We are IntechOpen,
the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Processing of Binaural Information in Human Auditory Cortex

Blake W. Johnson

Macquarie Centre for Cognitive Science, Macquarie University, Sydney
Australia

1. Introduction

The mammalian auditory system is able to compute highly useful information by analyzing slight disparities in the information received by the two ears. Binaural information is used to build spatial representations of objects and also enhances our capacity to perform a fundamental structuring of perception referred to as 'auditory scene analysis' (Bregman, 1990) involving a parsing of the acoustic input stream into behaviourally-relevant representations. In a world that contains a cacophony of sounds, binaural hearing is employed to separate out concurrent sound sources, determine their locations, and assign them meaning. In the last several years our group has studied how binaural information is processed in the human auditory cortex, using a psychophysical paradigm to elicit binaural processing and using electroencephalography (EEG) and magnetoencephalography (MEG) to measure cortical function.

In our psychophysical paradigm listeners are posed with monaurally identical broadband sounds containing a timing or level disparity restricted to a narrow band of frequencies within their overall spectra. This results in the perception of a pitch corresponding to the affected frequency band, concurrent with, but spatially separated from, the remaining background (Yost, 1991). The illusion of "hearing out" (termed "dichotic pitch") has a close analogy in the visual system, where retinal disparities in random dot stereograms can be used to achieve the "seeing out" of a shape displaced in depth from a random background (Julesz, 1971).

Using EEG and MEG to measure brain activity in human listeners, we have found that the hearing out of dichotic pitches elicits a sequence of auditory cortical responses over a time window of some 150-400 ms after the onset of a dichotically-embedded pitch. In a series of experiments (Johnson et al., 2003; Hautus & Johnson, 2005; Johnson et al., 2007; Johnson & Hautus, 2010) we have shown that these responses correspond to functionally distinct stages of auditory scene analysis. Our data provide new insights into the nature, sequencing and timing of those stages.

2. Dichotic pitch paradigm

Dichotic pitch is a binaural unmasking phenomenon that is theoretically closely related to the masking level difference (MLD), and involves the perception of pitches from stimuli that

contain no monaural cues to pitch (Bilsen, 1976; Cramer & Huggins, 1958). Dichotic pitch can be produced by presenting listeners with two broadband noises with interaurally identical amplitude spectra but with a specific interaural lag over a narrow frequency band (Dougherty et al., 1998). The interaurally-shifted frequency band becomes perceptually segregated from the noise, and the resulting pitch has a tonal quality associated with the centre frequency of the dichotically-delayed portion of the spectrum. Because the stimuli are discriminable solely by the interaural lag but are otherwise acoustically identical, the perception of dichotic pitch must ultimately depend upon the binaural fusion of interaural time differences (ITDs) within the central auditory system. The phenomenon of dichotic pitch demonstrates that the human auditory system applies its exquisite sensitivity for the fine-grained temporal structure of sounds to the perceptual segregation, localization, and identification of concurrently-presented sound sources.

Fig. 1 shows how dichotic pitches can be generated using a complementary filtering method described by Dougherty et al. (1998). Two independent broadband Gaussian noise processes, 500-ms in duration are digitally constructed, in this case with a sampling rate of 44,100 Hz. One noise process is bandpass filtered with a centre frequency of 600 Hz and 3-dB bandwidth of 50 Hz using a 4th-order Butterworth filter with corner frequencies of 575 and 625 Hz (Fig. 1: middle panels). The other noise process is notch filtered using the same corner frequencies as the bandpass filter (Figure 1: left panels). The sum of the filter functions for the notch and bandpass filters is equal to one for all frequencies.

The bandpass-filtered noise process is duplicated and, to produce the dichotic-pitch stimuli, one copy of the noise process is delayed by 500 µs. Control stimuli contain no delay. The notch and bandpass filtered noise processes are recombined, producing two spectrally flat noise processes, which are again bandpass filtered (4th-order Butterworth) with corner frequencies of 400 and 800 Hz (Fig. 1: right panels). All stimuli are windowed using a cos2 function with 4-ms rise and fall times. In our laboratory auditory stimuli are generated on two channels of a 16-bit converter (Model DAQPad 6052E, National Instruments, Austin, Texas, USA). Programmable attenuators (Model PA4, Tucker-Davis Technologies, Alachua, Florida, USA) set the binaural stimuli to 70 dB SPL prior to their delivery via earphones (In our lab, Etymotic insert earphones Model ER2 or ER3, Etymotic Research Inc., Elk Grove Village, Illinois, USA). For sequences of stimuli, a jittered interstimulus (offset to onset) interval is drawn from a rectangular distribution between 1000 and 3000 ms.

Comparable dichotic pitch perceptions can be elicited using interaural level (ILD) rather than timing differences. To produce ILD dichotic pitch, the relative amplitude of the two bandpass noises is adjusted to increase the level in one channel while reducing the level in the other, and the same is done for the two notched noises. The two noises for each channel are combined as for the ITD stimuli.

Fig. 2 illustrates some of the perceptions that can be evoked by dichotic pitch stimuli presented via earphones. Control stimuli (top row) contain an interaural time disparity (ITD) that is uniform over the entire frequency spectrum of the noise stimuli and results in a single percept of noise (represented as ###) lateralized to the side of the temporally leading ear. Dichotic pitch stimuli (bottom row) contain interaural disparities that are oppositely directed for a narrow notch of frequencies (e.g. 575-625 Hz) versus the remainder of the frequency spectrum. These stimuli evoke a perception of two concurrent but spatially separated sounds lateralized to opposite sides: a dichotic pitch (represented as a musical note) with a perceived pitch corresponding to the centre frequency binaurally delayed notch
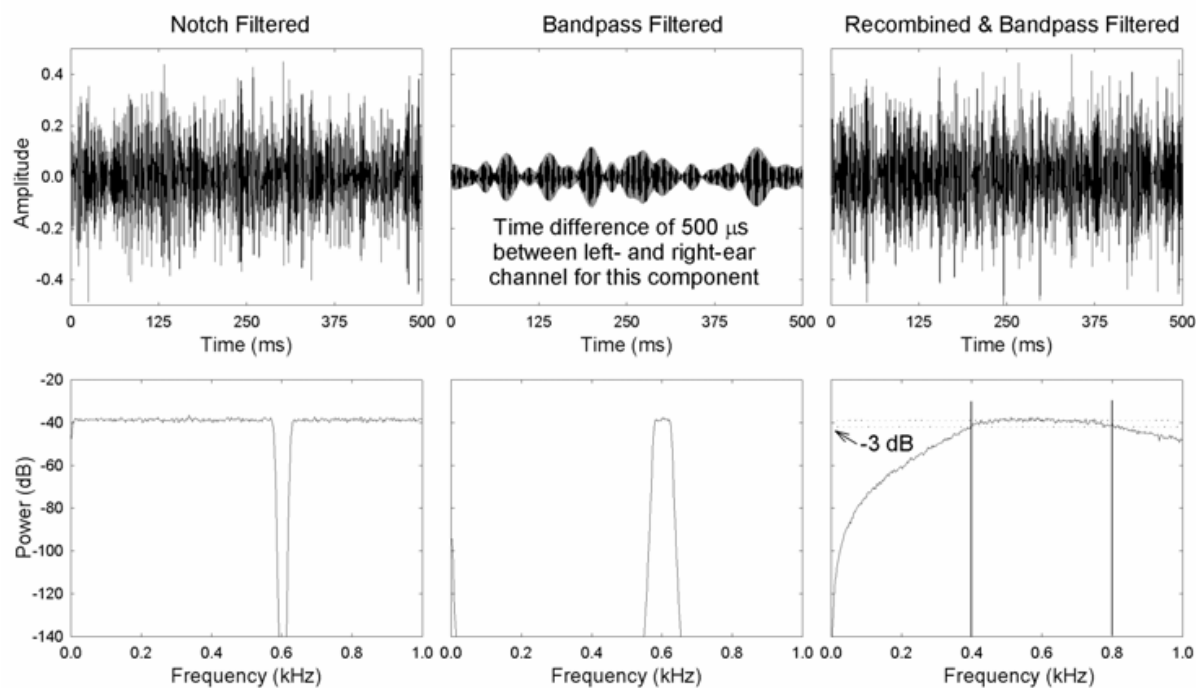
Fig. 1. Temporal and spectral representations of dichotic pitch stimulus. From Johnson et al., (2003) with permission.
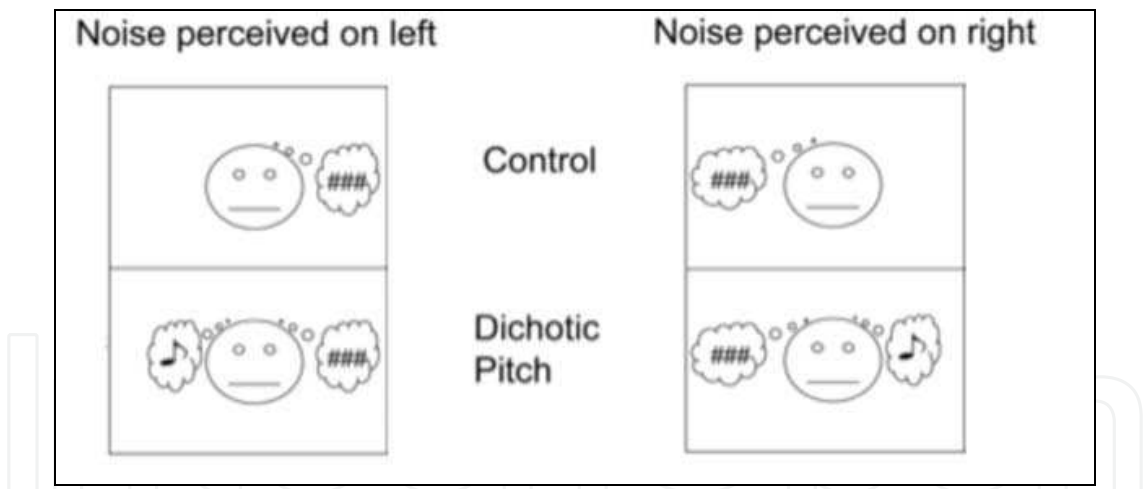

Fig. 2. Experimental stimuli and percepts of a listener. Adapted from Johnson and Hautus (2010) with permission.

