2948 \\ 0.6878 & 0.3122 \end{bmatrix}$$

If the current stimulus is a standard, the probability of the next stimulus being a standard is 90% (left matrix). However, the probability of a standard ten stimuli along is 71% (right matrix). As the Markov sequences used in the current study were ergodic, in that the probability of a stimulus change was always greater than zero, the stimulus probabilities converged on the global probabilities (70:30). The rate of convergence is typically assessed with the "total variation distance" (TVD), which, for the current sequences, is the absolute difference between the two values in a column (the result is the same, regardless of which column is used). For the strong-clustering condition, the TVD in the transition matrix is 0.67 (0.9 minus 0.23). The TVD after ten stimuli is 0.02. As the number of stimuli into the sequence is increased, the TVD will eventually approach zero. For the *n*-th stimulus, the TVD will be given by:

$$\Delta_n = k^n \cdot \Delta_0$$

where Δ_0 is the initial TVD, calculated from the transition matrix. The rate of convergence, *k*, is $1 - p - q$, where *p* is the transition probability of standard to deviant and *q* is the transition probability of deviant to standard. Thus, the greater the probability of a stimulus change, the smaller the value of *k* and the faster the

convergence. For the weak-clustering condition, *k* is 0.33, whilst for the strong-clustering condition, *k* is 0.67. Ninety-nine percent convergence occurs after 4.2 and 11.4 stimuli in the weak- and strong-clustering conditions, respectively. This places a lower bound on the memory span that a contextual effect would need to have to appear insensitive to clustering in the current conditions.

To measure the degree of clustering directly, one can find the mean number of standards or deviants that will occur consecutively in a sequence. This is simply the reciprocal of the probability of a stimulus change. In the strong-clustering case, the probability of a transition from a standard to a deviant is 0.1, therefore the mean number of standards that will occur consecutively is $\frac{1}{0.1}$ = 10. Similarly, the mean number of consecutive deviants is $\frac{1}{0.2333}$ = 4.3. A contextual effect would need to integrate over 14.3 stimuli to cover an average cycle of standards and deviants (note that, in this average cycle, standards occur 70% of the time and deviants 30%). These sequence lengths can be contrasted with those for the Bernoulli condition ("no clustering"), for which the mean number of consecutive standards and deviants is 3.3 and 1.4, respectively (cycle length of 4.7 stimuli). In the weak-clustering condition, it is 5 for standards and 2.1 for deviants (cycle length of 7.1 stimuli).

In all conditions, the acoustic frequencies of the standard and deviant were an octave apart (1 kHz ± 0.5 octaves) and the SOA was 500 ms. Each condition was presented twice, once with the lower frequency as the standard and once with the higher frequency as the standard. Participants completed Experiments 1 and 2, as well as the blocks within each session, in a random order. During recording, participants watched a subtitled DVD of their choice.

Electroencephalography and data pre-processing

The recording parameters, and the data pre-processing steps, were mostly identical to those used in Chapter 2 and in the majority of the experiments in Chapter 1. Neural responses from 32 electrodes were amplified, band-pass filtered and stored for offline analysis. Recordings were then lowpass filtered (35 Hz), down-sampled to 250 Hz and re-referenced to average reference. Data were segmented into epochs ranging from 100 ms before the start of each stimulus to 500 ms after the end of the stimulus. Epochs containing unusually large values were discarded and stereotyped artefacts were rejected using independent component analysis. Event-related potentials for each condition were imported into BESA and source waveforms were derived from a two-dipole model (one dipole at the centre of mass of primary region TE 1.0 in each hemisphere). The orientations of the two dipoles were fitted to a window encompassing the P1, N1 and P2 peaks of the grand-average ERP (0 to 200 ms after stimulus onset). Subsequently, source waveforms were averaged across hemispheres.

Data analysis

ERPs were triphasic, with peaks around 68, 112 and 172 ms, corresponding to the P1, N1 and P2 (Näätänen & Picton, 1987). The amplitude of the P1 was measured relative to a 50 ms pre-stimulus baseline, whilst the N1 was measured relative to the P1 (peak-to-peak) and the P2 was measured relative to the N1. This approach was chosen because, as in Chapters 1 and 2, both the N1 and the P2 were affected by the stimulus context. If a manipulation suppresses the N1 more than it suppresses the P2, the P2 could appear to increase, solely because it is less cancelled by the N1. This effect can be seen in the ERPs for Experiment 1 (Fig. 3.3). If the P2 were measured

relative to the pre-baseline period, the standard stimulus would appear to elicit a larger P2 than the deviant stimulus. However, this effect is artefactual, because the P2 rides on the much larger N1. In fact, the polarity reversal that follows the N1 is of much greater amplitude for deviants than standards, an effect captured by the peak-to-peak measure.

Subjects

Thirteen subjects (mean age of 21 years, range 19-23 years; 8 males) took part in both experiments. All subjects had pure-tone hearing thresholds at or below 20 dB HL at octave frequencies between 0.25 and 8 kHz and had no history of audiological or neurological disease. The study was approved by the Ethics Committee of the University of Nottingham School of Psychology.

3.B.ii Results

In this experiment, standard and deviant stimuli were presented with global probabilities of 70% and 30% in three conditions that differed in the local stimulus probabilities. Local probabilities in the "weak clustering" and "strong clustering" conditions were set to favour repeated occurrence of the same stimulus over stimulus change. If neural responses are sensitive to the local stimulus context, they should be affected by this manipulation. However, if they are only sensitive to the global stimulus context, responses should solely reflect whether the stimulus is a standard or a deviant.

Stimuli elicited triphasic neural responses, consisting of a P1, N1 and P2 peak (Fig. 3.2A). The scalp topographies of the P1 and P2 displayed a vertex positivity (Figs 3.2B-C), whilst the topography of the N1 displayed a vertex negativity (Fig. 3.2D). All three topographies inverted in polarity near the mastoids, consistent with generators in bilateral auditory cortices. Surprisingly, the P1 was quite large and appeared unaffected by either the global or the local probabilities (Fig. 3.3). This was confirmed with a linear mixed model, with fixed factors of global probability (70% standard or 30% deviant) and local probability ("none", "weak" or "strong" clustering) and a random effect of subject. Neither the main effects [global: $F(1,62) = 2.749$, $p = 0.102$; local: $F(2,62) = 0.677$, $p = 0.512$] nor the interaction [$F(2,60) = 0.085$, $p = 0.919$] were significant.

In contrast to the P1, the N1 was considerably larger for deviants than standards, displaying sensitivity to the global stimulus probabilities. It was not sensitive to the local probabilities, however, with the three clustering conditions eliciting peaks of similar size. A linear mixed model found a significant main effect of global probability [$F(1,62) = 166.022$, $p < 0.001$], but no effect of local probability

[$F(2,62) = 0.304$, $p = 0.739$] or an interaction of the two [$F(2,60) = 0.973$, $p = 0.384$]. On average, in the strong-clustering condition, 10 standards, and 4.3 deviants, would have occurred consecutively, so an average cycle of standards and deviants was 14.3 stimuli long (see Methods). Therefore, the N1 would have had to integrate over at least 14 stimuli to appear insensitive to the local probabilities.

Unlike the N1, the P2 did show sensitivity to the local stimulus probabilities; for deviants, the P2 decreased in amplitude as the degree of clustering increased. This meant that the deviant P2 was smaller when, on average, more deviants occurred consecutively. The average cycle of standards and deviants in the weak-clustering condition was 7.1 stimuli (5 consecutive standards and 2.1 consecutive deviants). It would appear, then, that the contextual effects on the P2 integrate over less than seven stimuli. A third linear mixed model found significant main effects of global probability [$F(1,60) = 41.929$, $p < 0.001$] and local probability [$F(2,60) = 3.310$, $p < 0.05$]. Pairwise comparisons found significant differences between the no-clustering and strong-clustering conditions ($p < 0.05$). However, there was also a significant interaction between the effects of global and local probability [$F(2,60) = 3.581$, $p < 0.05$]. This arose because the effect of local probability was significant for deviants [$F(2,60) = 6.676$, $p < 0.001$] but not for standards [$F(2,60) = 0.215$, $p = 0.807$]. For deviants, pairwise comparisons found a significant difference between the no-clustering and strong-clustering conditions ($p < 0.005$) and between the weak- and strong- clustering conditions ($p < 0.05$; no clustering versus weak clustering, $p = 0.110$).

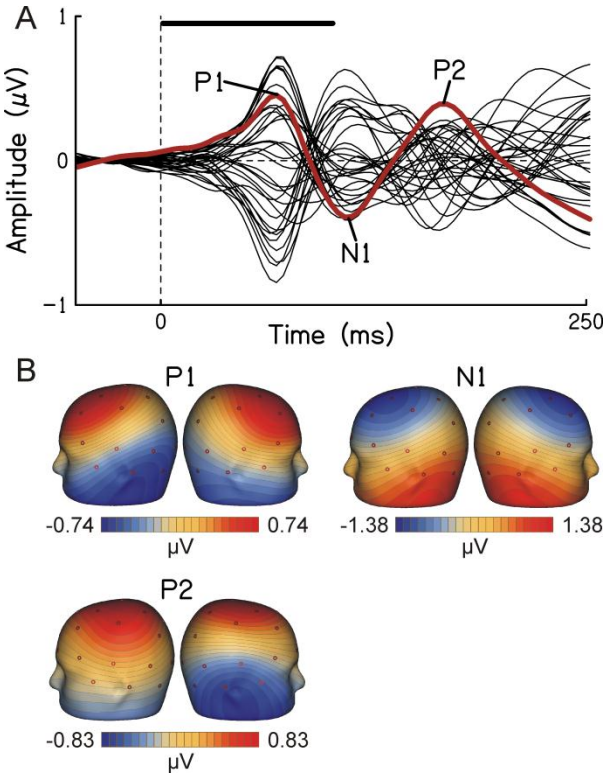

Figure 3.2. A: Mean event-related potential from Experiment 1, averaged across conditions. Each black line corresponds to a single recording electrode; the response from the vertex (Cz) electrode is plotted in red. B: Maps of field potential for the P1, N1 and P2 components, calculated over the 40-ms time window centred on each peak.

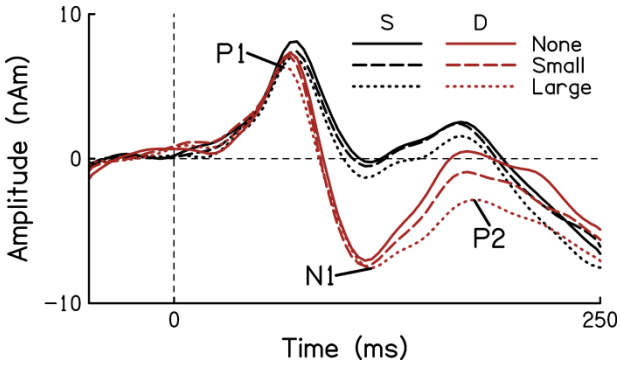

Figure 3.3. Source waveforms for Experiment 1, averaged across participants. Standard (black) and deviant (red) responses are plotted for each degree of clustering. Arrows mark the P1, N1 and P2 peaks of the deviant response in the large-clustering condition. The P1 appeared unaffected by the experimental manipulations. The N1 was influenced solely by the global stimulus probabilities, being larger for deviant stimuli than for standards. In contrast, the P2 to deviant stimuli was affected by the local stimulus probabilities, decreasing in amplitude as the degree of clustering increased.

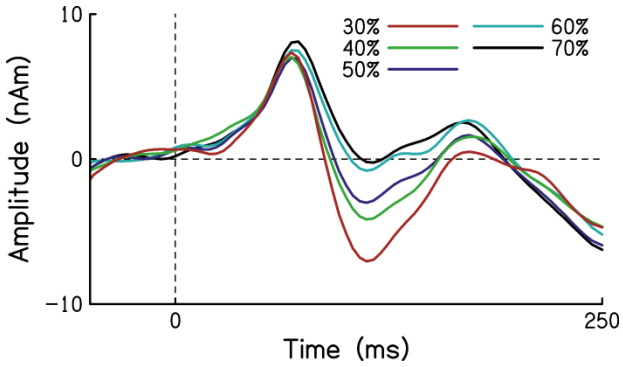

Figure 3.4. Source waveforms for Experiment 2, averaged across participants. The parameter is the stimulus probability. Whilst the P1 was unaffected by the experimental manipulations, the N1 and P2 increased in amplitude as the stimulus probability decreased.

