MECHANISMS AND DYNAMICS OF THE HUMAN AUDITORY STEADY-STATE RESPONSE

By
DANIEL J. BOSNYAK, B.A. (HONS)

A Thesis
Submitted to the School of Graduate Studies
In Partial Fulfillment of the Requirements
for the Degree

Doctor of Philosophy

McMaster University

MECHANISMS & DYNAMICS OF THE AUDITORY STEADY-STATE RESPONSE

DOCTOR OF PHILOSOPHY (2003)
(Psychology)

McMaster University
Hamilton, Ontario

TITLE : Mechanisms and Dynamics of the Human Auditory Steady-State Response

AUTHOR : Daniel J. Bosnyak, B.A. Hons (McMaster University)

SUPERVISOR : Professor Larry E. Roberts

NUMBER OF PAGES : VII, 102.

Abstract

Auditory evoked potentials recorded in the electroencephalogram consist of differentiable components that can be roughly classified as either *transient* or *steady state* responses. The latter, recorded at stimulus rates above about 3 Hz, have advantages for study of the auditory system. A large number of such responses can be collected rapidly to enhance the reliability of the resulting average; because the cortical sources of the response localize to the region of Heschl's gyrus, a picture of neural activity occurring in this region (primary auditory cortex, AI) is gained. An interesting feature of the steady-state response is that it reaches its peak amplitude at stimulation rates near 40 Hz. The experiments of this thesis investigated the mechanism and dynamics of the auditory steady-state response (SSR). The experiments were guided by the linear summation model which attributes the 40 Hz amplitude peak to the summation of transient "middle-latency" responses which contain a 40 Hz component and localize to Heschl's gyrus, overlapping the cortical sources of the SSR. Alternatively, it has been suggested that resonant properties of the auditory system may be responsible for the peak in SSR amplitude at 40 Hz.

Experiment 1 investigated how auditory evoked potentials change as stimulus rate increases from the transient to the steady state range. Stimuli consisting of 6 second long trains of stimulation at 1.5 Hz, 4 Hz or 13 Hz were presented to subjects in either the auditory, visual or somatosensory modality. The transient "on" response, "off response", and intervening responses evoked by stimuli in the train were recorded at each stimulus rate. Using the auditory recordings, frequency-amplitude functions were synthesized from the responses recorded at each of the stimulation rates. Only the 13 Hz response produced a frequency-amplitude characteristic with a prominent peak in response amplitude at 40 Hz. A new signal processing procedure, based on the Hotelling T^2 statistic but applied in a procedure similar to a spectrogram, was used to show dynamics in the steady-state response which were not observable in the time-domain averages.

In Experiment 2 a direct test of the linear summation model for the auditory steady-state response was performed. According to this model each of the individual stimuli in a train of steady-state stimulation induces an identical response referred to as the 'source transient'. Using a steady-state stimulus whose stimulation rate varied continuously from 10-50 Hz, I derived a 'source transient' using a deconvolution process (Gutschalk, Mase et al., 1999). The source transient was then used to reconstruct the 10-50 Hz response assuming that linear summation was responsible for the entire response. The results did not find evidence of a nonlinear contribution to the generation of the SSR at any stimulus rate, and were therefore consistent with a linear model even at stimulus rates near 40 Hz. The source transient was also used to reconstruct a traditional 40 Hz steady-state response and compared to a recorded version of this response. The phase and amplitude of the reconstructed response accurately predicted the recorded response on the basis of individual subjects. The morphology of the group average 'source transient' was comparable with the middle-latency portion of the transient response recorded to a 1.5 Hz stimulus although its amplitude was reduced.

In Experiment 3 I applied a steady-state stimulus procedure to evaluate the plasticity of transient and steady-state components of the auditory evoked potential. Non-musician subjects were trained to discriminate small increases in the pitch of a 1 second long segment of a 40 Hz amplitude modulated tone with a carrier near 2 kHz. Using appropriate methods of signal processing, I separated N1 and P2 components of the transient response thought to reflect activity in secondary auditory cortex (A2) from the SSR reflecting activity in primary auditory cortex (A1). Training enhanced the amplitude of the P2 component in both hemispheres and the N1c component in the right hemisphere. The amplitude of the SSR was not enhanced by training; however, the phase of the response was modified for a time window commencing near the rising edge of the P2 component. Control measurements confirmed that the P2 and SSR were separate brain events. The conclusion contains a discussion about how auditory evoked potentials may reflect neural mechanisms underlying remodeling of the auditory cortex by experience.

Overall the results of these studies suggest that linear summation gives an adequate description of the SSR when source transients are recorded at frequencies > 4Hz. My results are also consistent with neural generators for this response originating in Heschl's gyrus. Temporal properties of the steady-state response appear to be modifiable by training in adult non-musicians, although amplitude is more resistant to change. Exploration of a wider range of procedures may uncover additional dynamics in the SSR not detected here. These dynamics reflect fundamental mechanisms by which the brain encodes its sensory input.

Acknowledgements

I would like to sincerely thank my supervisor, Dr. L.E. Roberts for his invaluable support, assistance, and insight throughout the process of producing this document. Bill Gaetz provided an inexpressible amount of moral support without which I would never have reached the end. Rob and Amy Eaton assisted in gathering much of the data presented here, but their friendship was even more valuable. Graeme Moffat also provided both friendship and assistance. Finally I would like to acknowledge the support of my family and my wife Karen, all of whom provided support and encouragement through long years.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

Our species has evolved over the millennia such that each of us is equipped with sensory receptors that respond to physical continua present in the environment. Examples of such continua are the wavelength of light and the frequency of sound. We can be quite certain that receptor systems responsive to these continua are coded genetically, because organisms lacking such receptors would have been at a reproductive disadvantage relative to their conspecifics in the mating game. However, information conveyed to the brain by the sense organs changes on a millisecond time scale as events occur in the environment and as we move through physical space. In addition to their dynamic properties, the particular sensory patterns that we experience (for example, the sounds we hear or objects we see, and the sensory contexts of these events) are typically unique to each individual and therefore cannot be anticipated by a genetic code. Under these circumstances, how does the brain represent its dynamic sensory world?