(600 Hz) and a background noise corresponding to the remainder of the noise spectrum. From the point of view of an auditory researcher, the dichotic pitch paradigm has a number of features that make it useful for probing the workings of the central auditory system:

1. For experiments with dichotic pitch the control stimulus is simply one that has a uniform interaural disparity over its entire frequency range. Since the control and dichotic pitch stimuli are monaurally identical, any differences in perception or measured brain activity can be confidently attributed to differences in binaural processing;

2. Interaural disparities are first computed at the level of the medial superior olive in the brainstem (Goldberg & Brown, 1969; Yin & Chan, 1990) so perception of dichotic pitch can be confidently attributed to central rather than peripheral processes;
3. The perception of dichotic pitch depends on the ability of the auditory system to compute, encode, and process very fine temporal disparities (microseconds) and so provides a sensitive index of the temporal processing capabilities of the binaural auditory system. Consequently, dichotic pitch has been used to study clinical disorders such as dyslexia, that are suspected to involve central problems in auditory temporal processing (Dougherty et al., 1998);
4. The overall perceptual problem posed by dichotic pitch – that of separating a behaviourally relevant sound from a background noise or, more generally, that of segregating concurrent sound objects – is of considerable interest to those interested in how, and by what mechanisms, the brain is able to accomplish this important structuring of perception (Alain, 2007; Bregman, 1990).

Before proceeding to review experimental studies, we digress in the next section to describe for non-specialists the two main technologies used to measure auditory brain function in these studies, namely electroencephalography (EEG) and magnetoencephalography (MEG) and to introduce some terminology pertinent to these techniques.

3. EEG and MEG for measuring central auditory function

The methodologies for measuring brain function merit some consideration in any review of empirical studies, since the choice of method determines the type of brain activity measured (e.g. neuroelectric versus hemodynamic responses) and the spatial and temporal resolution of the measurements. These factors have a large impact on the types of inferences that can be derived from measured brain data.

Roughly speaking, EEG and MEG are the methods of choice when temporal resolution is an important or paramount requirement of a study. The reason for this is that the electromagnetic fields measured by these techniques are directly and instantaneously generated by ionic current flow in neurons. In contrast, positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) techniques measure the indirect, and temporally sluggish, metabolic and hemodynamic consequences of neuronal activity. Consequently PET and fMRI have inherently coarse temporal resolutions, on the order of one to many seconds. EEG and MEG are often described as having "millisecond" temporal resolution, but this is a technical limit imposed by the sampling capabilities of analogue-to-digital converters: by the nature of the measurements EEG and MEG can theoretically track ionic currents as fast as they occur. In practice though there are a number of additional limitations to the temporal capabilities of EEG and MEG: for example, time series are typically averaged over spans of tens or even hundreds of ms to improve the reliability of measurements and to reduce the dimensionality of the data. Even so, EEG-MEG are the methods of choice when one studies brain events that change rapidly and dynamically over time. For example, EEG-MEG techniques have long been an essential tool of psycholinguists studying the brain processes associated with language (Kutas et al., 2006).

The very properties that confer a high temporal resolution to impart fundamental limits on spatial resolution and EEG-MEG are generally considered inferior to PET and fMRI for localizing brain events in space. For both techniques the algebraic summation of electromagnetic fields limits their ability to resolve concurrent and closely-spaced neuronal

events. MEG has certain advantages over EEG in this regard because magnetic fields are not altered by conductive anisotropies and inhomogeneities. There are also advantages conferred to MEG by the fact that it is relatively less sensitive to distant sources and to neuronal sources in the crests of gyri (since these are oriented radially to the skull their magnetic fields do not exit the head). This lack of sensitivity is advantageous because MEG measurements present a relatively simpler picture of brain activity for researchers to interpret: simply put, there are fewer contributing brain sources that must be disentangled from measurements of fields on the surface of the head.

EEG-MEG measurements are typically carried out in event-related experimental designs in which stimuli are presented repeatedly (tens to hundreds or thousands of trials) and measurements are averaged over repeated trials to increase the signal-to-noise ratio of brain responses. In the case of EEG averaged signals are referred to as event-related potentials (ERPs) or evoked potentials (EPs). ERPs recorded on the surface of the head are often averaged across subjects as well to produce "grand-averaged" ERPs. In the case of MEG averaged signals are referred to as event-related magnetic fields (ERFs) but these are typically not analyzed as grand averages. This is because the higher spatial resolution of MEG means that it is not reasonable to assume that a given MEG sensor will record the same configuration of brain activations from subject to subject. For this reason MEG data is typically rendered into "source space" by computing the brain sources of the surface-recorded data, before performing descriptive and inferential statistics. Source analysis of EEG data is also possible and this is increasingly done by researchers. However the EEG source analysis problem is somewhat more complicated because of the need to specify the resistive parameters of the various tissue compartments of the head and brain. A final but essential piece of EEG-MEG nomenclature pertains to the naming of landmarks within ERP-ERF time series. ERP averages are presented as voltage deflections over time and deflections are named according to their polarity and latency (for example, "P100" may refer to a positive deflection at a latency of 100 ms after stimulus onset) or polarity and relative timing in a sequence (for example, P1-N1-P2 refers to a sequence of a positive and a negative and another positive deflection). ERP-ERF are also roughly subdivided into "middle" (about 20-70 ms) and "late" latency responses (greater than 80 ms or so). ERPs contain a third class of "early" (less than 10 ms) responses generated in CN VIII and the auditory brainstem. Because of the distance, MEG sensors are relatively insensitive to the sources of these early responses. Although this approach will not be further discussed in this review, we note in passing that it is also informative to analyse the frequency content of EEG and MEG signals and these are computed as "event-related spectral perturbations" (ERSPs).

4. Brain responses to dichotic pitch: the ORN

4.1 Passive listening conditions

Fig. 3 illustrates brain responses to dichotic pitch and control sounds, recorded with EEG from healthy adult subjects in a "passive" listening experiment (Johnson et al., 2003). In this experiment participants were instructed to attend to an engaging video viewed with the soundtrack silenced while they ignored experimental stimuli presented via insert earphones. Prior to the EEG recording session all subjects underwent a psychophysical screening procedure to ensure they could detect dichotic pitch (hereafter, "DP").

The left column of Fig. 3 shows ERPs averaged over 400 trials of each stimulus type and grand averaged over a group of 13 subjects, and recorded from electrodes placed at a frontal

midline location on the head and at two lateral positions about 4 cm to the left and right of the midline. ERPs are plotted as voltage time series over a time base of -100 ms before stimulus onset to 500 ms after stimulus onset. Voltages are plotted with negative up (a convention used in many EEG labs), and ERPs evoked by DP stimuli are overlayed on top of ERPs to control stimuli.

For both types of stimuli ERPs are characterized by positive-negative-positive sequence of responses labelled P1, N1 and P2, and with peak latencies of 76 ms, 108 ms and 172 ms respectively, typical of late cortical responses to a variety of acoustic stimuli with abrupt onsets. The DP and control ERPs begin to diverge just prior to the peak of the P2 component at a latency of about 150 ms, with the DP waveform becoming more negative than the control ERP. The differences between DP and control responses are best visualized in the subtraction waveforms in the centre column of Fig. 3, showing that the amplitude difference is maximal by about 210 ms with large differences persisting until about 280 ms, after which amplitude differences decline sharply and the two ERPs show similar amplitudes again by 380 ms latency. In this early study we referred to the difference wave simply as a "late negativity" or LN component.

As can be surmised from the fairly similar responses obtained at electrode sites as much as 8 cm apart, the ERPs have an extensive spatial distribution. In the right hand column of Fig. 3 the amplitude distribution of ERPs is shown as isovoltage contours on a schematic top view of the head. In the head schematic the dots represent electrode positions and the topographic maps are based on a fairly dense spatial sampling of 128 recording electrodes. The maps of the LN component (bottom row of contour maps) shows that this component has a broad distribution centred over the frontal midline of the head. This reinforces a point made in the preceding section about the relatively coarse spatial resolution of the EEG. However the EEG does quite a good job of localizing neural processing of DP in time: the difference waveforms demonstrate a robust phase of neural processing -- specific to DP -- that occurs some 150-350 ms after stimulus onset. As the two stimuli employed in this study were discriminable solely by a dichotic delay (in the DP stimulus) but were otherwise acoustically identical, we can be confident that the late cortical LN wave reflects neural processing that is dependent on binaural fusion within the brain. These results confirm that the late cortical ERPs are highly sensitive to the binaural processes underlying the perception of DP, and suggest that these may be a useful electrophysiological tool for assessing the binaural processing capabilities of the central auditory system.