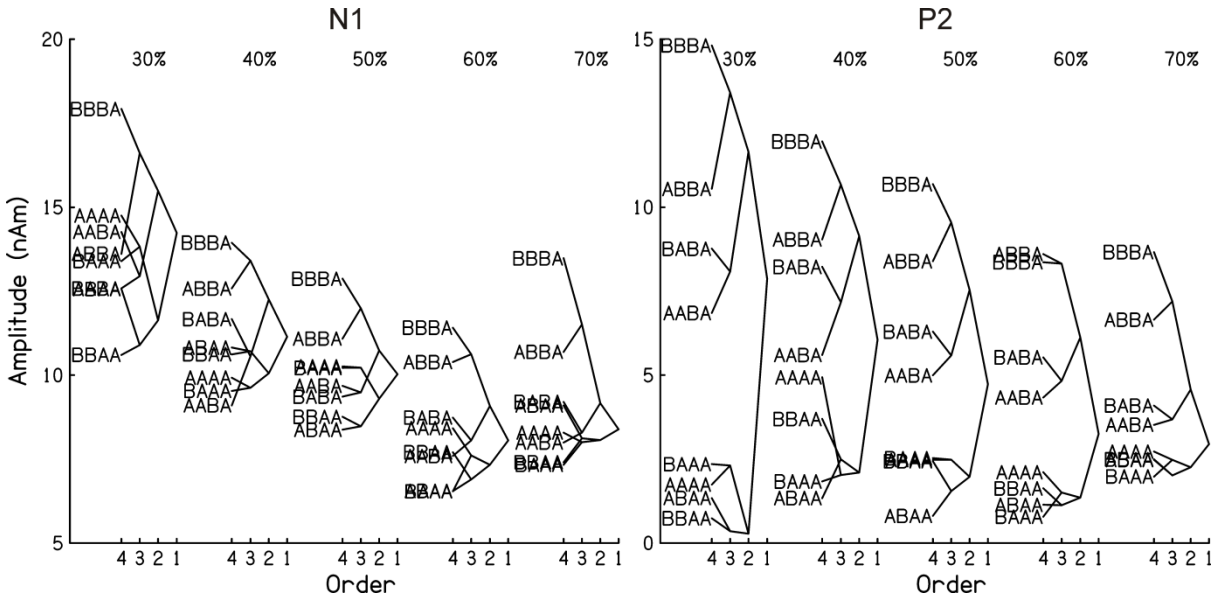

Figure 3.5. Stimulus history trees for the N1 (left) and the P2 (right) in Experiment 2. Trees are plotted separately for each global stimulus probability (30%, 40%, 50%, 60% and 70%). In each tree, the rightmost node reflects the mean response to a stimulus with the corresponding global probability. Nodes to the left reflect the response to the stimulus when it was preceded by a particular stimulus sequence (labelled to the left of each branch). In the sequence labels, the final "A" represents the current stimulus, preceding A's represent preceding stimuli with the same acoustic frequency, and preceding B's represent preceding stimuli with different acoustic frequencies. Thus, the node "BBBA" in the 30% tree is the response to a stimulus with a global probability of 30%, when that stimulus was preceded by three stimuli that differed from the current stimulus in acoustic frequency.

3.C EXPERIMENT 2

3.C.i Methods

Procedure

In the second experiment, stimuli were presented in Bernoulli sequences. There were two conditions presented in four blocks. In one condition, standard and deviant stimuli had probabilities of 60% and 40%, respectively. In a second condition, the two stimuli were equiprobable. As the first and second experiments tested the same participants, data from the Bernoulli condition of Experiment 1 were analysed alongside these data. In total, response sizes were obtained for stimuli with probabilities of 30%, 40%, 50%, 60% and 70%.

Electroencephalography and data pre-processing

Recording parameters were identical to those used in Experiment 1.

Data analysis

In Experiment 2, neural responses to stimuli with different global probabilities (30%, 40%, 50%, 60% and 70%) were measured as a function of the immediately-preceding stimulus history. From these measurements, history trees were plotted, with each node corresponding to a particular stimulus sequence (Squires et al., 1976; Ulanovsky et al., 2004). Permutation tests were used to compare nodes within and across the trees (Efron, 1993). For this, an ERP was calculated for each node and each subject. ERPs for the nodes of interest were randomly permuted within each subject, and then averaged across subjects to obtain "permutation ERPs", from which P1, N1 and P2 amplitudes were derived. This method allowed peak picking to be performed on ERPs with high signal-to-noise ratios, which improves the accuracy of

measurements because peak picking is a non-linear process. Subsequently, when comparing more than two nodes, the variance in amplitude across the permutation ERPs was calculated for each peak. When comparing only two nodes, the difference in amplitude was calculated. This process was carried out 1000 times to create a permutation distribution under the null hypothesis that there was no difference between nodes and thus the node labels were interchangeable. If the actual variance or difference between nodes of interest was greater than the extreme 5% of the permutation distribution, the null hypothesis was rejected.

In a subsequent analysis, an adaptation model with a single time constant was used to predict each stimulus history tree. The model assumed that the response to the final stimulus in each node's sequence (the "current" stimulus) was adapted by the immediately-preceding stimuli. The amount of adaptation was assumed to depend on the SOA between a preceding stimulus and the current stimulus, and on whether a preceding stimulus was the same as, or different from, the current stimulus. Adaptation is usually greatest when a preceding stimulus is identical to the current stimulus, and when the SOA between the stimuli is short (Chapter 2; Grill-Spector & Malach, 2001; Ulanovsky et al., 2004). As in Chapter 2, the decay of adaptation over time was modelled as an exponential process. This is in line with previous studies that attributed adaptation to the depletion of neurotransmitter pools at stimulated synapses (Wehr & Zador, 2005; Asari & Zador, 2009). Thus, the amount of adaptation induced by a preceding stimulus was modelled with the following function:

$$a_{\Delta f,t} = e^{mt + b_{\Delta f}}$$

where *t* is the time, in seconds, between the onset of the preceding stimulus and the onset of the current stimulus. The time constant of the decay of adaptation, *m*, was fitted to the data. The intercept of the function, *b*, was one when the preceding

stimulus was identical (A) to the current stimulus and between zero and one when the preceding stimulus was non-identical (B). The intercept for non-identical stimuli was fitted to the data. An intercept of one means that non-identical stimuli adapted the response to the current stimulus as much as identical stimuli adapted the current-stimulus response. An intercept of zero means that non-identical stimuli did not adapt the current-stimulus response at all. Therefore, the intercept specifies the frequency tuning of adaptation. It was assumed that the adapting effects of each preceding stimulus on the current-stimulus response simply accumulated (see Chapter 2). Thus, for a sequence that extends to order three, with an SOA of 500 ms, the response to the current stimulus, *p*, was calculated as:

$$p = s_1 \times \left(1 - a_{f\Delta,t=1.5}\right) \times \left(1 - a_{f\Delta,t=1.0}\right) \times \left(1 - a_{f\Delta,t=0.5}\right)$$

where s_1 is a scaling parameter that was also fitted to the data (three fitted parameters in total). All nodes entered the model for fitting, up to the fifth order. For each node, the associated stimulus sequence was extended to the seventh order, to encompass the full time course of adaptation. This involved weighting the intercept parameter, *b*, by the probability that a preceding stimulus was identical or non-identical to the current stimulus (see Results section). The three model parameters were fitted to each tree separately. Subsequently, for the P2, the 50% tree's parameters were used to predict responses for the other trees. *F*-tests compared the prediction errors obtained using the 50% tree's parameters with the errors obtained using a tree's own fitted parameters. An *F*-ratio was calculated for each tree, using the formula:

$$F = \frac{(SS_1 - SS_2)/(df_1 - df_2)}{SS_2/df_2}$$

where index 1 refers to the results obtained using the parameters from the 50% tree and index 2 refers to those obtained using a tree's own fitted parameters. *SS* stands for

the sum of the squared differences between the actual data points and the model data points. *df* stands for the model degrees of freedom (number of data points minus number of free parameters). The degrees of freedom associated with the *F*-ratio has a numerator of $(df_1 - df_2)$ and a denominator of df_2.

3.C.ii Results

Both the N1 and P2 peaks appeared to be affected by the global probability of the stimulus [Fig. 3.4; P1: $F(4,48) = 0.799$, $p = 0.532$; N1: $F(4,48) = 32.130$, $p < 0.001$; P2: $F(4,48) = 19.772$, $p < 0.001$]. Low-probability stimuli elicited larger responses than high-probability stimuli. For the N1, all pairwise comparisons were significant except those between the 50% and 60% conditions and between the 60% and 70% conditions. For the P2, the 50% and 70% conditions also did not differ significantly ($p > 0.05$). However, in these Bernoulli sequences, the global and local probabilities were identical for each stimulus.

To dissociate the local and global probabilities, response sizes were measured as a function of the immediately-preceding stimulus history, for each global probability separately. These measurements were used to construct stimulus history trees (Fig. 3.5). In each tree, the rightmost node is the average response to the stimulus (referred to as stimulus "A"). The neighbouring two nodes are the response to stimulus A when it was preceded by an identical stimulus ("AA") and the response when it was preceded by a different stimulus ("BA"). These two branches are referred to as "second order" branches. At the third order, there are four possible nodes, AAA, BAA, ABA and BBA. Stimulus history trees were constructed for the grand-average data and nodes were compared using within-subjects permutation tests (see Methods).

Local effects can be seen in each history tree. For example, the response to a stimulus preceded by an identical stimulus (AA) is smaller than the response to a stimulus preceded by a different stimulus (BA). This "one-back" effect was significant in each tree [N1: $p < 0.05$, P2: $p < 0.001$]. Local effects can be seen at higher orders also. For example, in most trees, the BBBA node is greater in amplitude than the ABBA node. More generally, for two nodes that differ only at the highest order, the node that starts with a B stimulus tends to have higher amplitude. This "n-back" effect was also found by Ulanovsky et al. (2004), in neurons of cat primary auditory cortex, and by Squires et al. (1976), for the P300 component of the human ERP. However, at least for the N1, local effects up to the fourth order could not fully account for neural response sizes. If the fourth order fully encompassed the memory span of the contextual effects, identical nodes would have the same response size regardless of global probability. This was clearly not the case. For the N1, the differences between the first-order nodes from each tree are still apparent at the fourth order (Fig. 3.5). For example, the BBBA node is highest for the 30% tree and smaller for higher global probabilities. The variance between the five BBBA nodes was significant ($p < 0.001$), as was the variance between the five AAAA nodes ($p < 0.005$). In contrast, for the P2, the variance between the five BBBA nodes was significant ($p < 0.005$), but the AAAA nodes did not differ significantly ($p = 0.952$). Whilst the upper P2 branches showed little sign of reaching an asymptote at the fourth order, the lower branches quickly stopped decreasing. This can be seen most clearly in the 30% tree. When the 30% stimulus was preceded by an identical stimulus (AA), the P2 was virtually abolished. As a result, higher-order nodes that ended in AA also elicited a very small P2, irrespective of the global stimulus probability. Further, whilst the upper P2 branches demonstrated an n-back effect, the lower branches did not.

To examine the time course of the contextual effects further, two summary measures were computed for each order of each tree. The nodes at each order were grouped according to the proportion of preceding A stimuli. For example, at the third order, node BBA was labelled "0%" as there were no A's in the stimulus history, whilst node AAA was labelled "100%". Nodes BAA and ABA were both labelled "50%" and their responses were averaged. Response sizes were regressed against the local proportion of A stimuli, yielding functions with negative slopes as responses generally decreased as the proportion of A stimuli increased. The regression lines were used to predict the response to a stimulus sequence that consisted of all A's (100%) and a sequence that consisted of all B's (0%; Fig. 3.6). The predictions from the trees representing different global probabilities should converge at an order that encompasses the memory span of the contextual effects. Therefore, the variance across the predictions should decrease to zero. For the N1, this did not happen even up to the fifth order. The variance between the predictions for different global stimulus probabilities was significant at each order for both the 0% and 100% predictions ($p < 0.005$). It was not possible to extend the predictions to higher orders. Beyond the fifth order, the number of response averages contributing to each node became too small, making amplitude measurements unreliable. Nevertheless, if the fifth order approached the memory span of the contextual effects on the N1, the variance across predictions should have shown signs of decreasing toward zero. As this decrease was not apparent, it can be assumed that the fifth order is not even close to the memory span of the contextual effects.

For the P2, the variance across predictions was significant at each order for the 0% predictions ($p < 0.005$) and, as for the N1, showed no sign of decreasing with increasing order. However, for the 100% predictions, the variance was not significant

after the first order (order 2: $p = 0.115$; order 3: $p = 0.933$; order 4: $p = 0.310$; order 5: $p = 0.993$). Therefore, there appear to be at least two contextual effects on the P2: one with a short memory span, which suppresses the response to repeated stimuli, and one with a longer memory span, which contributes to the larger response following a stimulus change. The suppressive mechanism could explain the effect of clustering on the P2 response to deviant stimuli in Experiment 1: Experiment 1 found that the deviant P2 was the smaller, the greater the degree of clustering. With greater clustering, the deviant stimulus would be more likely to be preceded by the same, rather than a different, stimulus. As the P2 is virtually abolished after only a single stimulus repetition, fewer deviants will elicit a sizeable P2, leading to a smaller average deviant response.

To see to what degree neural responses could be explained by local effects alone, an adaptation model with a single time constant was used to predict each stimulus history tree. For each node, the model assumed that the response to the current stimulus was adapted by the immediately-preceding stimuli. Adaptation was assumed to decay exponentially over time with a time constant that was fitted to the data. The amount of adaptation also depended on whether a preceding stimulus was the same as, or different from, the current stimulus (see Methods). The model was initially fitted to the nodes of the 50% tree, for the N1 and the P2 separately. As might be expected based on the results from the previous analysis, the model fit for the N1 was fairly poor, explaining only about a quarter of the variance ($r^2 = 0.275$), whilst for the P2, it was fairly good, explaining over 85% of the variance ($r^2 = 0.8761$; Fig. 3.7). The model yielded similar adaptation time constants (τ) for the N1 and the P2, with values of 1.1 s and 0.9 s, respectively. These values are close to the time constant of synaptic depression for neurons of primary auditory cortex (Wehr & Zador, 2005;

Asari & Zador, 2009). They are also close to the time constant of adaptation measured in Chapter 2 with the adaptation paradigm.