In order to solve the problem of sensory uniqueness, our brains must be equipped with mechanisms for representing the fine details of sensory information and for updating that information on a continuing basis. In addition, an efficient form of representation would demand an organization that dedicates a greater amount of cortical 'processing power' to the relevant stimulus patterns. The experiments of this thesis investigate the question of how auditory information is represented in the brain, and how these representations are modified by experience with behaviorally relevant auditory signals. I address these questions by investigating dynamic and neuroplastic properties of different components of the auditory evoked potential (AEP) which is evoked by acoustic

stimulation and recorded in the electroencephalogram (EEG) from electrodes placed on the scalp.

Organization of the Auditory System

In the human auditory system the means of transduction of the physical sound stimulus into a neural pattern via the mechanisms in the cochlea (the auditory sense organ) is fairly well understood. The basilar membrane of the cochlea performs a frequency analysis of the sound signal, so that the physical stimulus is transformed into a 'place' code where vibration on different positions along the length of the basilar membrane represents energy at different frequencies in the sound signal. The original account of this behavior, which was proposed by Helmholtz (1877/1954) and later tested by Von Bekesy (1960), has lately been modified by such complexities as the mechanical resonance of hair cell stereocilia (Lippe, 1986) and seemingly active intensity dependent peak shifts in inner hair cell response properties (Chatterjee and Zwislocki, 1997). Nevertheless the strict organization of the innervation of the receptors placed along the length of the basilar membrane creates a neural code in the auditory nerve where activity at certain frequencies in the stimulus is represented by activity on certain nerve fibres. Nearby fibres represent nearby frequencies, resulting in a smooth gradient of frequency representation across the nerve. This arrangement is known as a 'tonotopic' organization. This description is an oversimplification; information about the sound signal is also carried by the timing of the spikes in the fibres (Sachs, 1984; Cooper, Robertson et al.,

1993), and descending afferent input to the hair cells and other nuclei provides some form of cortical feedback whose function is not currently understood. However, by and large the representation at early levels is not dynamic. Similar sounds produce similar neural patterns regardless of the state of the organism.

In mammals the tonotopy found at the basilar membrane is preserved in some form at the various stages of the auditory projection. These stages include the cochlear nuclei (Bourk, Mielcarz et al., 1981), the superior olivary complex and the lateral leminiscus (Kelly, Liscum et al., 1998), the inferior colliculus (Clopton and Winfield, 1973; Adams, 1979), and the medial geniculate nucleus (Imig and Morel, 1984). This tonotopic arrangement extends into primary auditory cortex (AI) located in the superior temporal gyrus (Brodman's areas 41 and 42) and beyond. Cortical tonotopy has been detailed in nonhuman primates functionally (Imig, Ruggero et al., 1977; Merzenich and Brugge, 1973) and anatomically (Rauschecker, 1997; Rauschecker, Tian et al., 1997; Imig, Ruggero et al., 1977; Hackett, Preuss et al., 2001). It is generally accepted that auditory cortex can functionally and cytoarchitectonically be divided into three regions: (1) the core region, likely the primary auditory projection area, (2) the belt areas, consisting of seven or eight different regions surrounding the core region which receive some thalamic input but likely receive most input from the core area (Rauschecker, 1997) and perform higher level auditory processing, and (3) the parabelt region, which has connections to the belt regions as well as connections to possible sensory modality integration areas in prefrontal cortex (Romanski, Bates et al., 1999; Kaas and Hackett, 2000) and where no tonotopically organized areas have yet been described.

The core region, found along both the dorsal and lateral surfaces of the rostral half of the superior temporal gyrus, has two fields of neurons that respond to pure tones (Merzenich and Brugge, 1973; Imig, Ruggero et al., 1977) – a primary field, which receives projections from the medial geniculate, and a more rostral field. These fields are organized tonotopically, with high best frequencies for the neurons towards the front end of the primary field and low best frequencies at the caudal end. The arrangement in the second (more rostral) field is opposite (Morel, Garraghty et al., 1993). More recent studies (Kaas and Hackett, 2000, 1998; Hackett, Stepniewska et al., 1998) have detected a third tonotopic map within the core region, and have shown that all three tonotopically organized areas receive input from the ventral division of the medial geniculate nucleus. These core regions approximate primary auditory cortex (AI) in older but still useful terminology. The belt region also contains some tonotopically organized areas (Kaas and Hackett, 1998, 2000), but tonotopy is not characteristic of the parabelt region. These two regions roughly encompass "secondary auditory cortex" (AII) in earlier terminology.

The auditory system encodes more, however, than simply the spectral properties of sound. Neurons, particularly at the cortical levels, display complex tuning properties that respond to the loudness envelope of the sound signal, to conjunctions between temporal and spectral features (for example, speech formants), to time differences in the arrival of auditory signals between the two ears, and to variations within each of these properties including spectral bandwidth and sound intensity (Schreiner, Read et al., 2000). It is by means of the diversity of tuning properties that the brain is able to represent the complex features of sound cues that we encounter in the environment.

Plasticity of the Auditory System

Until about a decade ago it was widely assumed that the tonotopic organization of the auditory cortex and the cortical territories devoted to the processing of particular sound features was fixed after the early developmental years. However, studies of cortical representations for the sensory systems (visual and somatosensory as well as auditory) have lately been challenging the view that the ability of cortex to reorganize itself in response to a change in the nature of the input it receives is restricted to some early sensitive period (Hubel and Wiesel, 1970; Fox, 1992). It is becoming clear that this ability to reorganize continues at least to some degree into adulthood. Basic sensory areas that have presumably been fixed in their organization for some number of years have been shown to undergo substantial reorganization in time periods of a few weeks (Mogilner, Grossman et al., 1993) and even after only a few hours. From this, one might theorize that the organization of even primary sensory cortex is much more dynamic and fluid than the textbook Penfield homunculus indicates.

Although the basic map-like isomorphism between the arrangement of the sensor array (somatotopic, tonotopic, retinotopic) and the cortical surface always remains intact, and the early environment has a persistent influence on the conformation of the cortical (Zhang, Bao et al., 2001) or subcortical (Sanes and Constantine-Paton, 1985) map, a shift in the 'importance' of some part of the sensory milieu may result in a rearrangement of the map to more efficiently encode the environment. This reorganization can be a simple shift in the best frequency of pure-tone responding neurons, or a change more complex in nature. Such reorganization probably occurs at multiple levels with different time

constants. For example, structural changes involving the formation of new synaptic connections or the elimination of existing connections would take many days to evolve, but changes in the strength of existing synapses can occur as the result of a few afferent spikes.