4.2 Active listening conditions

While the LN response to DP is clearly based on binaural processing, it closely resembles an ERP response associated with a perceptual structuring based on a monaural cue – the inharmonicity of one component of a complex sound composed of multiple harmonics. Alain et al. (2002) measured ERPs from subjects presented with sounds containing tuned or mistuned harmonics. In two different listening conditions the subjects either actively attended to the sounds and indicated their perceptions (a single sound versus two sounds) with a button press, or ignored the acoustic stimuli while watching a silent movie. The perception of a mistuned harmonic as a separate sound was associated with a negative wave that has a peak latency of about 160 ms, which these authors termed the "object-related negativity" (ORN). The ORN was elicited in both active and passive listening conditions, while a later P400 wave was elicited by the mistuned harmonic stimuli only when subjects
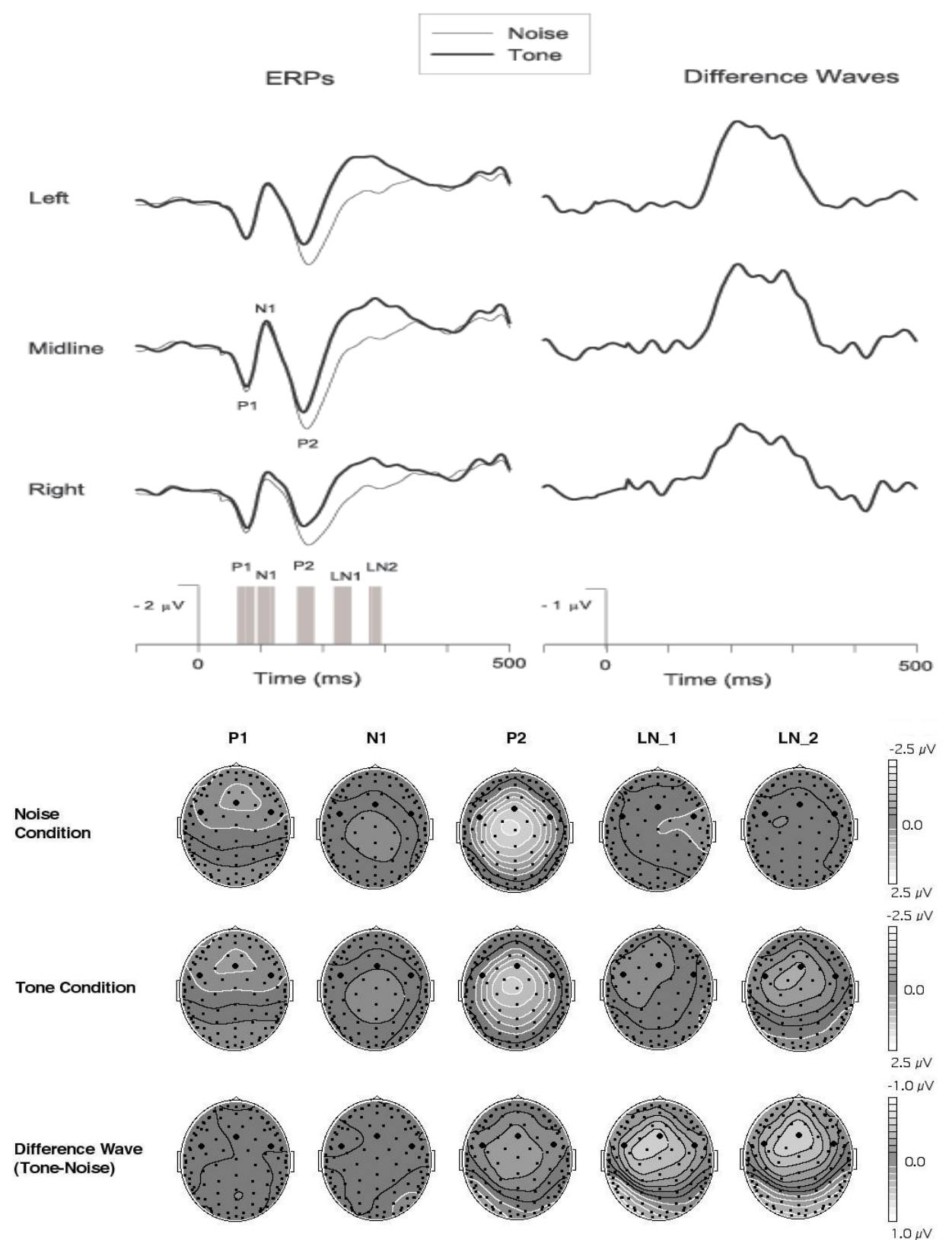
Fig. 3. Grand averaged ERPs in passive listening condition. "Tone" refers to dichotic pitch stimulus, "Noise" refers to control stimulus. Large dots indicate electrode positions for waveforms shown at top. Adapted from Johnson et al. (2003) with permission.

actively attended to these stimuli. The authors concluded that the two sequential components support a two-stage model of auditory scene analysis. In this model, complex sounds are first parsed based on an automatic process, reflected by the ORN and involving a detection of a mismatch between the harmonic template extracted from the incoming stimulus and the harmonic frequency expected based upon the fundamental of the complex sound. The second stage, indexed by the P400, is invoked only when subjects are actively attending to the stimuli and seems to reflect controlled processes responsible for identification of stimuli and selection of behavioural responses.

While the acoustic stimuli are radically different, the overall perceptual problem posed by the sounds in the DP paradigm used by us and the mistuned harmonic paradigm used by Alain et al. (2002) is the same: that of breaking a complex sound wave into components that correspond to two concurrent perceptual objects. Indeed, we found that under active listening conditions the mechanisms described in Alain et al.'s (2002) two-stage model are also deployed for the perception of DP.

Fig. 4 shows DP and control ERPs recorded under two passive listening conditions: one in which DP and control stimuli were randomly interleaved (P-R) and one in which they were presented in uniform blocks (P-B); and an "active" listening condition (A-R) in which listeners were required to actively attend to stimuli on each trial and to indicate with a button press whether a (randomly interleaved) DP or control stimulus had been presented. The ORN was elicited by DP stimuli in all three listening conditions, while a P400 response was elicited only in the active condition.

Fig. 5 is an instructive summary of the ERP data from this experiment because it clearly shows that successive segments of the ERPs show quite distinctive behaviours as a function of stimulus and listening conditions. The N1 component is modulated by attention but not stimulus type, showing a generalized increase in amplitude (i.e. greater negativity) when actively attending. The ORN component shows a generalized attention effect and a main effect of stimulus type (i.e. it is more negative when processing DP stimuli). Finally the P400 is manifest as an interaction between listening condition and stimulus type, because it is elicited only by DP stimuli and only when listeners are required to actively discriminate stimuli. These distinctive functional profiles clearly indicate that the three components reflect distinct stages of auditory processing; and conversely, that ERPs are capable of localizing different stages of processing in time.

These results show that the perception of DP is associated with two cortical processing stages – indexed by the ORN and the P400 – that are functionally comparable to those elicited by the mistuned harmonic stimuli used by Alain et al. (2002). Since the physical compositions and critical cues of these two classes of sounds are radically dissimilar, it is reasonable to conclude that these processing events are more related to the overall perceptual problem posed by both sounds: that of partitioning a complex acoustic input stream into two distinct perceptual objects. Alain et al. (2002) proposed that the ORN component indexes a transient automatic mismatch process between the harmonic template extracted from the incoming stimulus and the harmonic frequency expected based upon the fundamental of the complex sound. In the case of dichotic pitch, however, a neural mismatch response must be generated on the basis of location, since the noise processes by definition have no harmonic structure to match to. This suggests that the ORN reflects the activity of fairly general mechanisms of auditory stream segregation that can broadly utilize a range of cues to parse simultaneous acoustic events.
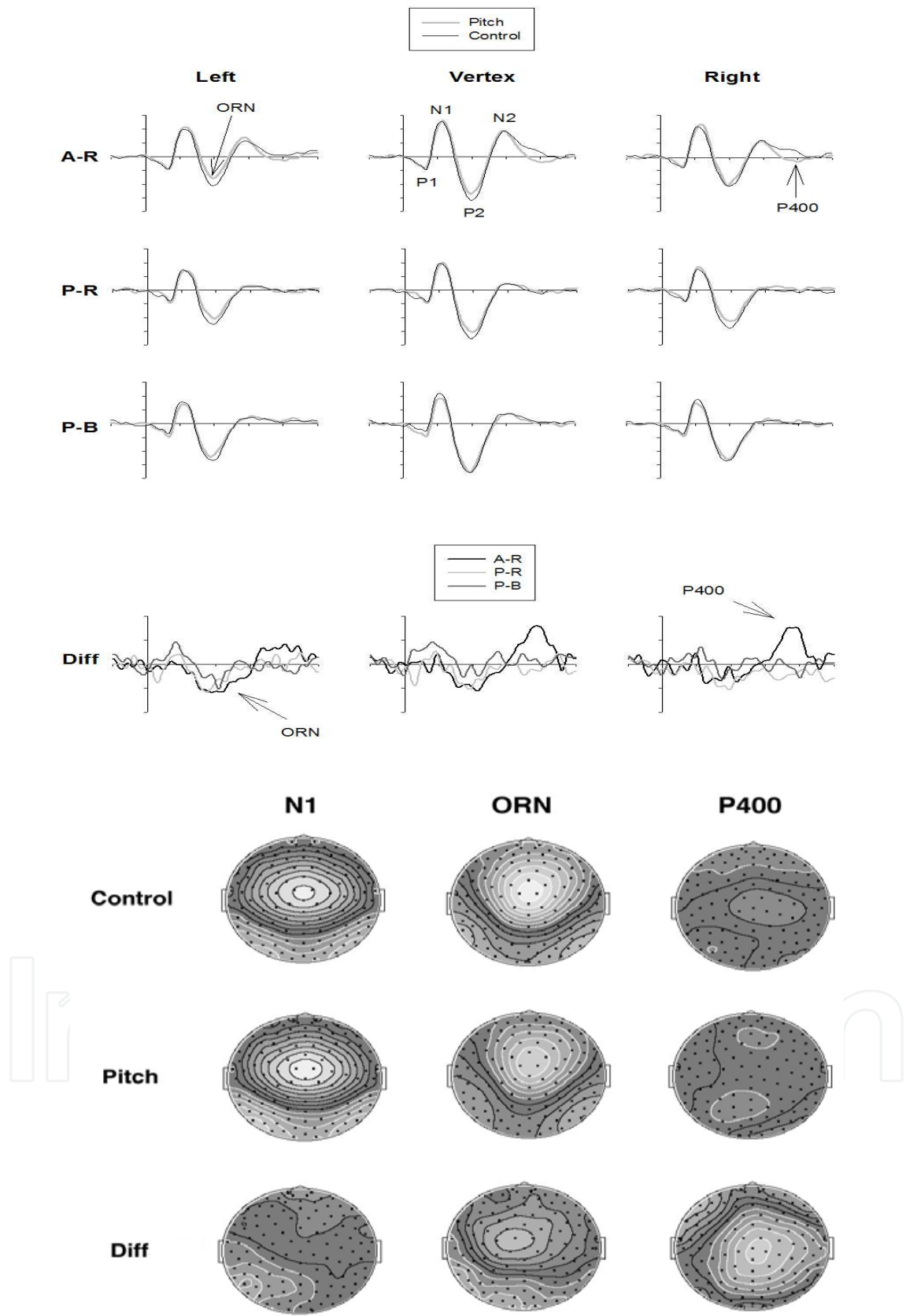
Fig. 4. Grand averaged ERPs under three listening conditions. Adapted from Hautus and Johnson (2005) with permission.
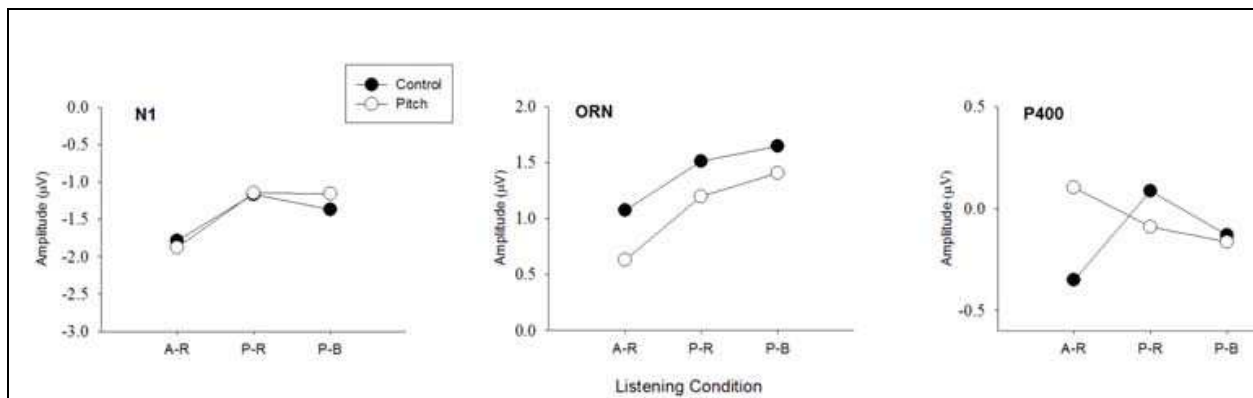
Fig. 5. Functional profiles of ERP components. Adapted from Hautus and Johnson (2005) with permission.