The frequency tuning parameter ($b_{\Delta f}$) differed between the N1 and the P2. This parameter specifies the amount by which non-identical stimuli adapt the response to the current stimulus, relative to the amount by which identical stimuli adapt the current-stimulus response. For the N1, $b_{\Delta f}$ was 0.9, whilst for the P2, $b_{\Delta f}$ was 0. This implies that the frequency tuning of adaptation was sharper for the P2 than for the N1. However, the poor model fit for the N1 would seem to make this result unreliable. Better fits were obtained when the model was fitted to the 40%, 60% or 70% N1 trees, with r^2 values of 0.58, 0.60 and 0.51, respectively. These yielded similar values for the adaptation time constant and frequency tuning parameter (40% tree: τ = 0.7, $b_{\Delta f}$ = 0.9; 60% tree: τ = 0.9, $b_{\Delta f}$ = 0.8; 70% tree: τ = 0.7, $b_{\Delta f}$ = 0.8). However, the fit for the 30% tree was poor (r^2 = 0.213) and yielded very different parameter values (τ = 3.8, $b_{\Delta f}$ = 0). The poor fits suggest that the N1 is sensitive to the relationships between the stimuli preceding the current stimulus. These relationships were not taken into account in the simple adaptation model, as the model assumed that the effects of multiple preceding stimuli on the current stimulus were independent. Some apparent anomalies in the 30% tree suggest that the N1 is sensitive to the global probabilities of short stimulus sequences. For example, under a purely local model, the AAAA node would be expected to elicit the smallest response. In terms of local probabilities, the most likely stimulus following a sequence of A's would be another A. Contrary to this expectation, the AAAA node elicited the second largest response in the 30% condition. The same was true for the AAA node. This finding would be expected if the N1 were sensitive to the global probabilities of short

stimulus sequences, as longer sequences of consecutive A's are unlikely when the global probability of A is low (30%).

For the P2, the adaptation model provided a good fit to each of the stimulus history trees, with r^2 values of 0.717, 0.831, 0.819 and 0.819 for the 30%, 40%, 60% and 70% trees, respectively. However, if a simple adaptation model can fully account for the P2 response size, it should be possible to use the parameters obtained for the 50% tree to predict trees of different global probabilities. This is because the adaptation model assumes that the response to a node only differs across trees when insufficient stimulus history is taken into account. For example, if adaptation is still present after four stimuli, the response to a BBBA sequence will be larger for a 30% tree than a 70% tree because, in the latter case, the sequence was more likely to have been preceded by an A stimulus. However, these differences can be taken into account by using the stimulus probabilities to extend each node's sequence. Thus, the BBBA sequence can be treated as an XBBBA sequence, where X is (0.3*A+0.7*B) for a 30% tree or (0.7*A+0.3*B) for a 70% tree (see Methods). In this way, sequences were extended to the seventh order, which corresponds to a reduction in the size of adaptation by 95%, assuming a decay time constant of one second. For each tree, the residual error was compared to that obtained when the tree's own fitted parameters were used. Fitting the model to a tree requires three free parameters, whilst all parameters are fixed when the values from the 50% tree are used. To take into account differences in model complexity, *F*-tests were used for model comparison. Essentially, the *F*-tests ask whether fitting a tree individually reduces the residual error significantly more than would be expected by chance from increasing the number of free parameters.

For all trees, except the 60% tree [$F(3,28) = 1.6332$, $p = 0.0966$], individual fits were significantly better than fits obtained using the parameters from the 50% tree [30% tree: $F(3,28) = 4.2302$, $p < 0.0005$; 40% tree: $F(3,28) = 3.9645$, $p < 0.0005$; 70% tree: $F(3,28) = 2.4318$, $p < 0.01$]. Interestingly, using the 50% tree parameters appeared to affect the upper branches of the trees more than the lower branches. Therefore, individual and 50% fits were also compared for the upper and lower branches separately. For the 30% and 40% trees, significant differences were found for nodes with a second-order B stimulus [30% tree: $F(3,13) = 23.2243$, $p < 0.0001$; 40% tree: $F(3,13) = 9.9972$, $p < 0.005$] but not for nodes with a second order A stimulus [30% tree: $F(3,13) = -0.7247$, $p = 1$; 40% tree: $F(3,13) = -0.6314$, $p = 1$]. A similar effect was found for the 70% tree, but it did not reach significance [lower branches: $F(3,13) = -1.0170$, $p = 1$; upper branches: $F(3,13) = 3.3221$, $p = 0.0536$]. Therefore, there appears to be a global effect on the upper, but not the lower, branches of the P2 trees, which cannot be accounted for using the simple adaptation model. When the global probability of the stimulus was low, the upper branches gave a larger response than predicted.

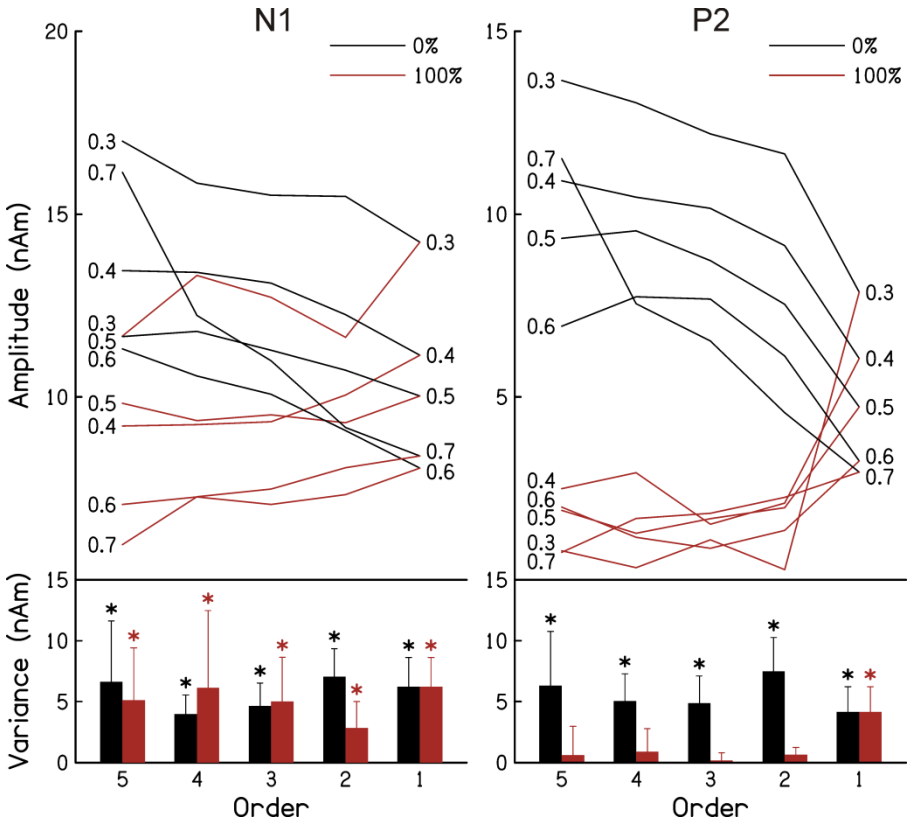

Figure 3.6. Top: Each pair of red and black lines shows the average response amplitude for a stimulus (A) preceded by only the same (100%; red lines) or only different (0%; black lines) stimuli; the parameter is the global probability of A in the stimulus sequence (labelled at the beginning and end of the lines). The response amplitudes were computed from the respective stimulus history trees by grouping nodes according to the local probability of A and regressing against these local probabilities (see text). Predictions from different trees should converge at an order that encompasses the entire memory span of the contextual effects. Bottom: The variance across predictions from different trees is plotted for each order. Error bars represent one standard error. Asterisks indicate predictions for which the variance was significant ($p < 0.05$).

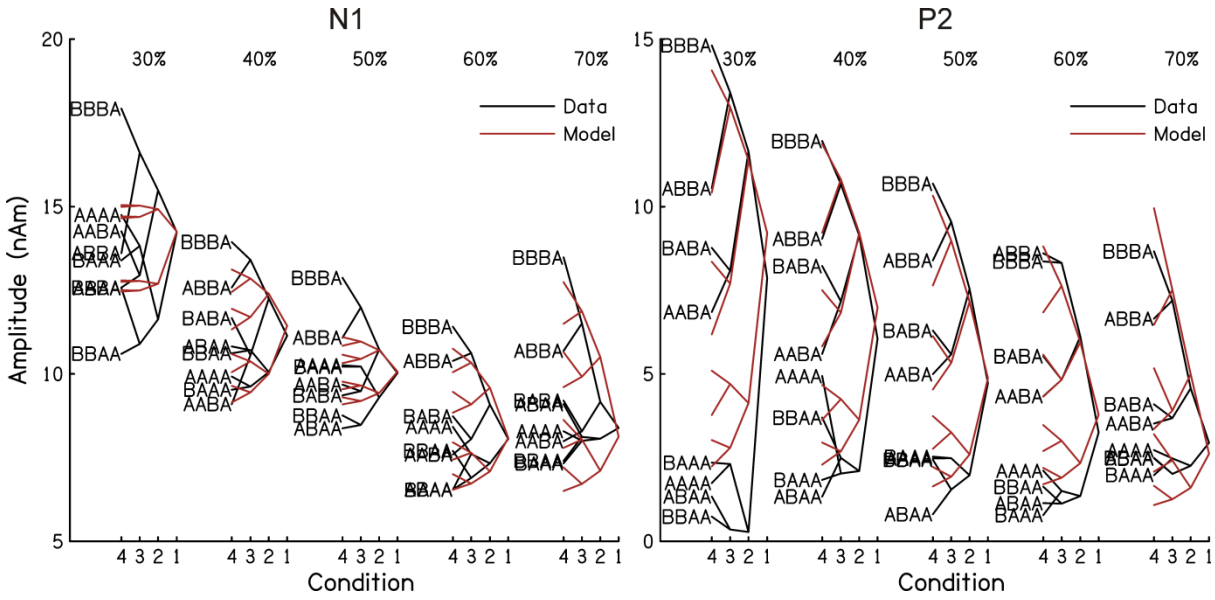

Figure 3.7. Stimulus history trees for Experiment 2 (black; see Fig. 3.5), overlaid with fits of a simple adaptation model (red). Whilst the model fits for the P2 trees were quite good (r^2 = 0.717-0.876), fits for the N1 trees were relatively poor (r^2 = 0.213-0.600), suggesting that the contextual effects on the N1 take into account the relationships between stimuli within a sequence (these relationships were not taken into account by the simple adaptation model).

3.D DISCUSSION

The effects of stimulus context differed markedly between the components of the auditory event-related potential. In the first experiment, the earliest component (the P1) appeared altogether insensitive to stimulus context, whilst the next component (the N1) was strongly affected by the global probabilities of stimuli. In contrast, the last component (the P2) was strongly affected by the local stimulus probabilities. The second experiment analysed the effects on the N1 and the P2 further. Both components were sensitive to the local stimulus context. The N1 was also influenced by longer-term effects that partly reflected the global probabilities in the stimulus sequences. The P2 was influenced by longer-term effects, but only for stimuli that were directly preceded by a non-identical stimulus. A local effect with a very short time constant suppressed the P2 for stimuli preceded by an identical stimulus. A previous study demonstrated that the MMR reflects a modulation of the N1 and P2 amplitudes (Chapter 2). The current study suggests that this modulation arises from a mixture of contextual effects with short and long memory spans.

P1

One of the most striking findings of the current study was the apparent lack of contextual effects on the P1. Unlike the N1 and the P2, the P1 was robust to variation in either the local or global stimulus probabilities. This is surprising, as previous studies using the adaptation paradigm have found a smaller P1 to the second stimulus than the first, even when the intra-pair SOA was identical to the SOA used in the current study (500 ms; Cardenas et al., 1997; Jerger et al., 1992; Waldo et al., 1992). One interpretation is that there is a contextual effect on the P1, but that it is stimulus unspecific. This means that the P1 would be suppressed to the same extent whether

the current stimulus was preceded by an identical or a non-identical stimulus. If there are stimulus-specific effects, they may have a relatively short time course, and thus could emerge when a shorter SOA is used. Using MEG, Huotilainen et al. (1998) localised the source of the P1 to a region of the superior temporal gyrus slightly anterior to the source of the N1 (see also Mäkelä et al., 1994). Buchwald et al. (1991; 1992), based on work with an animal model of the P1, suggested that the P1 also receives contributions from *sub-cortical* regions, such as the thalamus. The absence of prolonged, stimulus-specific effects suggests that the neural activity underlying the P1 could provide a veridical representation of the acoustic input for processing by later cortical areas. Stimulus-unspecific effects could act as a gain control mechanism to ensure that subsequent processing is not overwhelmed by rapid, high-amplitude input. This would be consistent with a hypothesised role for the P1 in sensory gating: participants with Schizophrenia, a disorder characterised by hypersensitivity to sensory input, show little P1 suppression in the adaptation paradigm (Waldo et al., 1992).

N1

Unlike the P1, the N1 was strongly influenced by stimulus-specific contextual effects. In Chapter 2, the time course of adaptation, measured in the adaptation paradigm, could be described by a decaying exponential with a time constant of around one second. A similar time constant has been found for synaptic depression in neurons of primary auditory cortex (Wehr & Zador, 2005; Asari & Zador, 2009). It is likely that a similar effect influences the N1 in the oddball paradigm. Experiment 2 demonstrated that the N1 was sensitive to the local stimulus context, and fits of a simple adaptation model estimated a time constant of around one second. However,

this effect cannot fully account for the current results. In the first experiment, the N1 was strongly influenced by the global stimulus probabilities and appeared insensitive to the local probabilities. As an average cycle of standards and deviants in the strong-clustering condition contained about 14 stimuli, and the SOA between stimuli was 500 ms, this result implies that the N1 has a memory span of at least 7.5 s. Adaptation, with a 1 s time constant, would have decayed to less than 1 % of its original size after only 4.6 s. A long memory span is also suggested by the results of the second experiment. In Experiment 2, the N1 amplitude differed as a function of the global stimulus probability, even when comparing stimuli that had the same local stimulus history. A summary measure found that the influence of the global stimulus probability had not even started to diminish after four stimuli, even though adaptation, with a 1 s time constant, would have decayed to less than 15% of its initial value after four stimuli.