Reorganization of cortical auditory areas has been demonstrated in animals in a number of studies. In cochlear lesion studies leading to missing frequency representations in auditory cortex, typically an overrepresentation of frequencies immediately adjacent to the missing frequencies is observed in the tonotopically organized primary (core) projection area, suggesting that the deafferented areas 'retune' themselves to surviving inputs (Rajan and Irvine, 1998; Seki and Eggermont, 2002; Kakigi, Hirakawa et al., 2000). In the case where some stimuli (typically tones of a particular frequency) are made more 'important' either through a classical or instrumental conditioning scheme, or by pairing stimuli with cortical or basal forebrain electrical stimulation, neurons have been shown to shift their best frequencies towards the 'important' frequency. For example, reorganization of tonotopicity in mouse colliculus due to focal cortical stimulation (Yan and Ehret, 2001) has been shown, as well as reorganization induced by basal forebrain stimulation in rats (Kilgard, Pandya et al., 2001) and guinea pigs (Dimyan and Weinberger, 1999) due to a classical conditioning procedure (Weinberger, Javid et al., 1993; Weinberger and Bakin, 1998).

Although the tonotopic organization of primary auditory cortex has been verified using single unit recordings in humans (Howard, Volkov et al., 1996), there have been no demonstrations of plasticity using indwelling electrode recording techniques in humans.

Study of human subjects requires noninvasive recordings such as those provided by electroencephalography (EEG) or its magnetic counterpart magnetoencephalography (MEG) that record the activities of populations of neurons in the underlying neocortical laminae.

EEG/MEG Measurement of Plasticity in Auditory Organization

Auditory stimulation evokes a cascade of activity in subcortical structures and regions of the auditory cortex that can be recorded by EEG or MEG. Responses recorded by EEG are known as auditory evoked potentials (AEPs) and those recorded by MEG as evoked auditory fields (AEFs). The EEG recording of the response to a single transient auditory stimulus, shown in Figure 1.1, typically consists of a series of positive and negative deflections that occur at characteristic times with respect to the stimulus (in this case a click) onset. As the signal-to-noise ratio of a response to a single stimulus is relatively low, these transient responses are generated by averaging together many individual responses.

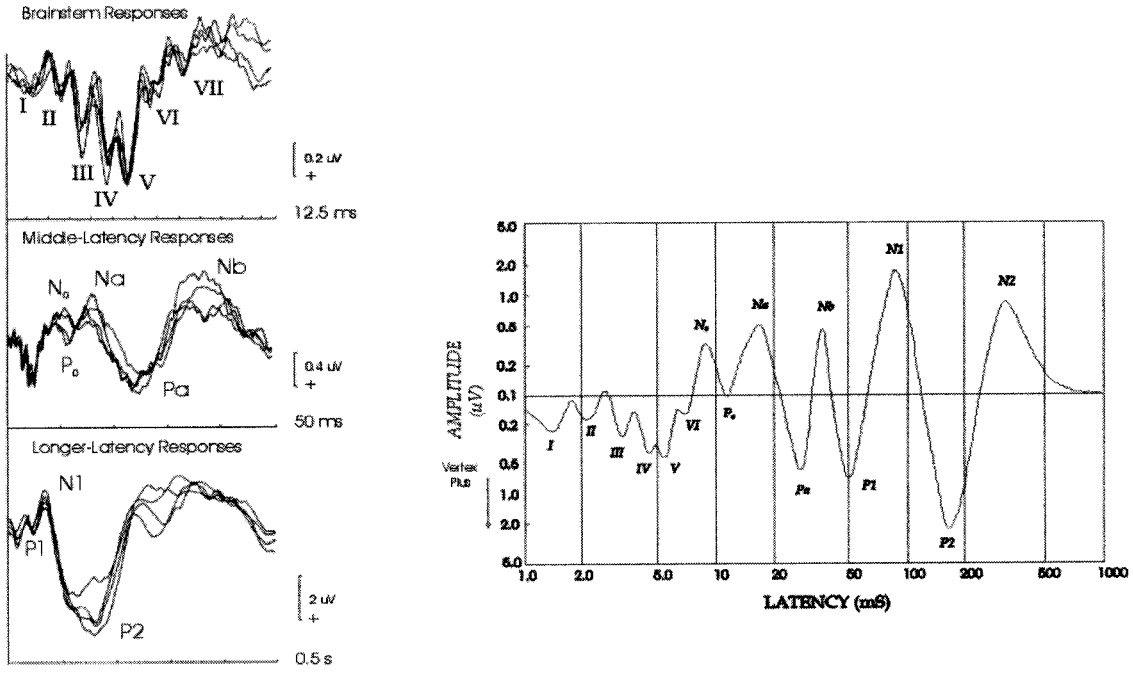


Figure 1.1. Transient responses of the auditory system. The left panel shows the brainstem, middle-latency and longer-latency responses at three different time and voltage scales. The panel on the right uses logarithmic time and voltage scales to depict all three groups of responses simultaneously. (adapted from Picton 1974)

Each of the deflections shown in Figure 1.1 probably reflects activity generated by clusters of neurons that perform local computations and exchange information with other regions of the brain. The various deflections have been localized to different brain areas. The earlier components occurring before about 10 ms, labeled I-VI, are due to activity in subcortical structures and are collectively known as the auditory brainstem response (ABR). These responses are somewhat difficult to collect, requiring a large number of repetitions and a high sampling rate to produce a stable average response (Stapells and Picton, 1981). The "middle latency responses" (MLRs), which consist of a series of three or four sinusoidal cycles at a frequency of about 40 Hz occurring between 8 ms and 50 ms post-stimulus, are thought to reflect activity in primary auditory cortex. The

components of this response complex have been labeled Po (12msec), Na (16msec), Pa (25msec), Nb (36msec), and P1 or Pb (50msec) (Picton, Hillyard et al., 1974). Neuromagnetic analogues of these components have also been recorded; both the P30m (Pa) and the P50m (Pb) localize in supratemporal auditory cortex (Makela, Hamalainen et al., 1994). Following the MLRs are the late AEPs consisting of the familiar N1 event with a latency around 100 ms and the P2 occurring at about 200 ms after stimulus onset. At least some of these responses appear to be tonotopically organized (Pantev, Hoke et al., 1988). In simultaneous MEG/EEG recordings the source location for the N1 and N1m was found at deeper locations for higher frequency stimuli, while the source locations for the Pa/Pam showed an opposite organization, having more superficial source locations for higher frequencies. The Pa tonotopic map was located in primary auditory cortex anterior to the N1m/N1 map (Pantev, Bertrand et al., 1995).