4.3 Cortical generators of the ORN

Consistent with what one would expect of mechanisms that play such a basic role in auditory perception, the results discussed in the preceding section, and those of previous researchers (Alain & Izenberg, 2003; Alain et al., 2002) suggest that the ORN indexes relatively automatic processes that function independently of attention. Such automaticity points to relatively low levels of the auditory system, and indeed physical modelling of the generators of the auditory P1-N1-P2 complex indicates that these waves are best modelled by current sources in or near the primary auditory cortices (Picton et al., 1999; Scherg et al., 1986). More evidence for the involvement of primary auditory cortex in stream segregation comes from a study by Dyson and Alain (Dyson & Alain, 2004), who reported that the first reliable effect of mistuning harmonic stimuli was a modulation of the amplitude of the Pa peak of the middle latency auditory evoked response. There is good evidence from convergent sources that the Pa wave is generated in primary auditory cortex (Liegois-Chauvel et al., 1991; McGee et al., 1992; Pantev et al., 1995; Scherg & Von Cramon, 1986), and the early latency of this response (about 32 ms) is consistent with the preattentive processing of acoustic cues to auditory stream segregation.

As noted previously, MEG is relatively most sensitive to cortical generators that are oriented tangentially to the surface of the head; consequently the auditory cortices of the supratemporal plane are well-visualized with MEG. Fig. 6A (left column) shows the distribution of AEF amplitude measured with MEG sensors (indicated as dots) and projected onto the surface of the head. In comparison to the EEG topographic maps described in previous sections, the MEG maps show some obvious differences. First, while the AEP fields were maximal over the superior surface of the head, the AEF fields are maximal over temporal regions. Second, the AEF fields have a much more focal distribution than the AEP fields. The first point of difference can be reconciled by appreciating that both electrical and magnetic fields are generated by populations of tangentially-oriented pyramidal cells in the supratemporal plane. A population of synchronously activated neurons can be well approximated by an "equivalent current dipole" (ECD), represented as the coloured ball and sticks in Fig. 6B (left column). The ball of the ECD represents the "centre of gravity" of a region of activated cortex, while the stick points in the direction of positive current flow. Accordingly, this ECD will generate an electric field with a negative polarity on the superior surface of the head and a positive polarity below the supratemporal plane. The magnetic field generated by the same ECD will follow the "right hand rule" of

electromagnetism: If current is flowing in the direction of your extended right thumb, then magnetic flux will circle the ECD in the direction of the right fingers. For the ECD of Fig. 6B (left column) this will result in the magnetic field pattern of Fig. 6A (left column): on the left hemisphere positive flux emerging from the posterior temporal region (shown as red) and negative flux re-entering the head in the anterior temporal region (shown as blue). Following this logic it is clear why the opposite flux pattern is seen over the right hemisphere. The second point of difference, the more focal distribution of the AEFs, is due the fact that magnetic fields are not subject to the smearing effects of conductive inhomogeneities: the brain, cerebrospinal fluid, skull and scalp are transparent to magnetic fields.

These considerations indicate that AEFs can be quite effectively modelled with very simple models of both the brain sources and the physical characteristics of the brain and head. With such computational models hundreds of channels of surface-recorded MEG data can be rendered into a simple configuration of brain sources that nonetheless capture quite significant dynamics of brain function. Fig. 6A (right column) shows grand-averaged source waveforms from the bilateral dipole model of AEFs elicited by broadband noise. The source waveforms show that the responses of the left and right hemispheres are distinctively different: the right hemisphere has a much reduced middle latency M50 (equivalent to the P1 AEP) response but a larger amplitude M100 (equivalent to the N1 AEP) with an earlier
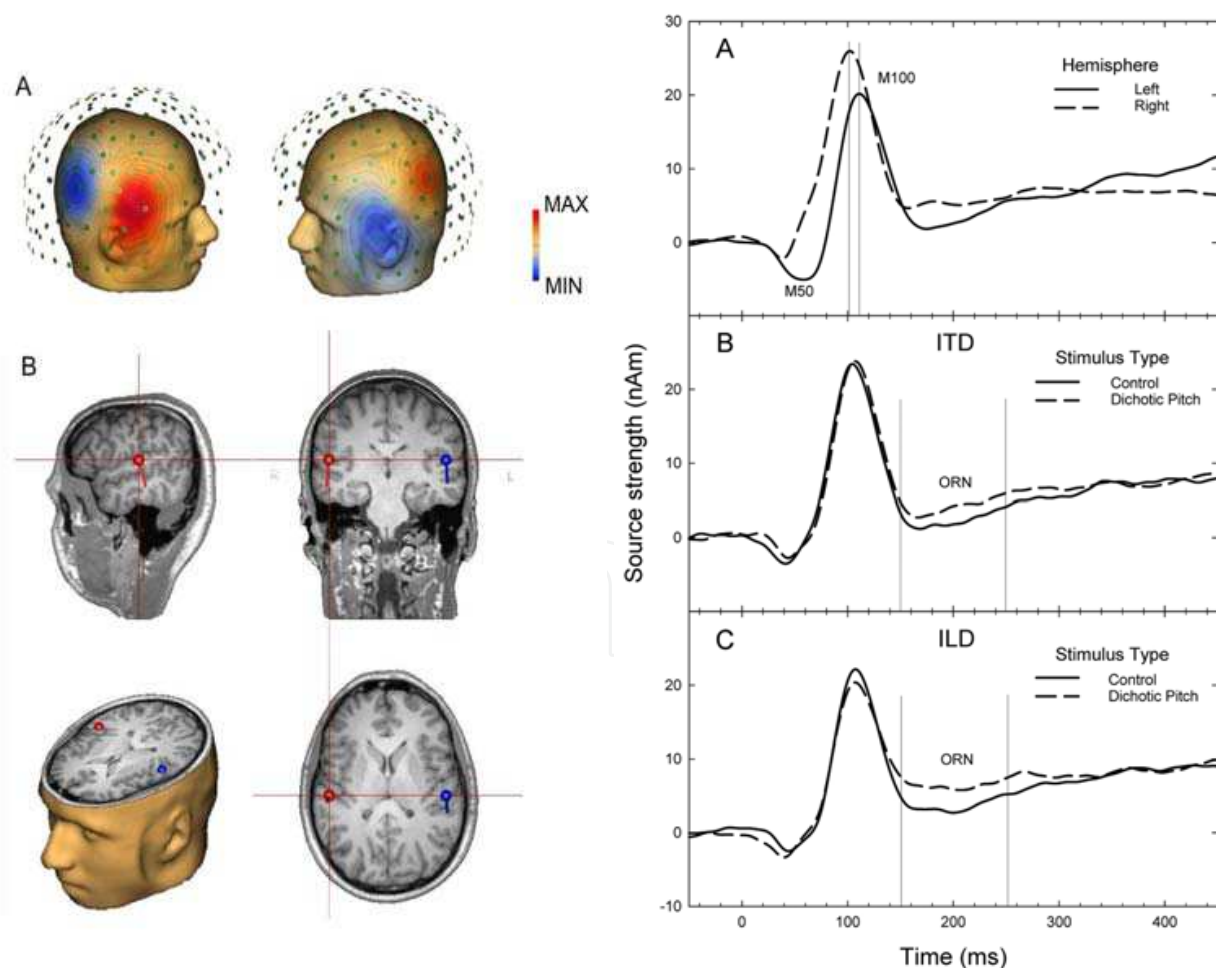

Fig. 6. ORN measured with MEG. Adapted from Johnson and Hautus (2010) with permission.

peak latency. Fig. 6B and C (right column) show the MEG source waveform version of the ORN elicited by interaural disparities in the left hemisphere. The timing of the MEG ORN coincides with the timing of the EEG ORN described previously.

5. Differential processing of interaural timing and level differences

If the ORN reflects the activity of fairly general mechanisms of auditory stream segregation that can draw on a range of cues to parse simultaneous acoustic events, we reasoned that it might be useful in helping to address an unresolved issue in binaural hearing concerning the processing of ITD and ILD cues in the auditory cortex (Johnson and Hautus, 2010). It has long been appreciated that ITDs are the dominant cues for low frequency sounds, while ILDs dominate for high frequency sounds, the so-called duplex theory of audition (Rayleigh, 1907). Since the formulation of the duplex theory, researchers have suggested the existence of separate neural processing channels for ITDs and ILDs. Indeed, there are several lines of physiological evidence for independent processing mechanisms from unit recordings in the auditory brainstem of animals (Phillips & Brugge, 1985; Smith et al., 1993; Yin & Kuwada, 1984) and also from surface recordings of auditory brainstem responses in humans (Pratt et al., 1997).

It would seem a logical requirement for the auditory system to eventually pool the spatial information extracted from disparate cues into a common code that can be used to solve the broad perceptual problems posed for auditory scene analysis. However it remains unclear when, or even if, information from ITDs and ILDs may be combined into a common code for spatial lateralization (Schroger, 1996; Ungan et al., 1997). On the one hand, psychophysical studies have shown that lateralization to one ear induced by one cue can be precisely counterbalanced by the complementary cue leading at the other ear, according to a systematic "trading ratio" (Hafter & Jeffress, 1968; Harris, 1960). This suggests that information from the two cues is eventually merged at some stage of the central nervous system. On the other hand, the trade-off between ITD and ILD does not seem to be complete: Listeners report that they experience distinctly different types of sound lateralization "images" for the two types of cues (Hafter & Carrier, 1972).