A long memory span is consistent with previous EEG studies that have presented blocks of identical stimuli. These studies have found that neural responses increase with increasing SOA up to about ten seconds (for a review, see Näätänen & Picton, 1987). Note that the paradigm used in these studies is equivalent to an oddball sequence with extreme stimulus probabilities (100% and 0% for the standard and deviant, respectively). Interestingly, Lütkenhöner & Steinsträter (1998) have localised the source of the N1 to planum temporale (posterior to primary auditory cortex), an area thought to be involved in perceptual streaming (Griffiths & Warren, 2002). A streaming mechanism would benefit from a relatively long memory span, as a long memory span would allow the mechanism to accumulate evidence about the objects present in the auditory environment. Psychophysical studies have found that

streaming builds up over a period of about ten seconds (Carlyon et al., 2001; Micheyl et al., 2005), consistent with the memory span of the N1.

The current results also suggest that the contextual effects on the N1 are relatively complex. The adaptation model, which assumed that the effects of multiple preceding stimuli on the current stimulus simply summed, provided a relatively poor fit to the N1 amplitudes. This implies that the N1 was influenced by the relationships between the stimuli in the preceding sequence. The poor fit appeared to reflect sensitivity to the global probabilities of particular short stimulus sequences: sequences that occurred infrequently in a block (such as "AAAA" when the global probability of A was low) elicited a large N1. This finding is consistent with previous studies demonstrating that the negative component of the MMR is sensitive to changes in a repeated sequence (Sussman et al., 1999; Horváth et al., 2001). For example, an occasional stimulus repetition ("AA") in a block of alternating stimuli ("ABAB…") elicits an enhanced negativity. Note that the contextual effects on the N1 would require a long memory span to detect the probabilities of short stimulus sequences. A mechanism that is sensitive to the probability of a stimulus sequence could facilitate auditory learning of sound patterns, as certain combinations of vowels and consonants typically occur together in speech. Alternatively, in terms of early-warning function, such a mechanism would be able to distinguish the pattern of sounds generated by the footsteps of a friend from the pattern generated by the footsteps of a foe.

P2

The current study also found longer-term effects on the P2, but this was only apparent for stimuli that were preceded by non-identical stimuli ("BA"). Responses to stimuli preceded by an identical stimulus ("AA") were greatly suppressed. There are

two striking aspects of this suppression. The first is its size. The P2 to AA stimuli was virtually zero, even for stimuli with a global probability of 30%. The second is its time course. The effect appears to have a memory span of just one stimulus (or half a second). The contextual effects on the N1 displayed neither of these properties. It may be that the suppressive effect is performing something akin to a first-order differentiation of the stimulus sequence. In this way, it would act as an "edge detector", responding only to stimulus change. In vision, edges in space help to define objects. Kubovy & van Valkenburg (2001) proposed that edges in time and spectral frequency help to define "auditory objects" – discrete units of sound such as words or musical notes. Griffiths & Warren (2004) note, however, that a single auditory object can have components separated in time and frequency. A mechanism that performs a differentiation over the stimulus sequence, rather than over time, could partially alleviate this problem. This hypothesis could be strengthened by varying the SOA within a sequence, to examine whether the suppressive effect has a memory span of one stimulus, or whether it has a memory span of half a second.

The nature of the longer-term effects on BA stimuli is unclear. Whilst the simple adaptation model provided a relatively poor fit to the N1 responses, the model fit the P2 responses quite well. This suggests that the P2 is not particularly sensitive to the relationships between stimuli in the local stimulus history. The simple adaptation model could not fully account for P2 responses however, as evidenced by the dependence of the model parameters on the global stimulus probability. Consistent with the presence of a global effect, a summary measure found that the influence of the global probability on BA stimuli had not even started to diminish after four stimuli (a result described earlier for the N1 to AA and BA stimuli). It may be that the global effect reflects a second, longer time constant of adaptation, as suggested by

Ulanovsky et al. (2004) for neurons of primary auditory cortex. Alternatively, the global effect could reflect an active facilitation of deviant responses or a reduction in the suppressive effects of adaptation on deviant responses (see Chapter 2). Finally, a number of studies have found long-term effects on the P2 amplitude that extended over periods of hours and days (Atienza et al., 2002; 2004; see also Crowley & Colrain, 2004). This suggests that the P2 is influenced by additional contextual effects that cannot be studied with the current paradigm.

Conclusions

Despite forty years of intensive study, the neural basis of the MMR remains poorly understood. Using the oddball paradigm, significant MMRs have been found for a diverse array of stimulus features and stimulation parameters (for a review, see Näätänen et al., 2007). The current study found that a number of contextual effects operate in the oddball paradigm, each of which could serve a different functional role. Therefore, part of the difficulty in interpreting the MMR might arise because different studies are more, or less, sensitive to different contextual effects. Further work, to disentangle the different contributions to the MMR, could help to improve its scientific, and clinical (e.g. Kane et al., 2000; Pekkonen, 2000; Bishop, 2007), utility.

Chapter 4. Transient and sustained auditory responses are differentially affected by the relative stimulus probabilities.

4.A INTRODUCTION

The behavioural significance of a stimulus depends on its context. A stimulus that occurs infrequently, such as the snap of a branch under a predator's foot, could signal a threat in the environment (Escera & Corral, 2007), whilst a stimulus that occurs frequently, such as a repeated syllable in speech, could be a cue for perceptual learning (Gutnisky et al., 2009; Agus et al., 2010). Many auditory studies have used the oddball paradigm to examine the effects of stimulus context on neural responses (see Chapter 3). In this paradigm, infrequent "deviant" stimuli are presented amongst frequent "standards". Even when participants are not attending to the stimuli, deviants typically elicit larger neural responses than standards. This difference, referred to as the "mismatch negativity" or "mismatch response" (MMR), increases as the probability of the deviant decreases (see Näätänen et al., 2007, for a review). The MMR is thought to reflect a mechanism that is sensitive to regularities in the auditory environment (Winkler et al., 1996; Schröger, 1997). When determining these regularities, the mechanism is thought to take into account the relationships amongst different types of stimuli within a sequence (Sussman et al., 2003). This means that the MMR should be sensitive to the relative, rather than the absolute, stimulus probabilities. This hypothesis was tested in the current study. In order to calculate the relative stimulus probabilities, the MMR would need to be sensitive to the longer-term ("global") stimulus context. This sensitivity was demonstrated in Chapter 3, alongside sensitivity to the short-term ("local") context.

Jääskeläinen et al. (2004) argued that the MMR arises solely from neural adaptation. Adaptation is ubiquitous in the cortex. It is closely related to synaptic

depression (Wehr & Zador, 2005; Asari & Zador, 2009) and is manifest as a reduction in the neural response to a stimulus when it is preceded by a similar stimulus. Adaptation is greatest when stimuli activate similar neural groups. Importantly, adaptation would not be expected to take into account the relationships between stimuli. Thus, if the MMR solely reflects adaptation, it would be primarily sensitive to the absolute stimulus probabilities. The study reported in Chapter 2 found that MMR does not arise due to adaptation alone, but instead reflects a change in the stimulus specificity of cortical neurons due to repetition. When a single adaptor stimulus preceded a probe stimulus, adaptation affected a wide range of acoustic frequencies. However, when the adaptor was repeated, so that it effectively became a standard, adaptation became more sharply tuned. It is unclear whether this "sharpening" effect reflected the relative or the absolute stimulus probabilities, as, after multiple adaptors, the probe not only occurred less often relative to the adaptors but also less often in the sequence.

In the traditional oddball paradigm, there is only one standard stimulus and one deviant stimulus. Therefore, changes in the *relative* probability of standards and deviants also lead to changes in the *absolute* stimulus probabilities. In Chapter 3, a modified oddball paradigm was used to examine whether the MMR is sensitive to the local or global stimulus context, yet, it too involved only a single standard and a single deviant. In many everyday listening environments, such as those encountered in a busy restaurant or marketplace, there are more than two stimuli present. Some of these stimuli will occur more often than others. Sensitivity to the relative stimulus probabilities could help the auditory system to focus on the most important stimulus in the environment, irrespective of how often it occurs in absolute terms.

To examine whether neural responses are sensitive to the relative or absolute stimulus probabilities, the current study compared neural responses in a traditional oddball paradigm with responses in a modified oddball paradigm, which contained more than two stimuli. The stimuli differed in acoustic frequency, which enabled changes in the stimulus specificity of the contextual effects to be examined. The current study used stimuli that had longer duration than those used in Chapters 2 and 3, in order to measure both their transient and sustained neural responses. The effects of stimulus context on the sustained response (SR) are particularly poorly understood. Early work by Picton et al. (1978a; 1978b) found only weak contextual effects on the SR, whilst a recent study, which used the oddball paradigm, found much stronger effects (Kretzschmar & Gutschalk, 2010). The current study assessed whether transient and sustained responses differ in their sensitivity to the relative stimulus probabilities.

4.B METHODS

The current study compared neural responses in a modified oddball paradigm to responses in the traditional oddball paradigm. The oddball paradigm was modified by introducing variability into either the standard or the deviant stimulus. Specifically, the spectral frequency of the modified stimulus was selected, with equal probability, from a set of four different values centred on the original stimulus frequency. This manipulation left the absolute probability of the unmodified stimulus unchanged. However, as it introduced a number of low-probability spectral frequencies into the sequence, the manipulation increased the relative probability of the unmodified stimulus.

Stimuli

As in Chapters 2 and 3, pure tones were presented diotically via headphones (60 dB SPL) alongside a continuous masking noise (20 dB SPL per ERB_N). However, pure tones had a longer duration in the current study (510 ms, including 10 ms cosine-squared onset and offset ramps), in order to elicit a measurable sustained response from the cortex.

Procedure

There were two "control" conditions and two "test" conditions. In each condition, the (nominal) acoustic frequencies of the standard and deviant were an octave apart (1 kHz ± 0.5 octaves). Each condition was presented in two 20-minute blocks, once with the lower, and once with the higher, frequency as the standard. The stimulus onset asynchrony (SOA) was always 1 s.

In the control conditions, standards and deviants were presented with absolute probabilities of 80% and 20% (“80:20” condition) and 50% and 50% (“50:50” condition). The “50:4D” test condition was derived from the 50:50 condition by replacing the deviant in the 50:50 condition with four equiprobable deviants, which differed slightly in acoustic frequency (±1/8 and ±3/8 octave around the original frequency; Fig. 4.1). The standard was not changed, which meant that its absolute probability also remained unchanged. However, as each deviant stimulus was now presented only 12.5% of the time, the standard became as likely as the standard in the 80:20 condition in relative terms, in that it was now four times as likely as the other stimuli in the sequence.

The “4S:20” test condition was derived from the 80:20 condition. In this case, the standard in the 80:20 condition was replaced with four standards, each presented 20% of the time. As a result, the absolute probability of the deviant remained unchanged, but its relative probability was increased to become equiprobable with the other stimuli in the sequence, as in the 50:50 condition.

Blocks were presented in a random order to each participant. During recording, participants watched a subtitled DVD of their choice.

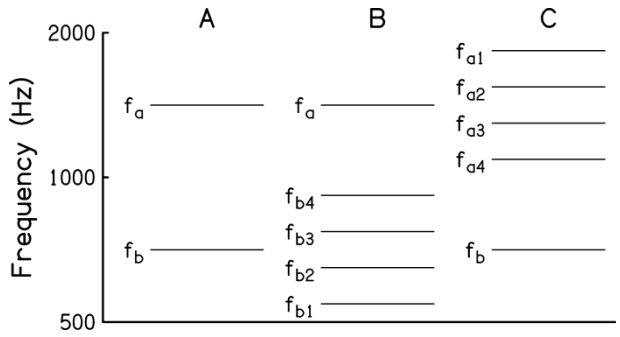

Figure 4.1. Schematic of the acoustic frequencies used in the current study. The control conditions (50:50 and 80:20) used the frequencies in column A (f_a and f_b) as standards and deviants. In the 50:4D and 4S:20 conditions, either the standard or deviant (each of which could be either f_a or f_b) was replaced by four frequencies, equally spaced around the original frequency on a logarithmic scale (B and C).

Electroencephalography and data pre-processing

The recording parameters, and the data pre-processing steps, were mostly identical to those used in Chapters 3 and 2, and in the majority of the experiments in Chapter 1. There were only two differences: (1) The data were segmented into epochs ranging from 100 ms before the start of each stimulus to 1000 ms after the end of the stimulus; (2) in the BESA source model, the orientations of the two dipoles were fitted to a window encompassing the entire stimulus duration (0 to 500 ms after stimulus onset).

Data analysis

The responses to both the standards and deviants consisted of a triphasic transient followed by a sustained response (SR), which decayed to baseline shortly after stimulus offset (fig. 4.2). The SR and the most prominent transient component, the N1, were analysed further. The amplitude of the SR was taken as the mean response size in the 100-ms window preceding stimulus offset (i.e. 400-500 ms after stimulus onset). This ensured that the response was not contaminated by the earlier transient. The amplitude of the N1 was measured relative to the preceding positive deflection (the "P1"). As in the previous chapters, this approach was chosen because the neural processes underlying the components of the transient response partially overlap in time (Näätänen and Picton, 1987; Makeig et al., 1997). Therefore, as the components have different polarities, they partially cancel each other out. This is a problem if a manipulation affects one component more than another component. The peak-to-peak measure minimises the impact of this problem.