Changes in the organization of auditory cortical areas consequent on a change in the input to the auditory cortex should be reflected in AEPs and AEFs. Such changes were first detected in the study of special populations. Examples are cortical reorganizations in humans due to high frequency hearing loss (Dietrich, Nieschalk et al., 2001), expansion of auditory cortex in blind subjects (Elbert, Sterr et al., 2002), and cortical changes following stapes substitution in otosclerotic patients (Tecchio, Bicciolo et al., 2000). Musicians have been shown to have larger middle-latency evoked responses (Schneider, Scherg et al., 2002) and magnetic N1 responses (Pantev, Oostenveld et al., 1998) than controls.

This thesis investigated whether changes in auditory cortical organization could be induced experimentally. Specifically, we studied (in chapter 4) how the AEP waveform was altered by training a novel pitch discrimination in nonmusician subjects. We also examined how a specific component of the AEP, the steady-state response (SSR), is affected by such training in the same subjects. The SSR was of special interest because the cortical sources of this response appear to be tonotopically organized and localize to Heschl's gyrus in the auditory core area (AI). We turn to this response next.

Steady-State Auditory Stimulation

The morphology of the human evoked auditory potential varies with the rate of auditory stimulation. If stimuli such as clicks or short tone bursts are presented at long inter-stimulus intervals, we record the 'transient' responses discussed above. Transient responses are responses recorded at interstimulus intervals (ISIs) long enough such that brain activity recovers before the next stimulus appears. If the ISI is shortened such that responses to successive stimuli begin to overlap, we can record what is known as a steady-state response (SSR). The transient response gives way to steady state pattern as the stimulation rate exceeds about 2 Hz. This is also the point at which the stimuli lose their individual identities and are perceived as a continuous stream. Unlike transient responses, whose frequency components change over time, the frequency components of the SSR ideally remain constant in amplitude and phase after the response stabilizes, although phase and amplitude will change systematically if the stimulus rate is changed (Regan, 1989). The collection of steady-state responses provides a number of technical

advantages including the ability to measure neural activity in primary auditory cortex where its cortical sources appear to be found (see Chapter 4 later, for a review of the localization evidence).

Interest in the response to auditory stimuli presented at steady-state rates stems from the work of Galambos (Galambos, Makeig et al., 1981; Galambos, 1982) who presented clicks and tone bursts at various repetition rates from 10-55Hz and found that 40Hz stimulation produced a much larger response in the EEG than stimulation at other rates. Galambos et al. observed that the SSR was present at sound intensities very close to the threshold of hearing and proposed its use in objective audiometric testing, improving on the then current technique of measuring auditory brainstem responses (ABR) for assessment of hearing in infants and children. This proposal prompted the development of a number of signal processing techniques which are able to detect the presence of a SSR at low stimulation levels and after relatively short stimulation periods (Stapells, Makeig et al., 1987; Picton, Vajsar et al., 1987; Tang and Norcia, 1993, 1995; Victor and Mast, 1991; Dobie and Wilson, 1990). The use of a steady state stimulus rather than the transient stimuli used in ABR testing allowed the use of these advanced objective techniques for closer estimation of actual threshold; also, an amplitude modulated sinusoidal stimulus concentrates most of the stimulus energy at the carrier frequency unlike transients which spread energy into a wide band of frequencies. The 40 Hz response in infants was found to be unreliable (Stapells, Galambos et al., 1988), but a response to stimuli in the 80-100 Hz range has been reported (Lins, Picton et al., 1995) and seems to be reliable in young children during sleep (Aoyagi, Kiren et al., 1993).

In adults the SSR is more robust, particularly when the rate of acoustic stimulation approaches 40 Hz. The fact that the SSR shows a peak in the frequency-amplitude characteristic at 40 Hz has made it a focus of interest for reasons other than methodological. The study of 40 Hz in the brain has been of interest to researchers for some time since the discovery that olfactory stimulation excited bursts of 40 Hz activity in the olfactory bulb (Adrian, 1942). The observation that the auditory steady-state response was preferentially excited by 40 Hz stimulation led to speculation on the nature of the response and specific theories indicating why the response should peak at 40 Hz. Spontaneous bursts of 40 Hz oscillation have been recorded intracranially in rat auditory cortex (Franowicz and Barth, 1995). Several authors have proposed that the 40 Hz oscillation is involved in perceptual integration (Gray, Konig et al., 1989; Gray and Singer, 1989). Understanding the source of the 40-Hz SSR has therefore become a question of some interest.

Generators of the Steady-State Response

Although the idea that the SSR response might represent a 'driven' version of the spontaneous 40 Hz oscillations has received some consideration, the most widely discussed hypothesis advanced to explain the steady-state response makes use of the characteristics of the response to transient auditory stimuli. In particular, middle latency responses (MLRs, Figure 1) consist of a series of three or four sinusoidal cycles at a frequency of about 40 Hz, occurring between 8 msec and 50 msec post-stimulus.

Galambos et al. (Galambos, Makeig et al., 1981) first proposed that the steady-state response might be composed of a series of MLRs overlapping each other and summing algebraically. They provided a simple graphical demonstration of how this response might be synthesized from multiple overlapping Nb-Pb waves taken from a recording of the response to 10 Hz stimulation. This model predicted the amplitude peak near 40 Hz, as well as minimal amplitudes seen at 25 and 55 Hz.