This suggests that segregation may be maintained, at least to some degree, to the level of conscious perception. Indeed, evidence from the cat (Phillips, 1993) and from brain-damaged human patients (Yamada et al., 1996) indicates that separate representations of ITDs and ILDs exist at the level of the auditory cortex. Further, an EEG study has reported different surface topographies for the circa 100 ms latency N1 event-related potential (ERP) component elicited by ITDs and ILDs, indicating spatially separated neural representations at a relatively late stage of processing in the cerebral cortex (Ungan et al., 2001). MEG recordings show that the two cues have independent effects on the amplitude of the M100, the magnetic counterpart of the N1 (Palomäki et al., 2005). A similar finding for the mismatch negativity (MMN) component of the ERP (Schroger, 1996) suggests at least partially independent cortical processing of timing and level cues up to 200 ms in latency.

Taken together, these studies provide good evidence for segregation of ITDs and ILDs to quite late stages of auditory processing. However, they shed little light on when (if ever) ITDs and ILDs might be incorporated into a common signal that could mediate perceptually relevant phenomena. We measured auditory brain function with MEG to determine if the ORN might represent a stage when information from the two cues is merged into a common code for auditory scene analysis (Johnson & Hautus, 2010).

5.1 Segregated processing of ITD and ILD cues

The experiment was a 2 x 2 x 2 design with variables location cue type (ITD or ILD), stimulus type (control or dichotic pitch), leading ear for background noise (noise perceived on left or noise perceived on right). Each location cue type could result in four possible percepts: a single noise on the left or right (control stimuli), or concurrent background noises and dichotic pitches (dichotic pitch stimuli), with the background noise perceived on the right or left and the dichotic pitch perceived on the opposite side.

The results showed that ITD and ILD cues elicit distinctive hemispheric patterns of activation during the M100 time window. Fig. 7 (top row) shows that left lateralized control sounds elicited larger amplitude M100s in the contralateral hemisphere, while right-lateralized sounds elicited similar amplitude responses in both hemispheres. In contrast, both left and right lateralized sounds for ILD cues elicited similar patterns of activation (larger amplitude on the right). Interestingly, the same stimulus x hemisphere interaction was obtained for ITDs only when the left-lateralized DP stimuli were compared to the right lateralized control stimuli. This is a striking result because the bulk of the stimulus energy of the left lateralized DP stimulus (the noise) was perceived on the right. We interpreted this pattern in terms of a stronger right-hemisphere bias for spatial information pitted against a left hemisphere bias for timing information: the left hemisphere holds its own if sounds originate solely from the opposite hemispace: the right hemisphere wins the tug of war if any sounds are present in the left hemispace (Johnson and Hautus, 2010).

While a unilateral ITD cue can engage the left hemisphere if presented in the right hemispace, no such effect was obtained for ILD cues (Figure 5, right). This finding reinforces the interpretation that it is timing information, contained in the ITD representation but not the ILD representation, that is crucial for engaging the left hemisphere. The greater capacity of ITD cues to activate the left hemisphere cannot be attributed to a greater salience of this cue, since our behavioural data showed that the ILD cues were in fact more detectable, despite our efforts at loudness and laterality matching. These results therefore support the hypothesis that, during the time window of the M100 response, ITD and ILD cues are processed in at least partially distinct channels of the auditory cortex (Wright & Fitzgerald, 2001), aimed at differentially elaborating the specific kinds of spatial information represented in each channel.

Our results are supported by an EEG topographic mapping study by Tardiff et al. (Tardif et al., 2006). These authors also found that ITD/ILD differences were primarily in terms of hemispheric lateralization, although they reported somewhat different hemispheric patterns than ours, with bilateral responses to ILD cues and left-lateralized responses to ITD cues. The differences between studies may be attributable to differences in EEG and MEG recording methodologies, but in any case both studies support a greater involvement of the left hemisphere in processing ITD cues than for ILD cues. An involvement of the left hemisphere in sound localization is supported by a recent study of patients with right or left hemisphere brain damage (Spierer et al., 2009). While these authors found a prominent right hemispheric dominance for auditory localization, significant spatial deficits were also observed in patients with strictly left hemisphere damage. These authors concluded, as do we, that right hemispheric dominance is better conceptualized as an asymmetrical activation of both hemispheres, rather than an absolute role for the right hemisphere.

5.2 Pooling of spatial information for auditory scene analysis

For both ITD and ILD location cues during a time window of 150-250 ms, sources in both hemispheres showed a prominent increase in amplitude (ORN) for sounds containing a
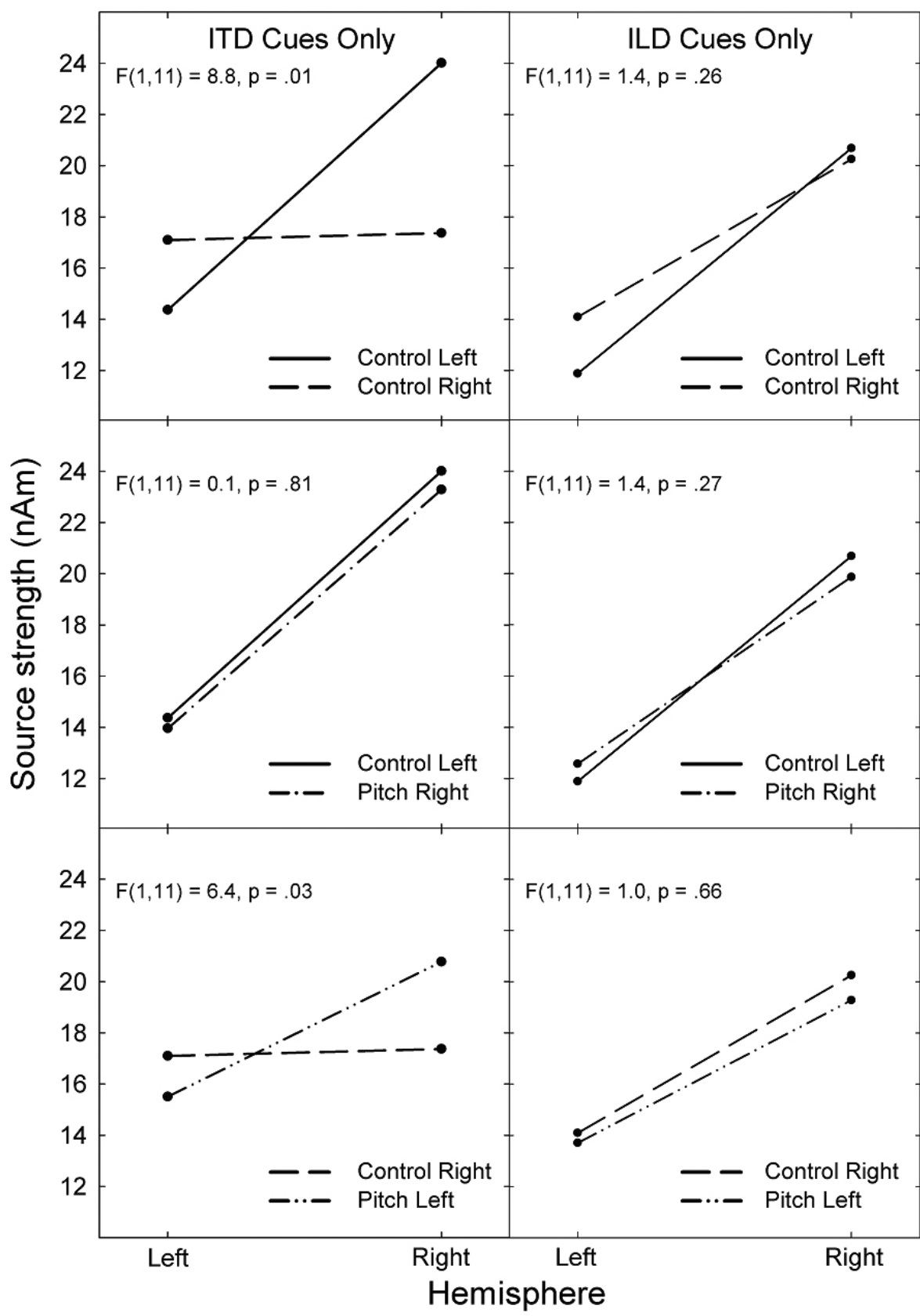
Fig. 7. Distinctive hemispheric activation patterns for ITD and ILD cues. Adapted from
Johnson and Hautus (2010), with permission.

dichotic pitch, in comparison to sounds that contained no binaurally-embedded pitch (Fig. 6 right panel, B & C). The finding that the ORN is elicited to ILD cues to sound location further strengthens the contention that ORN mechanisms are able to broadly draw on information derived from a variety of cues: that is, the generation of the ORN is not constrained to the operation of highly specific processes such as a spectral template matching mechanism. The mean time window of the ORN effect was essentially identical for both ITD and ILD cues. In contrast to the M100 component, the ORN component was not itself modulated in amplitude by location cues, nor was its lateralization influenced by the type of location cue. This negative result stands in contrast to previous findings for the M100 and later components (Schroger, 1996; Tardif et al., 2006). Therefore, the ORN can potentially place an important and novel temporal boundary on the extent to which ITD and ILD cues, or the information derived from them, remain segregated in separate channels of the cerebral cortex, since information that can be used to parse a pitch from a background noise must have been extracted from both cues prior to the generation of the ORN. Further investigations are required to address these important theoretical issues. We note also that the ORN component was superimposed upon brain activity in the same time window that was strongly modulated by cue type (i.e. there was a main effect of cue type but no interaction with stimulus type). It seems that the brain may continue to process information in independent streams even after spatial information has been extracted to support auditory scene segregation.

In summary, these results show a strong modulation of interhemispheric activity by ITD cues, but only when these cues are presented unilaterally from the right hemispace. These data support the interpretation of a relatively strong right hemisphere bias for spatial information in conjunction pitted against a relatively weaker left hemisphere preference for timing information. The hemispheric biases are large in comparison to the modest contralateral bias exhibited at the population level in primate auditory cortex (Werner-Reiss & Groh, 2008). In contrast, ILD cues lack the capacity of ITDs to engage the left hemisphere, presumably because their cortical representations lack the timing information that is preferentially processed in that hemisphere. Finally, spatial information that is common to both ITD and ILD cues seems to be extracted prior to the ORN time-window for use by the cerebral mechanisms of auditory scene segregation.