To examine the stimulus specificity of the local (short-term) effects on the transient N1 component, response sizes were measured as a function of the

immediately-preceding stimulus. Of particular interest were the responses to deviants in the 50:4D condition and the responses to standards in the 4S:20 condition. In the 50:4D condition, a deviant could be preceded by the exact-same deviant frequency ("DD"), a different deviant frequency ("dD") or the standard ("SD"). If the local effects are broadly tuned in frequency, treating all deviants similarly, the dD response should be similar in size to the DD response. However, if the local effects are sharply tuned, the dD response should be similar to the SD response. A stimulus-specificity index (SSI) was obtained with the following formula:

$$\frac{dD - DD}{SD - DD}$$

An SSI of zero would mean that the dD response was identical to the DD response (broad tuning). An SSI of unity would mean that the dD response was identical to the SD response (sharp tuning). Similarly, in the 4S:20 condition, a standard could be preceded by the same standard frequency ("SS"), a different standard frequency ("sS") or the deviant ("DS"). An SSI was obtained with the following formula:

$$\frac{sS - SS}{DS - SS}$$

SSIs were calculated using N1 amplitudes from the grand-average response waveforms (averaged across participants). A bootstrap distribution was obtained for each SSI by computing 1000 between-subjects resamples of the response waveforms. From each distribution, a 95%-*t*-confidence interval was derived. The SSI was compared between deviants (50:4D condition) and standards (4S:20 condition) using a between-subjects permutation test. Specifically, the difference between the two SSIs was compared to a permutation distribution, which was calculated in the following manner: (1) Each of the 24 participants was entered into a subject pool; twelve were

randomly assigned to group A and the other twelve to group B. (2) In each group, average ERPs were obtained for the three nodes of interest and an SSI value was calculated. (3) The difference between the SSI value from group A and the SSI value from group B was entered into the permutation distribution, and the process was repeated 1000 times. If the permutation distribution exceeded the SSI difference between the 50:4D and 4S:20 conditions in fewer than 5% of cases, the null hypothesis (that the two SSI values were identical) was rejected. The principal advantage of this method is that it allows peak picking to be performed on ERPs with high signal-to-noise ratios. This improves the accuracy of measurements because peak picking is a non-linear process.

For both the N1 and the SR, within-subjects permutation tests were used to compare responses to pairs of nodes within the local stimulus history trees. ERPs for the nodes of interest were randomly permuted within each participant, and then averaged across participants to obtain grand-average waveforms. The difference in response amplitude between the two waveforms was entered into a permutation distribution and the process was repeated 1000 times. The actual difference between the nodes of interest was compared to this distribution.

Subjects

Twelve subjects (mean age: 26 years; 4 males) completed the 50:50, 80:20 and 50:4D conditions and a further twelve subjects (mean age: 22 years; 7 males) completed the 50:50, 80:20 and 4S:20 conditions. All subjects had pure-tone hearing thresholds at or below 20 dB HL at octave frequencies between 0.25 and 8 kHz and had no history of audiological or neurological disease. This study was approved by the Ethics Committee of the University of Nottingham School of Psychology.

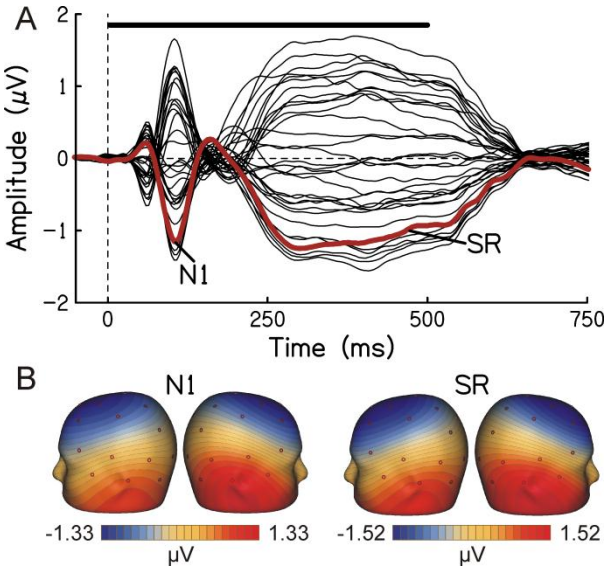

Figure 4.2. A: Grand-average event-related potential. Each black line corresponds to a recording electrode; the response of the vertex (Cz) electrode is plotted in red. Stimuli elicited a prominent N1 peak, followed by a sustained response (SR) that decayed to baseline shortly after stimulus offset. B: Maps of field potential for the N1 (computed over the 40-ms window centred on the grand-average N1 peak) and the SR (computed between 400-500 ms after stimulus onset).

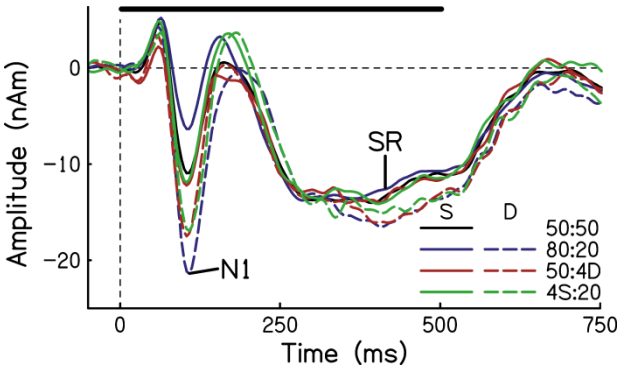

Figure 4.3. Source waveforms for each stimulus type, averaged across participants. The N1 was largest for the deviant in the 80:20 condition, and smallest for the 80:20 standard. The latter portion of the subsequent SR (400-500 ms after stimulus onset) was larger for deviants than standards, regardless of the stimulus condition.

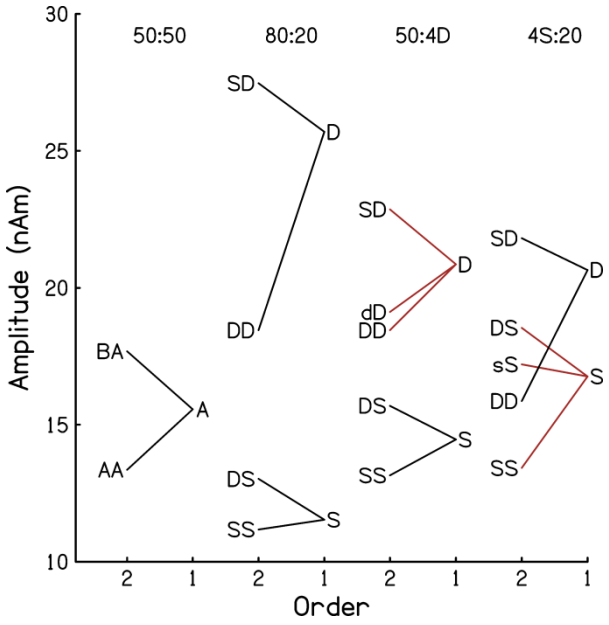

Figure 4.4. Stimulus history trees for each condition. Order 1 is the mean N1 amplitude to a standard ("S") or a deviant ("D"). Order 2 is the N1 amplitude as a function of the immediately-preceding stimulus. For example, "SD" is the response to a deviant when it was preceded by a standard. In the 50:4D condition, "DD" is the response to a deviant when it was preceded by the same deviant frequency, whilst "dD" is the response to a deviant when it was preceded by a different of the four deviant frequencies. Likewise, in the 4S:20 condition, "sS" is the response to a standard preceded by a different standard frequency.

4.C RESULTS

Stimuli elicited a triphasic transient response followed by a sustained response (SR) that decayed to baseline shortly after stimulus offset (fig. 4.2A). The most prominent transient component, the "N1", occurred around 104 ms after stimulus onset. The scalp topographies of the N1 and SR were similar (fig. 4.2B), with negativity at the vertex and polarity inversion around the mastoids. This is consistent with generators in auditory cortex bilaterally. For the N1 and SR separately, linear mixed models were used to compare response sizes across the seven stimulus types. The main effect of stimulus type was significant for both components [N1: $F(6,91.041) = 36.563$, $p < 0.001$; SR: $F(6,91.293) = 8.038$, $p < 0.001$].

N1

As expected from previous studies (Chapter 3), an MMR was produced in the 80:20 condition, in that the N1 was significantly larger for the deviant than the standard (Fig. 4.3; $p < 0.001$). This arose from both a *suppression* of the standard N1 ($p < 0.001$) and an *enhancement* of the deviant N1 ($p < 0.001$), relative to the 50:50 baseline (fig. 4.3). If the suppression reflects sensitivity to the relative probability of the standard, it should also be apparent in the 50:4D condition. However, the 50:4D standard was not significantly different from the 50:50 response ($p = 0.165$), suggesting that the suppressive effect in the standard response arose primarily from sensitivity to the absolute stimulus probabilities. In contrast, the mean deviant response in the 50:4D condition (averaged across all four deviant frequencies) was significantly larger than the 50:50 baseline ($p < 0.001$), suggesting that the enhancement effect in the deviant response was sensitive to the relative stimulus probabilities. However, this result could also reflect a local (short-term) effect,

because a deviant might have been less suppressed when preceded by a different deviant ("dD") than when preceded by the same deviant ("DD"). To examine whether this was the case, responses were analysed as a function of the immediately-preceding stimulus (fig. 4.4). Local effects of the preceding stimulus were observed in all conditions, in that the response to a stimulus was greater when it was preceded by a different stimulus than when it was preceded by the same stimulus. However, the dD response in the 50:4D condition was only slightly larger than the DD response (permutation test, $p = 0.160$). This implies that the local effects on deviants were broadly tuned, treating the different deviant frequencies similarly. The slight difference between dD and DD stimuli (0.7 nAm) cannot account for the large difference (6.4 nAm) between the mean responses to standards and deviants. Therefore, longer-term effects must have played a role. The tuning of the local effects was quantified with the stimulus-specificity index (SSI), which is defined as the difference between the dD and DD responses, as a proportion of the difference between the DD and SD responses (see Methods). An SSI of zero would imply broad frequency tuning, whilst an SSI of one would imply sharp tuning. In the current case, the SSI was 0.2, which was significantly different from unity but not from zero (95%-*t*-confidence interval: -0.4 to 0.7). This implies broad frequency tuning of the effect of the preceding stimulus in the 50:4D condition.

If the enhancement of the deviant N1 in the 80:20 condition arose solely from sensitivity to the relative stimulus probabilities, one might expect no enhancement in the 4S:20 condition. The 4S:20 deviant had the same absolute probability as the 80:20 deviant, but it had the same relative probability as the 50:50 deviant, as it was equiprobable with the other stimuli in the sequence. Although the 4S:20 deviant was smaller than the 80:20 deviant ($p < 0.005$), it was still larger than the 50:50 response

($p < 0.001$; fig. 4.3). This suggests that the deviant response is influenced by both the relative and the absolute stimulus probabilities. However, there is an alternative explanation for these findings, in that it could be that the regularity mechanism underlying the MMR treated the four standard stimuli as a group. The four standards are relatively close together in frequency, whilst the deviant stimulus is further apart (fig. 4.1). The enhancement of the deviant response might arise because the deviant occurred less often relative to the group of standards. The variability within the group of standards might have weakened the enhancement of the deviant response, accounting for the smaller response to the 4S:20 deviant than the 80:20 deviant.

The most striking finding arose when examining 4S:20 responses as a function of the immediately-preceding stimulus (fig. 4.4). It was noted earlier that the local effects on the 50:4D deviants were broadly tuned (SSI of 0.2). Surprisingly, the local effects on the 4S:20 standards appear to have been sharply tuned. The response to a standard preceded by a different standard ("sS") was similar to the standard response when preceded by a deviant ("DS"). The SSI value was 0.7, which was significantly different from zero but not from unity (95%-*t*-confidence interval: 0.3-1.2). A between-subjects permutation test (see Methods) showed that the SSI value for the 4S:20 standards was significantly greater (sharper tuning) than the SSI value for the 50:4D deviants ($p < 0.01$).

SR

As with the N1, there was a significant MMR in the 80:20 condition, in that the SR was larger for the deviant than the standard (fig. 4.3; $p < 0.001$). However, this arose solely from an enhancement of the deviant SR, relative to the 50:50 baseline ($p < 0.001$). Unlike the N1, the SR to the standard was not significantly suppressed ($p =$

0.297). Similarly, in both the 50:4D and 4S:20 conditions, the deviant response was enhanced ($p < 0.005$), but the standard response was unchanged (50:4D, $p = 0.836$; 4S:20, $p = 0.389$). The three deviant responses did not differ significantly from one another ($p > 0.05$). The SR was primarily sensitive to the global stimulus context. The response to a 50:50 stimulus was practically identical when the stimulus was preceded by the same frequency (12.3 nAm) as when it was preceded by a different frequency (12.2 nAm; permutation test, $p = 0.267$).

Interestingly, the size of the enhancement effect for deviants was a fairly small proportion of the overall sustained response. The average enhancement was 2.1 nAm, compared to the 12.3 nAm SR overall in the 50:50 condition. Thus, the enhancement was about 17% of the 50:50 amplitude. This contrasts with the N1, for which the enhancement for the 80:20 deviant was 65% of the 50:50 amplitude.

4.D DISCUSSION

The current study found striking differences between the contextual effects on the transient and sustained responses. This section will first address the effects on the transient responses. Effects on the sustained responses will be interpreted in the light of these findings.

Transient N1 response

The MMR in a traditional oddball sequence arose from both a suppression of the standard response and an enhancement of the deviant response. By presenting multiple standards or deviants in modified oddball sequences, the suppressive effect was found to reflect only the absolute stimulus probabilities, whilst the enhancement effect appeared to reflect both the relative and the absolute probabilities. By themselves, these findings are consistent with the traditional interpretation of the MMR, which assumes that the underlying regularity mechanism enhances deviant responses but does not affect responses to standards (Näätänen, 1992). However, analyses of the short-term effects on standard and deviant responses revealed a remarkable finding. The local suppressive effect of the immediately-preceding stimulus, which presumably reflects stimulus-specific adaptation, was *broadly* tuned for deviant stimuli but *sharply* tuned for standards. This result resembles the finding of Chapter 2 that the MMR arose partly from a change in the frequency specificity of adaptation with stimulus repetition. This change was attributed to a sharpening of the neural representation of the repeated (standard) stimulus. The current study suggests that the sharpening effect is sensitive to the relative stimulus probabilities. To be sensitive to the relative probabilities, the sharpening mechanism must also be sensitive to the global stimulus context.