The middle latency response complex has been fairly widely studied. It is thought to reflect activation of auditory cortex and possibly thalamus, distinct from the early ABR components and the late components (N1, P2, N2) occurring 80-300 msec after the stimulus which seem to be generated by a more diffuse activation of the auditory and other cortical areas. Both the P30m (Pa) and the P50m (Pb) components of the MLR localize in supratemporal auditory cortex (Makela, Hamalainen et al., 1994) and the P30m in particular to sources in Heschl's gyrus in the auditory core (Yvert, Crouzeix et al., 2001). There are a number of differences among the middle-latency components which suggest they might have different generators. Pa is much less sensitive to stimulus rate (and thus is a better candidate for summation) than P1/Pb, which essentially disappears above 1 Hz (Erwin and Buchwald, 1986a), although the findings on this point are not entirely consistent. The Pb also disappears during deep sleep and reappears during REM, whereas Pa remains stable (Erwin and Buchwald, 1986a). The scalp potential maps for Na and Pa suggest these components have different generators (Deiber, Ibanez et al., 1988).

The MLR superimposition analysis of Galambos has been replicated by other researchers (Hari, Hamalainen et al., 1989; Stapells, Linden et al., 1984; Stapells, Galambos et al., 1988). This model was also supported by the observations of Pantev and Elbert et al. (1993) who found that the region of cortical activation identified through dipole analysis for the steady-state field appeared indistinguishable from the source of the Pam wave, the magnetic equivalent of the Pa wave. The source of these two waves was found to be located approximately 1cm medial to the source of the M100, consistent with a cortical source in primary auditory cortex. Additional support for the hypothesis of identical generators for the MLR and SSR was provided by the observation that only the Pam wave among the components of the middle latency response is seen consistently for all subjects (Pelizzone, Hari et al., 1987) when recorded magnetically, as is the steady-state response.

Recently, however, a number of observations that conflict with the hypothesis that the steady-state response is generated by the summation of middle latency waves have called this view into question. In a series of papers (Azzena, Conti et al., 1995; Santarelli, Maurizi et al., 1995; Santarelli and Conti, 1999) researchers performed a much more detailed study of the correspondence between recorded steady-state signals and signals synthesized from summed overlapping transient responses. These researchers first recorded the MLR evoked by a stimulus presented at 7.9 Hz. They then compared responses measured with 30, 40, 50, and 60 Hz steady-state stimuli to synthesized versions of the same rates using the MLR recorded at 7.9 Hz. They found a close correspondence between measured and synthesized waves (both for amplitude and phase)

at 40 Hz, but at the other rates the predicted waves did not match the measured ones. A strict interpretation of the linear summation hypothesis would allow the SSR at any rate to be constructed from the linear summation of appropriately spaced MLRs. Additionally, it was found that the response to the last click in a train of eight clicks presented at 40 Hz contained extra components not contained in the middle latency response, suggesting a complex interplay between resonant and adaptational phenomena. These results were replicated using epidural recordings in rats and similar inconsistencies were noted between recorded and predicted waves (Conti, Santarelli et al., 1999). Pantev, Roberts et al. (1996) provided further evidence against the simple summation interpretation, showing that the tonotopic organizations of the SSR and the MLR were quite different. They found that the transient middle latency response was organized so that sources for higher frequencies were more lateral with respect to the saggital midline in the region of Heschl's gyrus than sources for lower frequencies, while for steady-state responses higher frequencies were more medial.

PLAN FOR THESIS

This thesis addresses the problem of how the auditory cortex encodes its dynamic sensory input. I will be presenting three studies that are organized as follows.

In order to study the dynamics of sensory representations, we must be able to make accurate and repeatable measurements of them. Experiment 1 (presented in Chapter 2) reports initial observations in which I examined transient and steady-state responses and the transition between them. I also describe signal processing procedures that were

developed to analyze the steady state response. My interest is focused mainly on the auditory SSR; however, for purposes of comparison visual and somatosensory SSRs are also analyzed.

Experiment 2 is presented in chapter 3. The goal of this study was to evaluate the linear summation hypothesis of the 40-Hz auditory SSR. A technique called temporal deconvolution (Gutschalk, Mase et al., 1999) was applied to determine whether regions of nonlinearity could be identified in the SSR frequency/amplitude characteristic when the stimulus rate was swept continuously from 10 to 50 Hz. Source waveforms deconvoluted from the swept SSR were compared to a continuous 40 Hz SSR and to transient MLRs recorded from the same subjects.

In Chapter 4 I report an experiment in which transient and steady state procedures were used to investigate remodeling of the auditory cortex by training at pitch discrimination, to determine if this training resulted in changes to representations in primary (AI) and secondary (AII) auditory cortex. The behaviors of the SSR localizing to AI and several transient responses localizing to the region of AII were investigated.

Chapter 5 presents a summary and discussion of the results in the previous chapters.

Chapter 2

Experiment 1
TRANSITION FROM TRANSIENT TO STEADY-STATE RESPONSES IN THREE SENSORY MODALITIES

Introduction

Most EEG or MEG studies of the auditory system use stimuli that can be classified as either 'transient' or 'steady-state'. Regan introduced a generally accepted criterion for distinguishing between these modes of stimulation (Regan, 1989, 1982). For stimuli to be considered transient they must "follow each other at sufficiently long intervals that [the response of the system] returns to its initial state before the next stimulus occurs". Steady-state stimulation, on the other hand, occurs at rates where responses to successive stimuli overlap. Application of this concept to the auditory system is instructive. Auditory EEG and MEG responses evoked by a brief click contain identifiable components occurring up to about 300-400 ms after the stimulus (Picton, Hillyard et al., 1974). This implies that if the click is presented at frequencies above about 3 Hz (that is, at interstimulus intervals less than 333 ms), successive transient responses will begin to overlap, such that the stimulus and its accompanying response begin to enter the "steady state" range above this frequency.