6. Sequential processing of ITDs for sound source segregation and spatial localization

In everyday life, ITDs play a fundamental role in two basic aspects of auditory perception. First, they serve as the primary cues for localization of low frequency sounds in azimuthal space (Blauert, 1997). Second, ITDs are one of a set of Gestalt-style grouping cues employed by the auditory system to sort and parse sound mixtures produced by concurrently active sound sources (Drennan et al., 2003) and therefore play a role in a structuring of perception referred to by Bregman (1990) as "auditory scene analysis." This analysis allows us (for example) to attend to the voice of one speaker among the babble of many others at a cocktail party.

Considered individually these two perceptual roles for ITDs have been well studied. However, the relationship between these fundamental processes has received scant attention from auditory scientists. Intuitively one may suppose that both perceptual results may be achieved by the same stage of processing, since identifying the locations of two

temporally concurrent (but spatially disparate) sounds would seem to automatically result in their spatial separation. However there are both empirical and theoretical reasons to believe that auditory scene analysis and spatial localization are achieved in distinct steps of auditory processing. Neuropsychological evidence comes from a recent report of a patient with an ischemic lesion of right temporo-parieto-frontal areas (Thiran & Clarke, 2003). This patient exhibited a severe impairment in localizing sounds in everyday life and was entirely unable to use ITD cues to localize sounds in a spatial release from masking task and several diotic tasks requiring spatial localization from ITDs. Despite her profound 'spatial deafness' she was nonetheless able to use ITDs to segregate concurrent sound sources.

Recent theoretical views of auditory perception also point to unique neural mechanisms for auditory segregation and localization. For example, Griffiths and Warren (2002) have suggested that the segregational processes of auditory scene analysis are accomplished by a 'computational hub' of the auditory system, which they suggest is located in the planum temporale. These authors proposed that the neurons of this region employ an algorithm akin to independent component analysis (Bell & Sejnowski, 1995) to separate the components of complex acoustic signals. On this view, the parsed outputs from this stage of processing are subsequently routed via anatomically and functionally separate pathways to higher order destinations for segregated processing of auditory attributes including object identity and spatial position (Rauschecker & Tian, 2000). In line with the neuropsychological evidence described above, this model suggests a sequence of processing in which binaural information in complex sounds is initially employed for sound segregation, followed by an analysis which results in the perceptual elaboration of sound location.

We attempted to address this issue by measuring ERPs in experimental conditions that required listeners to extract dichotically-embedded pitches lateralized to the right or left of auditory space. Following the sequential model of ITD processing outlined previously we predicted that the ORN, as a marker of an early step in auditory scene analysis, should be relatively unaffected by variations of location. On the other hand we expected location-specific processing to be manifest in relatively later components of the auditory ERP. A second objective was to determine if the neural processing of ITDs is influenced by the nature of the listening task. To test this, we compared DP ERPs obtained in a task that specifically required listeners to locate the stimuli in auditory space (localization task), to those from a task that could be performed without actually computing locations (detection task). We predicted that the relatively low-level and automatic processing indexed by the ORN should be relatively unaffected by task demands. However the relatively higher level and more controlled levels of auditory scene analysis, suggested to be indexed by the P400 component, could be expected to be more heavily influenced by the goals of the behavioural task.

6.1 The ORN and perceptual segregation

The experiment consisted of three conditions based on two different yes-no tasks: detection and localization.

Detection: Participants listened to a random sequence of control stimuli and DP stimuli. The a priori probability was 0.5. Within a block of presentations, the DPs were located consistently to one side of auditory space – either the right or the left. Thus, the detection task consisted of two conditions (detect-right and detect-left). The participant indicated on a button-box whether the stimulus was a dichotic pitch or a control stimulus.

<u>Localization</u>: (We note that spatial judgements of sounds perceived intracranially are properly referred to as "lateralization" judgements. We prefer to use localization in this context to avoid confusion when referring to hemispheric lateralization of neural activity.) Participants listened to a sequence of DP stimuli (no control stimuli) with the pitches randomly located with equal probability on either the left or right side of auditory space. The participant indicated on a button-box whether the pitch stimulus presented was located to their right or left. Across the two conditions there were three classes of stimuli, each with two types: detection DPs (left and right) and detection controls (for left and right pitches); localization DPs (left and right).

Fig. 8 shows that a robust ORN response was elicited in all experimental conditions. The finding that the ORN was not modulated by location nor by task supports the interpretation that this is a relatively automatic response to DPs.
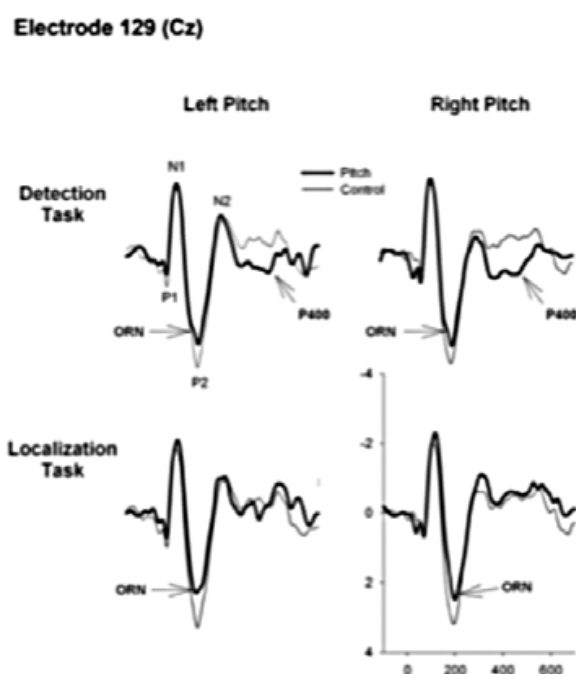

Fig. 8. Grand averaged ERPs for detection and localization tasks. Adapted from Johnson et al. (2007) with permission.

6.2 Location specific processing and the N2 component

Fig. 9 shows an amplitude modulation at lateral electrode sites which we termed "N2" because it overlaps during the time window of the vertex (Cz) N2 peak labelled in Fig. 8. We note that our N2 modulation is sustained several hundred milliseconds beyond the N2 peak (see Fig. 5) and has maximal amplitude distribution at lateral temporal sites contralateral to perceived pitch location. The most salient characteristic of the lateral temporal N2 component is its sensitivity to the spatial attributes of dichotic pitch, suggesting that this component reflects a location-specific phase of neural processing. We found a similar N2 contralateralization for lateralized pitches whether these were presented in the detection task, or if listeners were required to localize the pitches, suggesting that the spatial attributes of the stimuli were processed to some extent regardless of whether accurate task performance actually required a computation of spatial position.
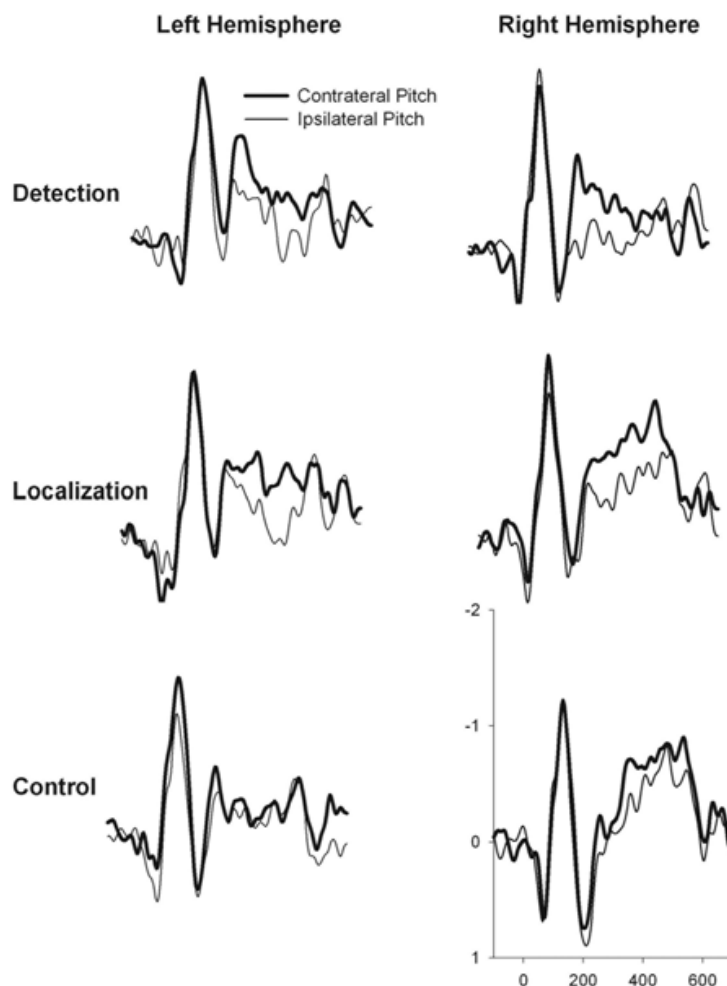
Fig. 9. Grand averaged ERPs recorded at lateral temporal electrodes on the left and right hemispheres. From Johnson et al. (2007) with permission.

6.3 The P400 and perceptual context

The ORN and N2 components were robustly evoked by ITDs in both tasks. In contrast, the P400 - a prominent feature of the detection ERPs – was entirely absent from the localization ERPs (Fig. 8). Previous studies have shown that, as with the ORN, the P400 component is tightly linked to concurrent sound segregation, but unlike the ORN is elicited only when listeners are required to actively attend and generate behavioural responses to sounds (Alain et al., 2002; Hautus & Johnson, 2005). As such, it seems clear that it indexes a controlled, rather than automatic, level of processing associated with the identification of concurrent sound sources and selection of a behavioural response (Alain et al., 2002).

The lack of a P400 component in the localization task shows that this component is not an obligatory consequence of auditory processes involved in concurrent sound segregation. Since both tasks require perceptual segregation of, and active responses to, concurrent sound sources, why do physically identical dichotic pitches elicit a P400 in the detection tasks but not in the localization task? One possibility is suggested by the fact that in the detection task the dichotic pitch stimuli were interleaved with control stimuli that did not contain any binaural cues, while in the location task all stimuli contained binaural cues. This

suggests that the P400 may be strongly influenced by the perceptual context in which the sounds are presented.