The following model is proposed to account for the results of the current study. There may be a global mechanism that can group stimuli over a variable spectral region, in order to identify a cluster of stimuli that occurs relatively frequently with respect to the other stimuli in the sequence. This element of the model resembles the regularity-detection mechanism that has previously been proposed to underlie the MMR (Winkler et al., 1996; Schröger, 1997). The global mechanism would then sharpen the neural representations of stimuli that belong to the group of frequent stimuli (standards). As a result, the amount by which standards adapt infrequent stimuli (deviants) would be reduced and deviants would elicit a larger response (contributing to the MMR). In contrast, the representations of deviants remain unchanged.

The global mechanism may use Gestalt principles (Koffka, 1935) when identifying the standard group. For example, the Gestalt "law of proximity" states that sensory elements (e.g. visual lines or acoustic frequencies) that are close together are more likely to be grouped than elements that are further apart. The influence of this principle can be seen when the acoustic frequencies used in the current study are plotted as visual lines (fig. 4.1). The four standards/deviants appear to form a perceptual group, whilst the single standard/deviant stands by itself. Presumably, when identifying the standard group, the global mechanism would prefer to group stimuli over the smallest possible spectral region. Thus, in the 50:4D condition, the mechanism would select the single, 50% standard, as it occurred more often than any other stimulus. As no individual stimulus occurred relatively often than the other stimuli in the 4S:20 condition, the mechanism would select the group of four standards due to their close frequency spacing. This is despite the fact that the "4S" stimuli were identical to the "4D" stimuli in terms of spectral frequency. The global

mechanism would take into account the relationships between all the stimuli in a sequence. This is compatible with the finding of Sussman et al. (2003) that, of three stimuli presented with different probabilities, only the most frequent stimulus was treated as a standard.

An advantage of the current model is that it would be able to deal with variability in the auditory environment. The global mechanism could still identify a stimulus as a standard even if the stimulus differed somewhat from presentation to presentation; the mechanism would include the different exemplars in the standard group. This would be an important feature in everyday environments, as stimuli vary somewhat across instances. It seems likely that the global mechanism is able to group stimuli across dimensions other than spectral frequency also. For example, Winkler et al. (1990) found a significant MMR using a standard stimulus that varied in intensity over nine different values, even though individual values occurred with the same probability (10%) as the deviant stimulus. Moreover, significant MMRs have been found for a diverse array of stimulus features, such as intensity, duration and location (for a review, see Näätänen et al., 2007). Whether the global mechanism changes the representations of different features in similar ways remains to be examined.

The current model differs from those previously proposed to account for the MMR in that it places greater importance on the standard stimulus. The global mechanism sharpens the neural representation of the standard. As noted in Chapter 2, the sharpening effect is reminiscent of a proposed neural correlate of sensory priming. Priming is an automatic form of perceptual learning, which manifests itself in faster, or more accurate, processing of a primed stimulus (i.e., a stimulus that has been presented before) than a novel stimulus (Tulving & Schacter, 1990). A number of authors have suggested that priming results from a sharpening of the neural

representation of the repeated stimulus (Ungerleider, 1995; Desimone, 1996; Wiggs & Martin, 1998). It may, therefore, be that the global mechanism proposed in the current model represents a form of short-term perceptual learning. By taking into account the relationships amongst stimuli in the auditory environment, the global mechanism would be able to focus on the most prominent stimulus (or group of stimuli). The most prominent stimulus is likely the most important stimulus to learn. For example, in a given language, certain vowel sounds will feature prominently. By sharpening the neural representations of these vowels, the auditory system may increase their discriminability.

As in traditional models of the MMR, the current model postulates a global mechanism that is sensitive to the relative stimulus probabilities. However, unlike traditional models, the current model attributes the MMR to an interaction between the global mechanism and stimulus-specific adaptation; the global mechanism is assumed to modulate adaptation *indirectly* by altering stimulus representations. Friston (2005) and Garrido et al. (2009) proposed that the MMR arises from an interaction between adaptation and top-down, predictive processes generated in frontal brain regions. It may be that the sharpening effect found in the current study is mediated by top-down modulation from frontal regions. Alternatively, the sharpening effect could be generated by a distinct mechanism in the sensory-driven area itself. Indeed, frontal contributions to the MMR appear to be more closely related to the reorienting of attention to highly deviant stimuli (Alho, 1995; Rinne et al., 2000; Escera & Corral, 2007).

Sustained response

An MMR was also found for the sustained component of the auditory event-related potential. However, unlike the transient MMR, the sustained MMR in the traditional oddball sequence arose solely from an enhancement of the deviant response. There was no concomitant suppression of the standard response. Likewise, in the 50:4D and 4S:20 conditions, deviant responses were enhanced whilst standard responses were unaffected. Further, the SR seemed to be dominated by the global stimulus context. The results can be understood by postulating a global mechanism that, in its initial stages, operates similarly to the mechanism proposed to account for the transient MMR. The global mechanism would group stimuli over a variable spectral region to identify the group of frequent stimuli, or standards. After this "regularity-detection" stage, the effects on the transient and sustained responses may be different. The sharpening effect, proposed to account for the transient MMR, depends on the presence of a local, suppressive effect. As the suppressive effect appeared to be absent for the sustained MMR, it may be that the global mechanism actively enhances the SR for stimuli that do not belong to the standard group.

According to the above model, the SR should be sensitive solely to the relative stimulus probabilities. Gutschalk et al. (2007) measured neural responses to long sequences of stimuli that alternated between two pitch strengths or pitch values, without a silent ISI. The size of the SR depended on the relative, rather than the absolute, durations of the two stimuli. This finding is consistent with those of the current study, despite the differences in the stimuli and paradigms. For example, Gutschalk et al. (2007) found that the SR was the same whether the two stimuli had durations of 720 ms or 1440 ms. When one stimulus had a duration of 720 ms, and the other 2160 ms, the response to the shorter stimulus was enhanced, relative to the

responses in the conditions in which the two stimuli had equal durations. However, the response to the longer stimulus was suppressed, an effect not found in the current study. The reasons for this discrepancy should be addressed in subsequent studies.

It is unknown whether the SRs in the current study came from the same group of neurons that generated the transient responses. MEG and EEG studies have typically localised the SR anterior to the transient N1 (Arthur et al., 1987; Scherg et al. 1989). Recently, Kretzschmar & Gutschalk (2010) recorded the SR in an oddball paradigm and found two sources in each auditory cortex, an anterior and a posterior source. The response from the anterior source displayed an MMR, with smaller amplitude for standards than deviants. In contrast, there was no MMR in the response from the posterior source, despite a large SR. Both anterior and posterior sources elicited a transient MMR. This suggests a partial dissociation between the transient and sustained responses. The results of Kretzschmar & Gutschalk (2010) might also explain why the effects on the SR in the current study were comparatively small; part of the SR, from the posterior source, may have been insensitive to stimulus context, or may have shown only stimulus-unspecific effects. Interestingly, Wang and colleagues (Barbour & Wang, 2003; Wang et al., 2005) found that the majority of auditory cortical neurons could elicit both sustained and transient responses. However, the sustained and transient responses were functionally dissociable. SRs were highly stimulus specific, whilst transient responses were more broadly tuned. It may be that the contextual effects on the two response components are likewise dissociable.

It is suggested that the global mechanism actively enhances the SR for deviant stimuli. The physiological processes that could underlie this enhancement are currently unclear. Bartlett & Wang (2005) found that many auditory cortical neurons elicited a larger response to their preferred stimulus when it was preceded by a

dissimilar stimulus than when it was presented alone. The range of preceding stimuli that led to this response "facilitation" varied from neuron to neuron. Interestingly, facilitation often began in the SR and lasted for the duration of the stimulus. However, it remains unclear whether this effect has a long enough time course for it to be sensitive to the relationships amongst stimuli in a sequence, particularly for sequences in which the standard stimulus is variable. In the majority of neurons, facilitation was absent for ISIs above about one second. Eytan et al. (2003) found longer-lasting facilitation when an oddball paradigm was presented, using electrical stimulation, to a network of cortical neurons developing *ex vivo*. The standard and deviant were presented every 5 and 50 seconds, respectively. Despite the slow stimulation rates, the standard response was suppressed over time, and the deviant response was enhanced, relative to the initial response sizes. The authors attributed this finding to a change in the balance of excitation and inhibition in the neural network, resulting from synaptic depression to the standard. However, in such a mechanism, the suppression of the standard response and the facilitation of the deviant response would be intrinsically linked. It remains to be seen whether suppression of the transient response could lead to facilitation of the SR.

Conclusions

The current results suggest that the transient MMR reflects an interaction between local and global contextual mechanisms. Local suppressive effects were more sharply tuned for stimuli that occurred relatively frequently than for stimuli that occurred relatively infrequently. Remarkably, the global mechanism was able to deal with variability in the standard stimulus. This suggests that it may play a role in the perceptual learning of prominent stimuli in the acoustic environment. Sustained

responses appeared to be influenced by only a global mechanism, which enhanced responses to stimuli that occurred relatively infrequently. The functional role of this effect, and its relationship to the effects on the transient responses, await further investigation.

General conclusions

The current project comprised four EEG studies that investigated the effects of stimulus context on auditory cortical responses. In **Chapter 1**, one of the principal contextual effects, stimulus-specific adaptation, was used to probe the neural representation of musical pitch. Previous fMRI studies have localised pitch processing to an area anterior and lateral to primary auditory cortex (Patterson et al., 2002; Penagos et al., 2004; Hall et al., 2006; Puschmann et al., 2009). However, it remained unclear how pitch was represented in this area. Using the adaptation paradigm, the current study found evidence for a cortical representation of pitch chroma. The chroma-sensitive response occurred around 40 ms later than the spectrally-sensitive response, and was generated in a more anterolateral region of auditory cortex. It seems likely, therefore, that the chroma-sensitive response originated in the non-primary, anterolateral pitch area identified previously with fMRI.

Whilst contextual effects are typically thought to reflect stimulus representations, the second study, reported in **Chapter 2**, found evidence that they may also modify representational information. The second study examined the position of Jääskeläinen et al. (2004) that the MMR can be accounted for solely by a passive release from adaptation. Responses to probe stimuli after single adaptors were used to predict the contribution of adaptation to probe responses after multiple adaptors. Whilst the predictions fit the data near-perfectly when the probe was identical to the adaptors, predictions underestimated responses to deviant probes. This finding refutes the position of Jääskeläinen et al. (2004), providing evidence for a "true" MMR. However, contrary to traditional theories, the MMR reflected a modulation of the sensory-driven response to the deviant stimuli, rather than an independent, or endogenous, response (see Näätänen, 1992). Plotting the probe

response as a function of the acoustical difference between the adaptor and probe suggested that the enhanced deviant response came about through a sharpening of the neural representation of the adaptor with repeated presentation. A similar mechanism has previously been proposed to account for sensory priming in the visual system (e.g. Wiggs & Martin, 1998). By analogy, this suggests that the MMR might play a role in auditory perceptual learning. The study presented in **Chapter 2** also found sustained activity during a stimulus sequence, which localised anterior to the transient, sensory-driven responses. This sustained activity might reflect a monitoring mechanism involved in orienting attention to unexpected stimuli in the environment (Downar et al., 2000), or sharpening neural responses to the stimuli that occur frequently.

The third study, reported in **Chapter 3**, examined the memory spans of the contextual effects on auditory responses in the oddball paradigm. Whilst the MMR is typically thought to reflect the longer-term ("global"), rather than the short-term ("local"), stimulus probabilities (Horváth & Winkler, 2004), the local and global probabilities are confounded in the traditional oddball paradigm. The third study dissociated the local and global stimulus probabilities by presenting stimuli in Markov sequences, and by examining responses in Bernoulli sequences as a function of the local stimulus history. Striking differences were observed between the contextual effects on the different components of the auditory evoked response. The P1 component appeared insensitive to either the local or the global stimulus probabilities. In contrast, the N1 appeared to be influenced by contextual effects with a memory span of the order of ten seconds, or even longer. The N1 effects were also rather complex, partially reflecting the global probability of short stimulus sequences. This is consistent with previous work on the negative component of the MMR (Sussman et al., 1999; Horváth et al., 2001). The effects could play a role in auditory streaming, or

the perceptual learning of sound patterns. The P2 appeared to be influenced both by longer-term effects as well as by a short-term suppressive effect that virtually abolished the P2 response to repeated stimuli. The magnitude and time course of this suppressive effect suggested that it could play a role in detecting changes (temporal edges) within sound patterns, and might thus contribute to the formation of auditory objects.

The final study, reported in **Chapter 4**, examined whether the MMR is sensitive to the relative, rather than the absolute, stimulus probabilities. This would be expected if the underlying mechanism takes into account the relationships amongst stimuli in a sequence. However, in the traditional oddball paradigm, there is only one standard stimulus and one deviant stimulus. As a result, the relative and absolute stimulus probabilities co-vary. The fourth study presented stimuli in modified oddball sequences, which contained multiple standard or deviant stimuli. Longer-duration stimuli (500 ms) were used to elicit both transient and sustained responses. For the transient N1 response, the most striking finding was that adaptation was broadly tuned to deviant stimuli, but sharply tuned to standards. This is consistent with the "sharpening" model put forward to account for the results of the second study (**Chapter 2**). The fourth study extended this model by demonstrating that the sharpening effect is sensitive to the relative stimulus probabilities, and is able to accommodate variability in the standard stimulus. If the sharpening mechanism plays a role in perceptual learning, these properties would enable the auditory system to optimise processing of the most prominent stimulus, or group of stimuli, in the environment. Interestingly, it seemed that the contextual effects on the sustained response could not be accounted for by the sharpening model. Instead, there appeared to be an active facilitation of the sustained response to deviant stimuli, although this

effect was much smaller than the effects on the transient N1 response. The facilitation was sensitive to the relative stimulus probabilities. The role of this effect remains unclear, though it provides evidence for a functional difference between the transient and sustained responses.