This analysis invites the hypothesis that steady-state responses consist of temporally overlapping components evoked by transient stimulation of the auditory system. Galambos et al. (1981) evaluated this hypothesis by recording the steady state potential evoked in human subjects by repetitive auditory stimuli occurring at rates from 10 Hz to 50 Hz. Their results, which are reproduced in Figure 2.1, showed that the largest amplitude response was obtained to stimulus repetition rates around 40 Hz. A subsequent analysis by these investigators showed that this frequency/amplitude characteristic

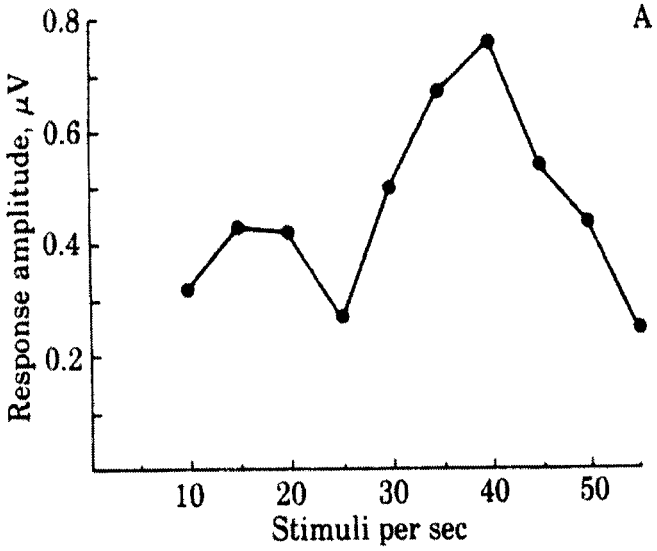


Figure 2.1 Function relating SSR amplitude to SSR stimulation rate, showing a peak at 40 stimuli/s. From Galambos (1981)

(sometimes called a "modulation rate transfer function") could be reproduced by summating transient auditory "middle latency responses" (MLRs) at the various stimulus repetition rates. Because MLRs have a response latency and wave period of about 25 ms, successive MLR waveforms are in phase and survive signal summation when presented at about 40 Hz but are subject to cancellation when stimuli are presented at other frequencies. The thrust of this analysis is that the steady-state response does not contain new information but is made up of linearly summated components of the auditory transient response.

There are, however, reasons to question this analysis. The MLRs used by Galambos et al. (1981) for demonstration of this theory were recorded to clicks occurring at 10 Hz, which is not a transient stimulus using the criterion of Regan (1982, 1989). Whether summation of AEPs evoked in the transient range will yield the 40-Hz SSR waveform can therefore be questioned. The critical frequency range for emergence of steady state responses is notable in other respects. Studies of paired-pulse facilitation using recordings from single neurons indicate that membrane dynamics begin to change above about a 4 Hz rate. Postsynaptic responses begin to broaden, allowing for temporal

integration of synaptic inputs that may alter network behaviour in the auditory cortex (Buonomano, 2000). This is also the frequency range over which discrete pulses lose their individual identities and seem to merge into a single percept. Bullock (Bullock, Karamursel et al., 1994) noted that the termination of visual steady state stimuli presented above about 4 Hz evokes a distinctive "off" response that is not observed when trains of transient stimulation are terminated after a period of time. These properties of the steady-state response have been noted for the visual system but may not be restricted to the visual modality.

Experiment 1 was a preliminary study undertaken to provide a point of departure for investigation of the auditory steady-state response. One goal was to examine the transition of the AEP waveform evoked by trains of repetitive stimulation over stimulus rates varying from 1.5 to 13 Hz. Transient "on" and "off" responses were examined as well as the waveform between these events during the stimulus train. A guiding question was whether changes in transient responses recorded at low stimulus frequencies could account for the steady-state waveform observed at higher stimulus frequencies. Although my main focus was on the auditory system, visual and somatosensory responses were investigated for purposes of comparison. A second goal of Experiment 1 was to develop signal processing tools for use in later studies.

Method

Stimuli

The auditory stimulus was provided free-field via a speaker situated about 1m in front of the subject. The stimulus was presented at 78 db SPL. It consisted of a 10ms long burst of 1 kHz sine wave, with a 2ms rise and fall (cos2 envelope). The waveform was synthesized digitally at a 100 kHz sampling rate and generated by a Tucker-Davis DA-2 digital-to-analogue converter.

The visual stimulus was provided by a 14mm diameter red LED positioned about 1m from the subject, surrounded by a black board to minimize extraneous visual input. The LED was flashed on for 10ms for each stimulation. The lights in the room were dimmed during the experiment. The mean peak intensity of the LED was 168 cd/m2 and the LED appeared very bright. Subjects wore headphones delivering white noise during the stimulation procedure.

The somatosensory stimulus was provided by a solenoid driving a plastic rod with a 1.5mm diameter metal tip. The subjects arm rested on a foam pad with their right index finger placed over a small hole through which the probe could protrude. At rest, the tip of the probe was not in contact with the subject's finger. The solenoid was activated via a transistorized driver for 10ms for each stimulation, bringing the probe into contact with the tip of the subject's finger. Subjects wore headphones delivering white noise to mask any noise from the solenoid which was housed in a sound-attenuating case.

Stimulus Procedure

The stimulus procedure was the same for all three modalities. Stimuli were generated at three different rates: 1.5Hz, 4Hz, and 13Hz. Each trial consisted of six seconds of stimulation, followed by a 2 second gap timed from the first omitted stimulus, with the next trial beginning immediately for 40 trials in a block. At 1.5Hz the subject would experience 9 light flashes, tone pips, or finger pokes (depending on the condition); at 4Hz 24 stimuli; at 13Hz 78 stimuli. After each block, the subject rested briefly and the stimulus rate was changed. Each subject experienced three blocks at each rate (one block at 1.5 Hz, pause, one block at 4 Hz, pause, one block at 13 Hz, pause, repeated three times), for a total of 120 trials. Subjects were instructed to count the gaps in the stimulation, and reported their counts at the end of each block. Subjects were instructed not to close their eyes during the experiment.

Subjects

Subjects were either volunteers or were recruited from the Introductory Psychology subject pool. Written informed consent was obtained. Ten subjects participated in the auditory condition, twelve subjects in the somatosensory condition, and thirteen in the visual condition.

Recording

EEG was recorded from 21 electrodes placed according to the 10-20 system, using an Electrocap with tin electrodes. The CZ electrode was used as the reference, and earlobe electrodes were recorded for off-line re-referencing. The EEG was recorded

continuously at a digitization rate of 1 kHz, DC-200Hz. Digital markers were recorded for the first and the last stimulus in each trial for off-line epoching.