To test this possibility we performed a second order analysis in which we re-averaged ERPs based on the type of sound that immediately preceded a given sound. The results supported the conjecture that the P400 to DP is highly sensitive to the perceptual context in which the sounds are presented. Maximal P400 amplitudes were elicited when the dichotic pitch was presented after a control stimulus, but were much reduced in amplitude when the dichotic DP followed another DP. No P400 was elicited when the sounds in a sequence were all DPs (the localize condition). This pattern of results suggests that the P400 component may index a change in the number of perceptual objects in the acoustic environment. This appears to be a unidirectional phenomenon, since no P400 is obtained when the change is from a DP to a control stimulus. Thus, the P400 may be specifically sensitive to a change from one object (the centrally located noise) to two objects (the noise plus a lateralized pitch).

6.4 ERP components as indices of sequential processing

The sequence of processing events revealed by the results of this ERP study are summarized by the functional profiles of Fig. 10. The earliest, N1 component exhibits no modulation by experimental variables (cf. Fig 5, where attentional factors modulated the N1). The ORN was elicited by ITDs regardless of location or task. In contrast, the later N2 response (250-350 ms) was strongly contralateralized to the perceived location of a DP. Finally, DP stimuli in the detection task elicited a P400 at a latency of 400-500 ms, but this response was entirely absent from ERPs elicited by identical stimuli in the localization task. The sequence of cortical processing events shown here supports the prediction that operations associated with spatial localization of sounds are functionally distinct from, and are preceded by, operations associated with the perceptual segregation of those sounds. This functional-temporal dissociation is consistent with a model of auditory system functioning that suggests that the components of auditory information are initially separated on the basis of Gestalt-style auditory grouping cues (Bregman, 1990) including location (specified here by ITDs) and inharmonicity. This initial structuring of the acoustic waveform is considered a crucial computational step in auditory perception (Griffiths & Warren, 2002), which feeds parsed outputs to higher cortical areas for perceptual elaboration of key sound features including spatial position and object identity (Rauschecker & Tian, 2000).

7. Conclusion

Nature has elected to provide most of its creatures with a pair of ears rather than economizing with a single hearing sensor (Schnupp and Carr, 2009). Binaural hearing -- like stereo vision -- analyses informational disparities between two sensors to build spatial representations of objects and the underlying neural computations in the two modalities may employ common algorithms (Wagner, 2004). Both stereo vision and binaural hearing serve to enrich our perceptions of the world. In the auditory modality these improvements enhance our capacity to perform a fundamental structuring of perception referred to as 'auditory scene analysis' (Bregman, 1990), involving a parsing of the acoustic input stream into behaviourally-relevant representations. In a world that contains a cacophony of sounds, binaural hearing is employed to extract out a single sound source, determine its location,
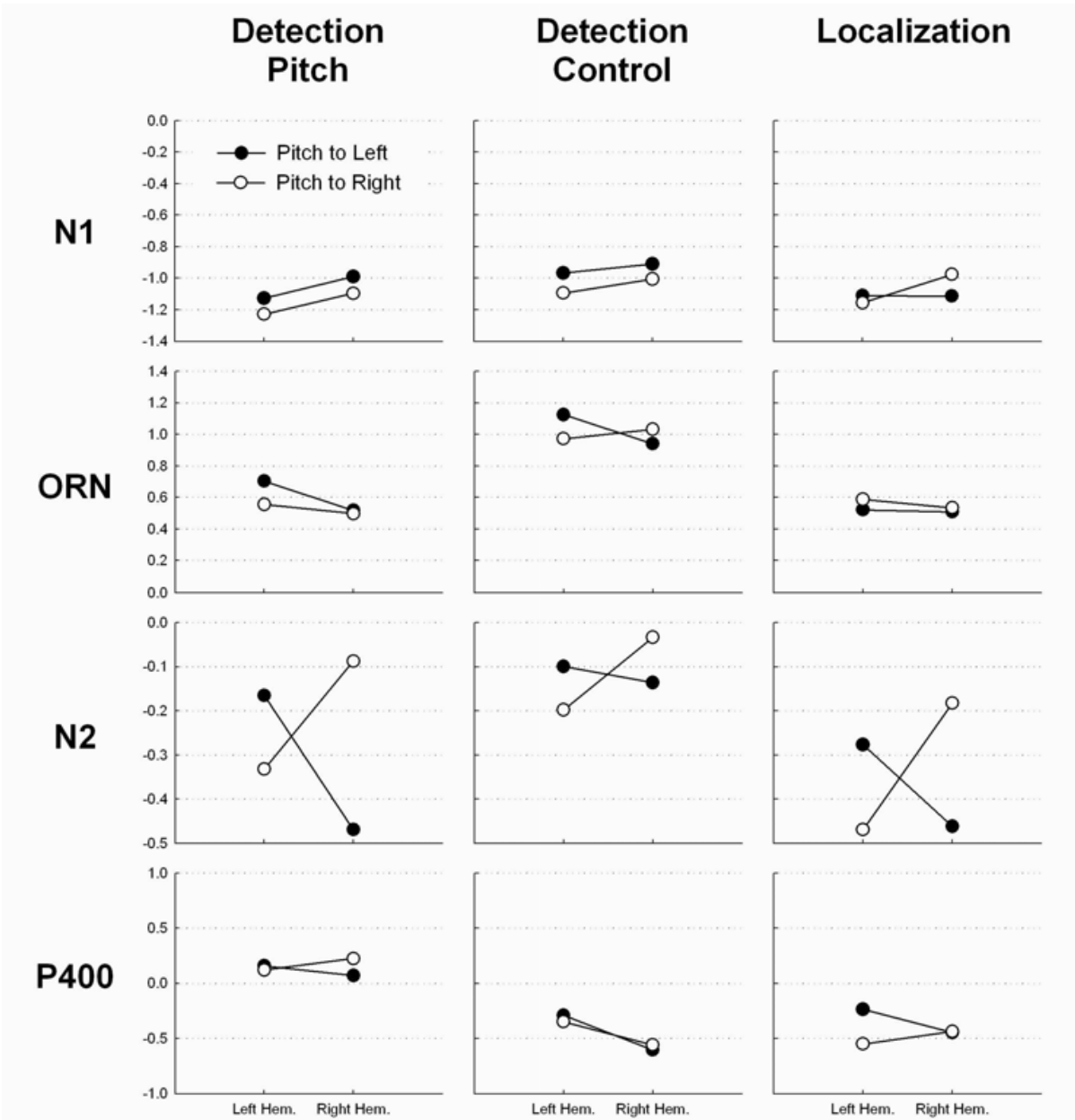
Fig. 10. Functional profiles for ERP components, plotted for left and right hemispheres. From Johnson et al. (2007) with permission.

and assign it meaning (Erikson and McKinley, 1997, p. 722). In the last several years our group has studied how binaural information is processed to these ends in the human auditory cortex, using a psychophysical paradigm to elicit binaural processing and using electroencephalography (EEG) and magnetoencephalography (MEG) to measure cortical function.

The dichotic pitch paradigm has a number of features that have made it useful for probing the workings of the binaural auditory system. Interaural disparities are first computed at the level of the brainstem so the perception of DP can be confidently attributed to central rather than peripheral processes. Further, the overall perceptual problem posed by DP –that of segregating concurrent sound objects – is of considerable interest to those interested in how, and by what mechanisms, the brain is able to accomplish this important structuring of

perception. The experiments reviewed here show that EEG and MEG responses to DP consist of a sequence of auditory cortical responses that provide important markers of a number of functionally distinct stages of auditory scene analysis in the human brain.: (1) The M100 ERF seems to reflect the operation of right-hemispheric mechanisms for analysis of spatial information pitted against left hemisphere mechanisms for analysis of timing information; (2) The ORN ERP and ERF reflect the operation of fairly automatic and generalized brain mechanisms for auditory scene segregation. The ORN mechanisms can broadly draw on information about scene analysis from a variety of acoustic cues, including inharmonicity, ITDs, and ILDs. As such, the ORN appears to represent a stage of auditory processing that draws on information extracted from disparate cues into a common code that can be used to solve the broad perceptual problems of auditory scene analysis. (3) The P400 ERP is an electrophysiological signpost of a later, more controlled stage of processing, involving identification and generation of a behavioural response. This stage is highly dependent on the task and context in which stimuli are presented. (4) The N2 ERP recorded at lateral sites over the temporal lobes is highly sensitive to the spatial attributes of dichotic pitch, suggesting that this component reflects a location-specific phase of neural processing. The N2 has not been observed in MEG responses, likely because the generators have a radial orientation that the MEG is relatively less sensitive to than EEG.

Future work can leverage these electrophysiological markers to gain clearer insights into clinical conditions in which one or more of these important central processing stages may have gone awry. For example, psychophysical studies have reported that DP detection is significantly impaired in individuals with developmental dyslexia compared to normal readers (e.g. Dougherty et al., 1998). A current study in our laboratory is measuring concurrent EEG-MEG responses to DP in dyslexic and normal reading children (Johnson et al., submitted), to determine if auditory processing deficits in reading impaired children can be localized to one or more of the processing stages delineated in studies of healthy adults.

8. Acknowledgements

9. References

Alain, C. (2007). Breaking the wave: effects of attention and learning on concurrent sound perception. *Hearing Research*, 229, 1-2., (July 2007) 225-236, 0378-5955 (Print).

Alain, C., & Izenberg, A. (2003). Effects of attentional load on auditory scene analysis. *Journal of Cognitive Neuroscience*, 15, 7, 1063-1073. 0898-929X (Print) 1530-8898 (Electronic)

Alain, C., Schuler, B. M., & McDonald, K. L. (2002). Neural activity associated with distinguishing concurrent auditory objects. *Journal of the Acoustical Society of America,* 111, 990-995, 0001-4966 (Print) 1520-8524 (Electronic).

Bell, A. J., & Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation,* 7, 6, 1129-1159, 0899-7667 (Print).

Bilsen, F. A. (1976). Pronounced binaural pitch phenomenon. *Journal of the Acoustical Society of America,* 59, 2, 467-468, 0001-4966 (Print)

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*, MIT Press, 0-262-02413-6 , Cambridge, MA.

Bregman, A. (1990). *Auditory scene analysis: The perceptual organization of sound*, MIT Press, 0-262-52195-4 , Cambridge, MA.

Cramer, E., & Huggins, W. (1958). Creation of pitch through binaural interaction. *Journal of the Acoustical Society of America,* 30, 413-417, 0001-4966 (Print) 1520-8524 (Electronic).

Dougherty, R. F., Cynader, M. S., Bjornson, B. H., Edgell, D., & Giaschi, D. E. (1998). Dichotic pitch: a new stimulus distinguishes normal and dyslexic auditory function. *Neuroreport,* 9(13), 3001-3005, 0959-4965 (Print) 1473-558X (Electronic).