A major contribution of this work has been to properly consider the representation of the standard stimulus in the oddball paradigm. Previous studies have focussed on the deviant stimulus, under the assumption that the primary role of pre-attentive auditory processing is to act as an early warning system, alerting higher-order brain areas to significant changes in the environment. Whilst this may be an important function of pre-attentive processing, the current work suggests that one of the most well-studied correlates of pre-attentive processes, the MMR, actually arises due to a sharpening of the neural representation of the repeated standard stimulus, rather than due to an additional, change-related response to deviant stimuli. This sharpening could reflect automatic learning of the principal features of the acoustic environment, enabling an organism to distinguish between frequently-presented sounds, such as different vocalisations.

The four studies reported in this thesis were conducted on human participants, who were awake but not attending to the stimuli. Despite this, the contextual effects on auditory responses were found to both reflect and modify stimulus representations. The methods employed in the current study could thus provide powerful tools for investigating neural representations in groups of participants that would be difficult to test behaviourally. In the auditory domain, the results could also contribute to theories of streaming and object formation, as both phenomena are strongly influenced by stimulus context (Carlyon et al., 2001; Micheyl et al., 2005). Moreover, due to the ubiquity of contextual effects in the cortex, the results could contribute to general

theories of perceptual learning. One future direction for this work could be to examine the interactions between the contextual effects found in the current study and the effects of top-down attention. Fritz et al. (2003), recording from ferret auditory cortex, found rapid, facilitative changes in the neural representations of behaviourally-relevant stimuli. Recently, Kauramäki et al. (2007) provided evidence that attention increases the stimulus specificity of neurons tuned to the attended stimulus, in human auditory cortex. This effect resembles the sharpening effect found in the current project for frequent stimuli (standards). It may be that the effects of attention, and the contextual effects described in the current project, have independent influences on auditory responses. Alternatively, attention might modify the contextual effects, optimising them to suit the current task demands.

References

Abdi, H. (2007). Bonferonni and Šidák corrections for multiple comparisons. In N. J. Salkind (Ed.), *Encyclopedia of measurement and statistics* (pp. 103-107). Thousand Oaks, CA: Sage.

Adelman, J. P., & Herson, P. S. (2004). Making scents of olfactory adaptation. *Nature Neuroscience, 7*, 689-690.

Adrian, E. D. (1928). *The basis of sensation.* London, UK: Christophers.

Agus, T. R., Thorpe, S. J., & Pressnitzer, D. (2010). Rapid formation of robust auditory memories: insights from noise. *Neuron, 66*, 610-618.

Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., Levänen, S., Lin, F.-H., Sams, M., Shinn-Cunningham, B. G., Witzel, T., & Belliveau, J. W. (2006). Task-modulated "what" and "where" pathways in human auditory cortex. *Proceedings of the National Academy of Sciences USA, 103*, 14608-14613.

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

ANSI. (1994). *American National Standard Acoustic Terminology.* New York, NY: American National Standards Institute.

Arthur, D. L., Flynn, E. R., & Williamson, S. J. (1987). Source localization of long-latency auditory evoked magnetic fields in human temporal cortex. *Electroencephalography and Clinical Neurophysiology Supplement, 40*, 429-439.

Asari, H., & Zador, A. M. (2009). Long-lasting context dependence constrains neural encoding models in rodent auditory cortex. *Journal of Neurophysiology, 102*, 2638-2656.

Atienza, M., Cantero, J. L., & Dominguez-Marin, E. (2002). The time course of neural changes underlying auditory perceptual learning. *Learning and Memory, 9,* 138-150.

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

Bachem, A. (1937). Various types of absolute pitch. *Journal of the Acoustical Society of America, 9,* 146-151.

Barbour, D. L., & Wang, X. (2003). Auditory cortical responses elicited in awake primates by random spectrum stimuli. *Journal of Neuroscience, 23,* 7194-7206.

Bartlett, E. L., & Wang, X. (2005). Long-lasting modulation by stimulus context in primate auditory cortex. *Journal of Neurophysiology, 94,* 83-104.

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation, 7,* 1129-1159.

Berens, P. (2009). CircStat: A Matlab toolbox for circular statistics. *Journal of Statistical Software, 31,* 1-21.

Bishop, D. V. M. (2007). Using mismatch negativity to study central auditory processing in developmental language and literacy impairments: Where are we, and where should we be going? *Psychological Bulletin, 133,* 651-672.

Bonhoeffer, T., & Grinvald, A. (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature, 353,* 429-431.

Brimijoin, W. O., & O'Neill, W. E. (2010). Patterned tone sequences reveal non-linear interactions in auditory spectrotemporal receptive fields in the inferior colliculus. *Hearing Research, 267,* 96-110.

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

Buchwald, J. S., Erwin, R., Lancker, D. V., Guthrie, D., Schwafel, J., & Tanguay, P. (1992). Midlatency auditory evoked responses: P1 abnormalities in adult autistic subjects. *Electroencephalography and Clinical Neurophysiology, 84*, 164-171.

Buchwald, J. S., Rubinstein, E. H., Schwafel, J., & Strandburg, R. J. (1991). Midlatency auditory evoked responses: Differential effects of a cholinergic agonist and antagonist. *Electroencephalography and Clinical Neurophysiology, 80*, 303-309.

Budd, T. W., & Michie, P. T. (1994). Facilitation of the N1 peak of the auditory ERP at short stimulus intervals. *Neuroreport, 5*, 2513-2516.

Burns, E. M. (1999). Intervals, scales, and tuning. In D. Deutsch (Ed.), *The psychology of music* (2nd ed., pp. 215-264). San Diego, CA: Academic Press.

Butler, R. A. (1968). Effect of changes in stimulus frequency and intensity on habituation of the human vertex potential. *Journal of the Acoustical Society of America, 44*, 945-950.

Calford, M. B., & Semple, M. N. (1995). Monaural inhibition in cat auditory cortex. *Journal of Neurophysiology, 73*, 1876-1891.

Carandini, M., Heeger, D. J., & Senn, W. (2002). A synaptic explanation of suppression in visual cortex. *Journal of Neuroscience, 22*, 10053-10065.

Cardenas, V. A., Gill, P., & Fein, G. (1997). Human P50 suppression is not affected by variations in wakeful alertness. *Biological Psychiatry, 41*, 891-901.

Carlyon, R. P. (1998). Comments on "A unitary model of pitch perception". *Journal of the Acoustical Society of America, 104*, 1118-1121.

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences, 8*, 465-471.

Carlyon, R. P., Cusack, R., Foxton, J. M., & Robertson, I. H. (2001). Effects of attention and unilateral neglect on auditory stream segregation. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 115-127.

Chase, S. M., & Young, E. D. (2007). First-spike latency information in single neurons increases when referenced to population onset. *Proceedings of the National Academy of Sciences USA, 104*, 5175-5180.

Chung, S., Li, X., & Nelson, S. B. (2002). Short-term depression at thalamocortical synapses contributes to rapid adaptation of cortical sensory responses in vivo. *Neuron, 34*, 437-446.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201-215.

Crottaz-Herbette, S., & Ragot, R. (2000). Perception of complex sounds: N1 latency codes pitch and topography codes spectra. *Clinical Neurophysiology, 111*, 1759-1766.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology, 115*, 732-744.

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

de Cheveigné, A. (2005). Pitch perception models. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception* (pp. 169-233). New York, NY: Springer.

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

Demany, L., & Armand, F. (1984). The perceptual reality of tone chroma in early infancy. *Journal of the Acoustical Society of America, 76*, 57-66.

Desimone, R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences USA, 93*, 13494-13499.

Deutsch, D. (1969). Music recognition. *Psychological Review, 76*, 300-307.

Deutsch, D. (1972). Octave generalization and tune recognition. *Perception and Psychophysics, 11*, 411-412.

Deutsch, D. (2006). The enigma of absolute pitch. *Acoustics Today, 2*, 11-19.

di Lorenzo, P. M., & Lemon, C. H. (2000). The neural code for taste in the nucleus of the solitary tract of the rat: Effects of adaptation. *Brain Research, 852*, 383-397.

Douglas, K. M., & Bilkey, D. K. (2007). Amusia is associated with deficits in spatial processing. *Nature Neuroscience, 10*, 915-921.

Dowling, W. J. (1999). The development of music perception and cognition. In D. Deutsch (Ed.), *The psychology of music* (pp. 603-625). San Diego, CA: Academic Press.

Downar, J., Crawley, A. P., Mikulis, D. J., & Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nature Neuroscience, 3*, 277-283.

Efron, B. (1993). *An introduction to the bootstrap.* New York, NY: Chapman & Hall.

Eggermont, J. J., & Roberts, L. E. (2004). The neuroscience of tinnitus. *Trends in Neurosciences, 27*, 676-682.

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

Eytan, D., Brenner, N., & Marom, S. (2003). Selective adaptation in networks of cortical neurons. *Journal of Neuroscience, 23*, 9349-9356.

Formisano, E., Kim, D.-S., Salle, F. D., van de Moortele, P.-F., Ugurbil, K., & Goebel, R. (2003). Mirror-symmetric tonotopic maps in human primary auditory cortex. *Neuron, 40*, 859-869.

Freeman, T. C. B., Durand, S., Kiper, D. C., & Carandini, M. (2002). Suppression without inhibition in visual cortex. *Neuron, 35*, 759-771.

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

Fritz, J., Shamma, S., Elhilali, M., & Klein, D. (2003). Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex. *Nature Neuroscience, 6*, 1216-1223.

Garrido, M. I., Kilner, J. M., Stephan, K. E., & Friston, K. J. (2009). The mismatch negativity: A review of underlying mechanisms. *Clinical Neurophysiology, 120*, 453-463.

Glasberg, B. R., & Moore, B. C. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*, 103-138.

Griffiths, T. D., Büchel, C., Frackowiak, R. S., & Patterson, R. D. (1998). Analysis of temporal structure in sound by the human brain. *Nature Neuroscience, 1*, 422-427.

Griffiths, T. D., Uppenkamp, S., Johnsrude, I., Josephs, O., & Patterson, R. D. (2001). Encoding of the temporal regularity of sound in the human brainstem. *Nature Neuroscience, 4*, 633-637.

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences, 25*, 348-353.

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

Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences, 10*, 14-23.

Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica, 107*, 293-321.

Gutnisky, D. A., Hansen, B. J., Iliescu, B. F., & Dragoi, V. (2009). Attention alters visual plasticity during exposure-based learning. *Current Biology, 19*, 555-560.

Gutschalk, A., Patterson, R. D., Rupp, A., Uppenkamp, S., & Scherg, M. (2002). Sustained magnetic fields reveal separate sites for sound level and temporal regularity in human auditory cortex. *Neuroimage, 15*, 207-216.

Gutschalk, A., Patterson, R. D., Scherg, M., Uppenkamp, S., & Rupp, A. (2007). The effect of temporal context on the sustained pitch response in human auditory cortex. *Cerebral Cortex, 17*, 552-561.

Hall, D. A., Edmondson-Jones, A. M., & Fridriksson, J. (2006). Periodicity and frequency coding in human auditory cortex. *European Journal of Neuroscience, 24*, 3601-3610.

Hall, D. A., Johnsrude, I. S., Haggard, M. P., Palmer, A. R., Akeroyd, M. A., & Summerfield, A. Q. (2002). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex, 12*, 140-149.

Hall, D. A., & Plack, C. J. (2007). The human 'pitch center' responds differently to iterated noise and Huggins pitch. *Neuroreport, 18*, 323-327.

Hall, D. A., & Plack, C. J. (2009). Pitch processing sites in the human auditory brain. *Cerebral Cortex, 19*, 576-585.

Haynes, J.-D., & Rees, G. (2005). Predicting the orientation of invisible stimuli from activity in human primary visual cortex. *Nature Neuroscience, 8*, 686-691.

He, C., & Trainor, L. J. (2009). Finding the pitch of the missing fundamental in infants. *Journal of Neuroscience, 29*, 7718-8822.

Horváth, J., Czigler, I., Sussman, E., & Winkler, I. (2001). Simultaneously active pre-attentive representations of local and global rules for sound sequences in the human brain. *Cognitive Brain Research, 12*, 131-144.

Horváth, J., & Winkler, I. (2004). How the human auditory system treats repetition amongst change. *Neuroscience Letters, 368*, 157-161.

Huotilainen, M., Winkler, I., Alho, K., Escera, C., Virtanen, J., Ilmoniemi, R. J., Jääskeläinen, I. P., Pekkonen, E., & Näätänen, R. (1998). Combined mapping of human auditory EEG and MEG responses. *Electroencephalography and Clinical Neurophysiology, 108*, 370-379.

Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., Lin, F.-H., May, P., Melcher, J., Stufflebeam, S., Tiitinen, H., & Belliveau, J. W. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences USA, 101*, 6809-6814.

Jacobsen, T., & Schröger, E. (2001). Is there pre-attentive memory-based comparison of pitch? *Psychophysiology, 38*, 723-727.

Jerger, K., Biggins, C., & Fein, G. (1992). P50 suppression is not affected by attentional manipulations. *Biological Psychiatry, 31*, 365-377.

Jesteadt, W., Bacon, S. P., & Lehman, J. R. (1982). Forward masking as a function of frequency, masker level, and signal delay. *Journal of the Acoustical Society of America, 71*, 950-962.