Data Analysis

A number of different data epochs were used for analysis. For analysis of transient off and on responses, an epoch of 4300 ms including the last 1500 ms of stimulation, the 2 second off interval, and the first 800 ms of the next stimulus train was used. Epochs were considered contaminated by artifact if the FP1 or FP2 electrodes exceeded ±100uV, and were rejected from averaging. An average of the accepted epochs was generated for each subject and then the subjects were combined into grand averages. The grand averages were low-pass filtered at 20 Hz using a zero-phase shift FIR filter. For analysis of the SSR, a waveform for the entire stimulation period -500 ms pre-stimulus extending for 8100 ms was constructed by adding together five overlapping 2 second long sub-epoch averages, constructed after artifact rejection and averaging procedures similar to that described above. This was done to ensure that a large enough number of artifact-free trials were combined in each average; an epoch consisting of the full 8100 ms period would almost invariably contain some artifact.

Results

Experiment 1 was undertaken to describe changes in the evoked response waveform with stimulus frequency and to develop tools for experiments presented later in

the thesis. I focus on descriptive analyses of the results and on attempts to simulate SSRs from transient waveforms.

Transition from Transient to SSR

The time-domain averages for the three stimulation rates and modalities are shown in Figure 2.2. At the onset of stimulation, all rates and modalities produce an 'on' response which changes somewhat depending on the stimulation rate. Differences in the 'on' response between the 1.5 Hz and the 4 Hz condition must be due to the subjects' experience with the stimulus rate, as the stimulation received by the subject is the same in

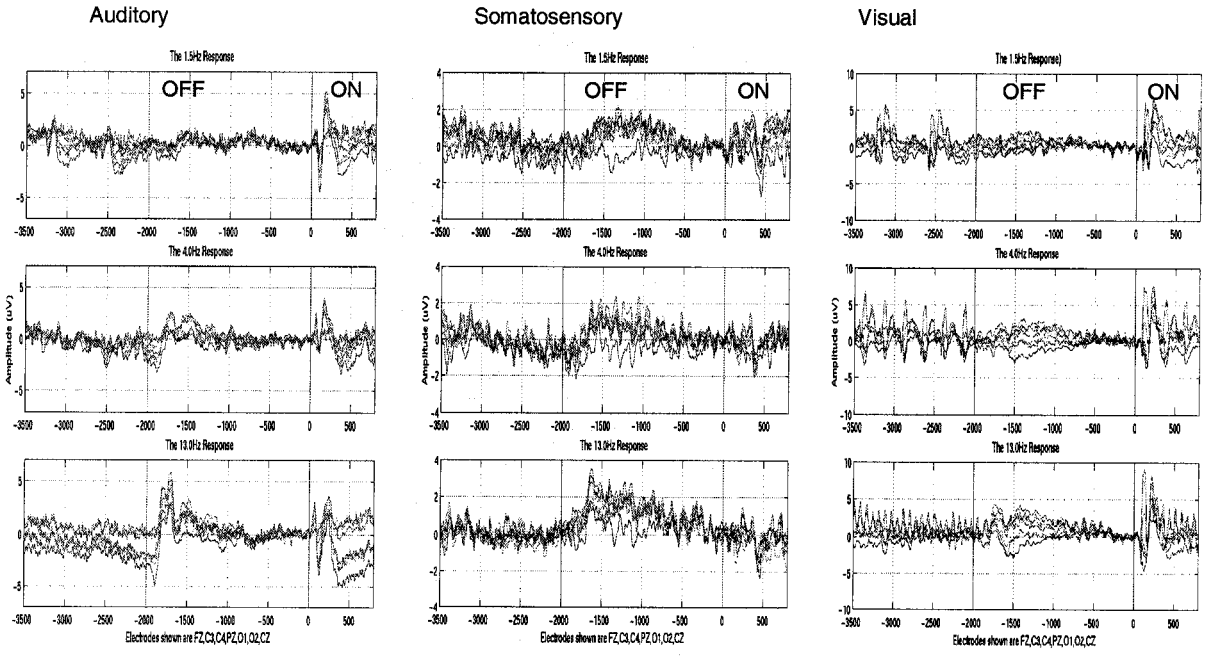


Figure 2.2 Time-Domain group average responses, shown with the 'off' period, between the blue vertical lines, in approximately the centre of each graph. All electrodes are shown, referenced to linked earlobes. Responses to the individual stimuli can easily be seen in the visual case (right), while they are less obvious for auditory stimulation (left), and imperceptible for somatosensory stimulation (centre).

both cases up until 250 ms post-onset. In the 13 Hz condition the subject has received several discrete stimuli during the period of the 'on' response. The 'off' response also appears for all rates and modalities, and appears to increase in amplitude as stimulation rate increases for all modalities. The ongoing response at the higher two stimulation rates is only detectable for the visual stimulus. The auditory stimulus produces a detectable response to each discrete pulse for the 1.5 Hz stimulus but not the other rates. The somatosensory stimulus produces only detectable 'on' and 'off' responses but not any obvious ongoing response.

SSR Analyses

The results in this section rely on the following method of analysis. A 'source waveform' derived from the data, is used to simulate the SSR waveform for a given rate of stimulation. This simulation is accomplished by the linear addition of multiple copies of the 'source waveform' shifted by suitable time intervals, for example shifted by 25 ms to produce a synthesized 40 Hz SSR. If the linear summation hypothesis is correct and the 'source waveform' represents accurately the components that are linearly summated, the synthesized SSR waveform will be similar to a waveform generated by actual stimulation at the synthesized rate.

A number of source waveforms are used in this analysis. All were generated from the group average waveforms for the auditory stimulation data. The waveforms were band-pass filtered 1.5 Hz to 55 Hz using a 5th order IIR filter applied in the forward/reverse directions (zero phase). Source wave 'A' was generated from the response to the first

stimulus in the 1.5 Hz train cut from the stimulus onset to 500 ms post-stimulus when the response had returned to baseline. This is a transient wave representing a stimulus rate of approximately 0.5 Hz as the stimulus occurs after the 2 second off period. Source wave 'B' was generated from the average of the responses to the last 8 stimuli in the 1.5 Hz train, cut from stimulus onset to 500 ms post stimulus. It is a transient from stimulation at 1.5 Hz, and has a reduced amplitude N1/P2 component from the 0.5 Hz transient. Source wave 'C' was generated from the average of the responses to the last 22 pulses from the 4 Hz train, cut from stimulus onset to 250 ms post stimulus when the next stimulus occurred. Source wave 'D' was generated from the average of the responses to the last 75 pulses from the 13 Hz train, cut from stimulus onset to 77 ms post stimulus when the next stimulus occurred.