Drennan, W. R., Gatehouse, S., & Lever, C. (2003). Perceptual segregation of competing speech sounds: the role of spatial location. *Journal of the Acoustical Society of America,* 114, 2178-2189, 0001-4966 (Print).

Dyson, B. J. & Alain, C. (2004). Representation of concurrent acoustic objects in primary auditory cortex. Journal of the Acoustical Society of America, 115, 280-288, 0001-4966 (Print).

Erikson, M., & McKinley, R. (1997). The intelligibility of multiple talkers separated spatially in noise. In R. Gilkey & T. Anderson (Eds.), *Binaural and Spatial Hearing in Real and Virtual Environments* (pp. 701-724). New Jersey, 13: 978-080581654, Lawrence Erlbaum.

Goldberg, J. M. & Brown, P. B. (1969). Response of binaural neurons of dog superior olivary complex to dichotic tonal stimuli: some physiological mechanisms of sound localization. *Journal of Neurophysiology,* 32, 613-636, 0022-3077 (Print) 1522-1598 (Electronic).

Griffiths, T. D., & Warren, J. D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences,* 25, 348-353, 0166-2236.

Hafter, E. R., & Carrier, S. C. (1972). Binaural interaction in low-frequency stimuli: the inability to trade time and intensity completely. *Journal of the Acoustical Society of America,* 51, 6, 1852-1862, 0001-4966 (Print).

Hafter, E. R. & Jeffress, L. A. (1968). Two-image lateralization of tones and clicks. *Journal of the Acoustical Society of America,* 44, 2, 563-569, 0001-4966 (Print).

Harris, G. (1960). Binaural interactions of impulsive stimuli and pure tones. *Journal of the Acoustical Society of America,* 32, 685-692, 0001-4966 (Print).

Hautus, M. J., & Johnson, B. W. (2005). Object-related brain potentials associated with the perceptual segregation of a dichotically embedded pitch. *Journal of the Acoustical Society of America,* 117, 275-280, 0001-4966 (Print).

Johnson, B. W., Hautus, M., & Clapp, W. C. (2003). Neural activity associated with binaural processes for the perceptual segregation of pitch. *Clinical Neurophysiology*, 114, 2245-2250, 1388-2457 (Print) 1872-8952 (Electronic).

Johnson, B. W., & Hautus, M. J. (2010). Processing of binaural spatial information in human auditory cortex: neuromagnetic responses to interaural timing and level differences. *Neuropsychologia,* 48, 2610-2619, 0028-3932 (Print) 1873-3514 (Electronic).

Johnson, B. W., Hautus, M. J., Duff, D. J., & Clapp, W. C. (2007). Sequential processing of interaural timing differences for sound source segregation and spatial localization:

evidence from event-related cortical potentials. *Psychophysiology*, 44, 541-551, 0048-5772 (Print) 1540-5958 (Electronic).

Johnson, B.W., McArthur, G., Hautus, M., Reid, M., Brock, J., Castles, A., Crain, S. (submitted). Development of lateralized auditory brain function and binaural processing in children with normal reading ability and in children with dyslexia.

Julesz, B. (1971). *Foundations of Cyclopean Perception*, University of Chicago Press, 0-262-10113-0, Chicago.

Kutas, M., Van Petten, C. & Kluender, R. (2006). Psycholinguistics Electrified II: 1994-2005. In: *Handbood of Psycholinguistics*, M. Traxler & M. Gernsbacher (Eds.) (2nd ed.), (659-724), Elsevier, 0-12-369374-8, New York.

Liegois-Chauvel, C., Musolino, A., & Chauvel, P. (1991). Localization of the primary auditory areas in man. *Brain,* 114, 139-153, 0006-8950 (Print) 1460-2156 (Electronic).

McGee, T., Kraus, N., Littman, T., & Nicol, T. (1992). Contribution of the medial geniculate body subdivision to the middle latency response. *Hearing Research,* 61, 147-152, 0378-5955 (Print).

Palomäki, K. J., Tiitinen, H., Mäkinen, V., May, P. J., & Alku, P. (2005). Spatial processing in human auditory cortex: the effects of 3D, ITD, and ILD stimulation techniques. *Cognitive Brain Research*, 24, 364-379, 0926-6410 (Print).

Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., et al. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology,* 94, 26-40, 0013-4694 (Print) 0424-8155 (Electronic).

Phillips, D., & Brugge, J. (1985). Progress in the neurobiology of sound direction. *Annual Review of Psychology,* 36, 245-274, 0066-4308 (Print) 1545-2085 (Electronic).

Phillips, D. P. (1993). Representation of acoustic events in the primary auditory cortex. *Journal of Experimental Psychology: Human Perception and Performance,* 19, 203-216, 0096-1523 (Print) 1939-1277 (Electronic).

Picton, T. W., Alain, C., Woods, D. L., John, M. S., Scherg, M., Valdes-Sosa, P., et al. (1999). Intracerebral sources of human auditory-evoked potentials. *Audiology & Neurotology,* 4, 64-79, 1420-3030 (Print).

Pratt, H., Polyakov, A., & Kontorovich, L. (1997). Evidence for separate processing in the human brainstem of interaural intensity and temporal disparities for sound lateralization. *Hearing Research*, 108, 1-8, 0378-5955 (Print) 1878-5891 (Electronic).

Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proceedings of the National Academy of Sciences of the United States of America,* 97, 11800-11806, 0027-8424 (Print).

Rayleigh, L. J. (1907). On our perception of sound direction. *Philosophical Magazine (Series 6),* 13, 74, 214-232, 1941-5982 (Print).

Scherg, M., Vajsar, J., & Picton, T. W. (1986). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience,* 1, 326-355, 0898-929X (Print) 1530-8898 (Electronic).

Scherg, M., & Von Cramon, D. (1986). Evoked dipole source potentials of the human auditory cortex. *Electroencephalography and clinical Neurophysiology,* 65, 344-360. 0013-4694 (Print).

Schnupp, J., & Carr, C. (2009). On hearing with more than one ear: lessons from evolution. *Nature Neuroscience*, 12(6), 692-697. 0022-3077.

Schroger, E. (1996). Interaural time and level differences: Integrated or separated processing? *Hearing Research,* 96, 191-198, 0378-5955 (Print) 1878-5891 (Electronic).

Smith, P. H., Joris, P. X., & Yin, T. C. (1993). Projections of physiologically characterized spherical bushy cell axons from the cochlear nucleus of the cat: evidence for delay lines to the medial superior olive. *Journal of Comparative Neurology,* 331(2), 245-260, 0021-9967 (Print).

Spierer, L., Bellmann-Thiran, A., Maeder, P., Murray, M. M., & Clarke, S. (2009). Hemispheric competence for auditory spatial representation. *Brain,* 132(Pt 7), 1953-1966. 1460-2156 (Electronic).

Tardif, E., Murray, M. M., Meylan, R., Spierer, L., & Clarke, S. (2006). The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans. *Brain Research,* 1092, 161-176. 0006-8993 (Print).

Thiran, A. B., & Clarke, S. (2003). Preserved use of spatial cues for sound segregation in a case of spatial deafness. *Neuropsychologia,* 41, 1254-1261. 0028-3932.

Ungan, P., Yagcioglu, S., & Goksoy, C. (2001). Differences between the N1 waves of the responses to interaural time and intensity disparities: scalp topography and dipole sources. *Clinical Neurophysiology,* 112, 485-498, 1388-2457 (Print) 1872-8952 (Electronic).

Ungan, P., Yagcioglu, S., & Ozmen, B. (1997). Interaural delay-dependent changes in the binaural difference potential in cat auditory brainstem response: implications about the origin of the binaural interaction component. *Hearing Research,* 106, 66-82, 0378-5955 (Print) 1878-5891 (Electronic).

Wagner, H. (2004). A comparison of neural computations underlying stereo vision and sound localization. *Journal of Physiology Paris*, 98, 135-145. 0928-4257 (Print).

Werner-Reiss, U., & Groh, J. M. (2008). A rate code for sound azimuth in monkey auditory cortex: Implications for human neuroimaging studies. *Journal of Neuroscience,* 28, 3747-3758. 0270-6474.

Wright, B. A., & Fitzgerald, M. B. (2001). Different patterns of human discrimination learning for two interaural cues to sound-source location. *Proceedings of the National Academy of Sciences USA,* 98, 12307-12312, 0027-8424 (Print).

Yamada, K., Kaga, K., Uno, A., & Shindo, M. (1996). Sound lateralization in patients with lesions including the auditory cortex: comparison of interaural time difference (ITD) discrimination and interaural intensity difference (IID) discrimination. *Hearing Research,* 101, 173-180, 0378-5955 (Print).

Yin, T., & Kuwada, S. (1984). Neuronal mechanisms of binaural interaction. In G. Edelman (Ed.), Dynamic *Aspects of Neocortical Function*. New York, 0471805599, Wiley.

Yin, T. C., & Chan, J. C. (1990). Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology,* 64, 465-488. 1522-1598 (Electronic), 0022-3077 (Print).

Yost, W. A. (1991). Thresholds for segregating a narrow-band from a broadband noise based on interaural phase and level differences. *Journal of the Acoustical Society of America,* 89, 838-844. 0001-4966 (Print).

Advances in Sound Localization

Edited by Dr. Pawel Strumillo

Sound source localization is an important research field that has attracted researchers' efforts from many technical and biomedical sciences. Sound source localization (SSL) is defined as the determination of the direction from a receiver, but also includes the distance from it. Because of the wave nature of sound propagation, phenomena such as refraction, diffraction, diffusion, reflection, reverberation and interference occur. The wide spectrum of sound frequencies that range from infrasounds through acoustic sounds to ultrasounds, also introduces difficulties, as different spectrum components have different penetration properties through the medium. Consequently, SSL is a complex computation problem and development of robust sound localization techniques calls for different approaches, including multisensor schemes, null-steering beamforming and time-difference arrival techniques. The book offers a rich source of valuable material on advances on SSL techniques and their applications that should appeal to researches representing diverse engineering and scientific disciplines.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Blake W. Johnson (2011). Processing of Binaural Information in Human Auditory Cortex, Advances in Sound Localization, Dr. Pawel Strumillo (Ed.), ISBN: 978-953-307-224-1, InTech, Available from: http://www.intechopen.com/books/advances-in-sound-localization/processing-of-binaural-information-in-human-auditory-cortex

INTECH
open science | open minds