Kalmus, H., & Fry, D. B. (1980). On tune deafness (dysmelodia): Frequency, development, genetics and musical background. *Annals of Human Genetics, 43*, 369-382.

Kane, N. M., Butler, S. R., & Simpson, T. (2000). Coma outcome prediction using event-related potentials: P3 and mismatch negativity. *Audiology and Neurootology, 5*, 186-191.

Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature, 377*, 155-158.

Kauramäki, J., Jääskeläinen, I. P., & Sams, M. (2007). Selective attention increases both gain and feature selectivity of the human auditory cortex. *PLoS One, 2*, e909.

Koffka, K. (1935). *Principles of Gestalt Psychology.* New York: Harcourt, Brace & World.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97*, 3155-3164.

Kretzschmar, B., & Gutschalk, A. (2010). A sustained deviance response evoked by the auditory oddball paradigm. *Clinical Neurophysiology, 121*, 524-532.

Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lütkenhöner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex, 13*, 765-772.

Kubovy, M., & van Valkenburg, D. V. (2001). Auditory and visual objects. *Cognition, 80*, 97-126.

Langner, G., Albert, M., & Briede, T. (2002). Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla (Chinchilla laniger). *Hearing Research, 168*, 110-130.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation, 11*, 417-441.

Levitin, D. J., & Rogers, S. E. (2005). Absolute pitch: Perception, coding, and controversies. *Trends in Cognitive Sciences, 9*, 26-33.

Lütkenhöner, B., & Steinsträter, O. (1998). High-precision neuromagnetic study of the functional organization of the human auditory cortex. *Audiology and Neurootology, 3*, 191-213.

Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., & Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proceedings of the National Academy of Sciences USA, 94*, 10979-10984.

Mäkelä, J. P., Hämäläinen, M., Hari, R., & McEvoy, L. (1994). Whole-head mapping of middle-latency auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology, 92*, 414-421.

McAlpine, D. (2004). Neural sensitivity to periodicity in the inferior colliculus: Evidence for the role of cochlear distortions. *Journal of Neurophysiology, 92*, 1295-1311.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2010). Individual differences reveal the basis of consonance. *Current Biology, 20*, 1035-1041.

McLaughlin, D. F., & Kelly, E. F. (1993). Evoked potentials as indices of adaptation in the somatosensory system in humans: A review and prospectus. *Brain Research Reviews, 18*, 151-206.

Meddis, R., & Hewitt, M. J. (1991a). Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery: I: Pitch identification. *Journal of the Acoustical Society of America, 89*, 2866-2882.

Meddis, R., & Hewitt, M. J. (1991b). Virtual pitch and phase sensitivity studied using a computer model of the auditory periphery: II: Phase sensitivity. *Journal of the Acoustical Society of America, 89*, 2883-2894.

Micheyl, C., & Oxenham, A. J. (2010). Objective and subjective psychophysical measures of auditory stream integration and segregation. *Journal of the Association for Research in Otolaryngology, 11*, 709-724.

Micheyl, C., Tian, B., Carlyon, R. P., & Rauschecker, J. P. (2005). Perceptual organization of tone sequences in the auditory cortex of awake macaques. *Neuron, 48*, 139-148.

Moore, B. C. J. (1978). Psychophysical tuning curves measured in simultaneous and forward masking. *Journal of the Acoustical Society of America, 63*, 524-532.

Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., & Zilles, K. (2001). Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. *Neuroimage, 13*, 684-701.

Näätänen, R. (1990). The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function. *Behavioral Brain Sciences, 13*, 201-288.

Näätänen, R. (1992). *Attention and brain function.* Hillsdale, NJ: Erlbaum.

Näätänen, R. (2003). Mismatch negativity: Clinical research and possible applications. *International Journal of Psychophysiology, 48*, 179-188.

Näätänen, R., & Escera, C. (2000). Mismatch negativity: Clinical and other applications. *Audiology and Neurootology, 5*, 105-110.

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

Näätänen, R., Pakarinen, S., Rinne, T., & Takegata, R. (2004). The mismatch negativity (MMN): Towards the optimal paradigm. *Clinical Neurophysiology, 115*, 140-144.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: A review and an analysis of the component structure. *Psychophysiology, 24*, 375-425.

Näätänen, R., Sams, M., Järvilehto, T., & Soininen, K. (1983). Probability of deviant stimulus and event-related brain potentials. In R. Sinz, & M. R. Rosenzweig (Eds.), *Psychophysiology 1980* (pp. 397-405). Amsterdam, Netherlands: Elsevier.

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

Norris, J. R. (1998). *Markov Chains.* Cambridge, UK: Cambridge University Press.

Novitski, N., Huotilainen, M., Tervaniemi, M., Näätänen, R., & Fellman, V. (2007). Neonatal frequency discrimination in 250-4000-Hz range: Electrophysiological evidence. *Clinical Neurophysiology, 118,* 412-419.

Nunez, P. L., & Srinivasan, R. (2006). *Electric fields of the brain: The neurophysics of EEG.* New York: Oxford University Press.

Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlargement phenomenon. *Journal of the Acoustical Society of America, 73,* 1694-1700.

Opitz, B., Rinne, T., Mecklinger, A., von Cramon, D. Y., & Schröger, E. (2002). Differential contribution of frontal and temporal cortices to auditory change detection: fMRI and ERP results. *Neuroimage, 15,* 167-174.

Patterson, R. D. (1986). Spiral detection of periodicity and the spiral form of musical scales. *Psychology of Music, 14,* 44-61.

Patterson, R. D., Uppenkamp, S., Johnsrude, I. S., & Griffiths, T. D. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron, 36,* 767-776.

Pearce, J. M. S. (2005). Selected observations on amusia. *European Neurology, 54,* 145-148.

Pekkonen, E. (2000). Mismatch negativity in aging and in Alzheimer's and Parkinson's diseases. *Audiology and Neurootology, 5,* 216-224.

Penagos, H., Melcher, J. R., & Oxenham, A. J. (2004). A neural representation of pitch salience in nonprimary human auditory cortex revealed with functional magnetic resonance imaging. *Journal of Neuroscience, 24,* 6810-6815.

Peretz, I., Ayotte, J., Zatorre, R. J., Mehler, J., Ahad, P., Penhune, V. B., & Jutras, B. (2002). Congenital amusia: A disorder of fine-grained pitch discrimination. *Neuron, 33*, 185-191.

Peretz, I., & Hyde, K. L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences, 7*, 362-367.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978a). Human auditory sustained potentials. I. The nature of the response. *Electroencephalography and Clinical Neurophysiology, 45*, 186-197.

Picton, T. W., Woods, D. L., & Proulx, G. B. (1978b). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology, 45*, 198-210.

Plack, C. J., & Oxenham, A. J. (2005). The psychophysics of pitch. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception* (pp. 7-55). New York, NY: Springer.

Plack, C. J., Oxenham, A. J., & Drga, V. (2006). Masking by inaudible sounds and the linearity of temporal summation. *Journal of Neuroscience, 26*, 8767-8773.

Pressnitzer, D., & Patterson, R. D. (2001). Distortion products and the pitch of harmonic complex tones. In D. Breebaart, A. Houtsma, A. Kohlrausch, V. Prijs, & R. Schoonoven (Eds.), *Proceedings of the 12th international symposium on hearing, physiological and psychological bases of auditory function* (pp. 97-104). Maastrict, Netherlands: Shaker BV.

Puschmann, S., Uppenkamp, S., Kollmeier, B., & Thiel, C. M. (2010). Dichotic pitch activates pitch processing centre in Heschl's gyrus. *Neuroimage, 49*, 1641-1649.

Rinne, T., Alho, K., Ilmoniemi, R. J., Virtanen, J., & Näätänen, R. (2000). Separate time behaviors of the temporal and frontal mismatch negativity sources. *Neuroimage, 12*, 14-19.

Schellenberg, E. G., & Trehub, S. E. (1994). Frequency ratios and the discrimination of pure tone sequences. *Perception and Psychophysics, 56*, 472-478.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*, 235-248.

Scherg, M. (1990). Fundamentals of dipole source potential analysis. In F. Grandori, M. Hoke, & G. Romani (Eds.), *Auditory evoked potentials and fields: Advances in audiology* (pp. 40-69). Basel, Switzerland: Karger.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience, 1*, 336-355.

Schröger, E. (1997). On the detection of auditory deviations: A pre-attentive activation model. *Psychophysiology, 34*, 245-257.

Schröger, E., & Wolff, C. (1996). Mismatch response of the human brain to changes in sound location. *Neuroreport, 7*, 3005-3008.

Schulze, H., Hess, A., Ohl, F. W., & Scheich, H. (2002). Superposition of horseshoe-like periodicity and linear tonotopic maps in auditory cortex of the Mongolian gerbil. *European Journal of Neuroscience, 15*, 1077-1084.

Schulze, H., & Langner, G. (1997). Periodicity coding in the primary auditory cortex of the Mongolian gerbil (Meriones unguiculatus): Two different coding strategies for pitch and rhythm? *Journal of Comparative Physiology A, 181*, 651-663.

Seither-Preisler, A., Krumbholz, K., Patterson, R., Seither, S., & Lütkenhöner, B. (2004). Interaction between the neuromagnetic responses to sound energy

onset and pitch onset suggests common generators. *European Journal of Neuroscience, 19,* 3073-3080.

Shackleton, T. M., & Carlyon, R. P. (1994). The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *Journal of the Acoustical Society of America, 95,* 3529-3540.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28,* 1059-1074.

Squires, K. C., Wickens, C., Squires, N. K., & Donchin, E. (1976). The effect of stimulus sequence on the waveform of the cortical event-related potential. *Science, 193,* 1142-1146.

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

Sussman, E., Sheridan, K., Kreuzer, J., & Winkler, I. (2003). Representation of the standard: Stimulus context effects on the process generating the mismatch negativity component of event-related brain potentials. *Psychophysiology, 40,* 465-471.

Swartz, B. E., & Goldensohn, E. S. (1998). Timeline of the history of EEG and associated fields. *Electroencephalography and Clinical Neurophysiology, 106,* 173-176.

Talavage, T. M., Sereno, M. I., Melcher, J. R., Ledden, P. J., Rosen, B. R., & Dale, A. M. (2004). Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. *Journal of Neurophysiology, 91,* 1282-1296.

Tiitinen, H., Alho, K., Huotilainen, M., Ilmoniemi, R. J., Simola, J., & Näätänen, R. (1993). Tonotopic auditory cortex and the magnetoencephalographic (MEG) equivalent of the mismatch negativity. *Psychophysiology, 30*, 537-540.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301-306.

Ueda, K., & Ohgushi, K. (1987). Perceptual components of pitch: Spatial representation using a multidimensional scaling technique. *Journal of the Acoustical Society of America, 82*, 1193-1200.

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

Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science, 270*, 769-775.

van Hoey, G., Vanrumste, B., D'Havé, M., de Walle, R., Lemahieu, I., & Boon, P. (2000). Influence of measurement noise and electrode mislocalisation on EEG dipole-source localisation. *Medical and Biological Engineering and Computing, 38*, 287-296.

Waldo, M., Gerhardt, G., Baker, N., Drebing, C., Adler, L., & Freedman, R. (1992). Auditory sensory gating and catecholamine metabolism in schizophrenic and normal subjects. *Psychiatry Research, 44*, 21-32.

Wang, A. L., Mouraux, A., Liang, M., & Iannetti, G. D. (2008a). The enhancement of the N1 wave elicited by sensory stimuli presented at very short inter-stimulus intervals is a general feature across sensory systems. *PLoS One, 3*, e3929.

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

Wang, X., Lu, T., Snider, R. K., & Liang, L. (2005). Sustained firing in auditory cortex evoked by preferred stimuli. *Nature, 435*, 341-346.

Wang, Y., & Gotman, J. (2001). The influence of electrode location errors on EEG dipole source localization with a realistic head model. *Clinical Neurophysiology, 112*, 1777-1780.

Warren, J. D., Uppenkamp, S., Patterson, R. D., & Griffiths, T. D. (2003). Separating pitch chroma and pitch height in the human brain. *Proceedings of the National Academy of Sciences USA, 100*, 10038-10042.

Wehr, M., & Zador, A. M. (2005). Synaptic mechanisms of forward suppression in rat auditory cortex. *Neuron, 47*, 437-445.

Wiggs, C. L., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current Opinion in Neurobiology, 8*, 227-233.

Winkler, I., Karmos, G., & Näätänen, R. (1996). Adaptive modeling of the unattended acoustic environment reflected in the mismatch negativity event-related potential. *Brain Research, 742*, 239-252.

Winkler, I., Paavilainen, P., Alho, K., Reinikainen, K., Sams, M., & Näätänen, R. (1990). The effect of small variation of the frequent auditory stimulus on the event-related brain potential to the infrequent stimulus. *Psychophysiology, 27*, 228-235.

Winter, I. M. (2005). The neurophysiology of pitch. In C. J. Plack, A. J. Oxenham, R. R. Fay, & A. N. Popper (Eds.), *Pitch: Neural coding and perception* (pp. 99-146). New York, NY: Springer.

Woldorff, M. G. (1993). Distortion of ERP averages due to overlap from temporally adjacent ERPs: Analysis and correction. *Psychophysiology, 30*, 98-119.

Yost, W. A. (1996). Pitch of iterated rippled noise. *Journal of the Acoustical Society of America, 100,* 511-518.

Yost, W. A., Patterson, R., & Sheft, S. (1996). A time domain description for the pitch strength of iterated rippled noise. *Journal of the Acoustical Society of America, 99,* 1066-1078.

Zentner, M. R., & Kagan, J. (1998). Infants' perception of consonance and dissonance in music. *Infant Behaviour & Development, 21,* 483-492.