Synthesis of Frequency/Amplitude Characteristic

Auditory stimulation at rates from 10-50 Hz evokes responses whose amplitudes are known to show a rate/amplitude characteristic that produces maximal response amplitudes at 40 Hz stimulation (see Figure 2.1). The SSR at rates from 4 to 70 Hz in 0.1 Hz steps was simulated using source waves 'B', 'C' and 'D', for electrode Fz. The RMS energy in a 1 second segment of the simulated wave was determined at each rate. The results are given in Figure 2.3. The source 'B' wave characteristic, generated from the transient at 1.5 Hz, shows a peak at approximately 10 Hz, since the 100 ms N1/P2 component is predominant in this wave. A smaller peak is also shown around the 40 Hz rate, due to the summation of the MLRs in the transient. The source 'C' wave shows only low frequency peaks. The source 'D' wave, having no N1/P2 components, shows

only the expected 40 Hz peak, although the amplitude of the simulated response is very small, possibly due to poor signal-to-noise ratio in the source wave.

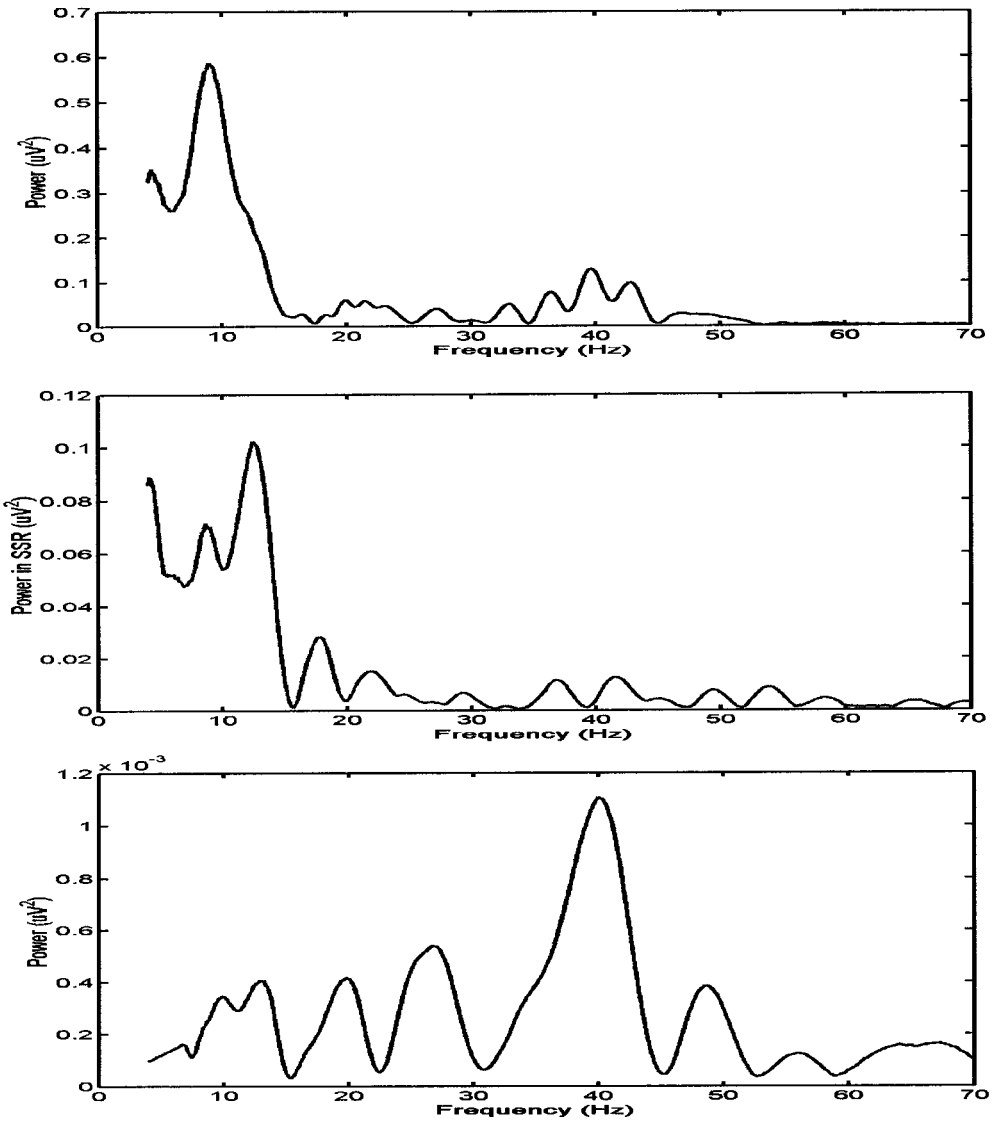


Figure 2.3 Synthesized frequency amplitude/frequency characteristics are shown using three source waveforms: Wave 'B' (1.5 Hz), top; Wave 'C' (4 Hz), middle; and Wave 'D' (13 Hz) bottom.

Synthesis of Onset and Offset of 13 Hz SSR

This analysis was performed to determine whether the on or off responses for the 13 Hz auditory stimulation condition could be captured by linear summation of transient source waves. The onset and offset of stimulation was simulated as described above for 13 Hz using both source waves 'A' and 'B', for a group of frontal and occipital electrodes with respect to a common average reference. The results are shown in Figure 2.4. The middle of the 3 sets of traces shows the recorded 13 Hz auditory response; above and below are the waveforms simulated using wave 'B' and 'A' respectively. The wave 'A' simulation (upper panel) best simulates the 'on' response, but produces an ongoing SSR that is very large compared to the almost undetectable ongoing response in the recorded wave. The wave 'B' (lower panel) simulation produces a smaller but still too large ongoing response, but fails to capture the 'on' response. Neither simulation captures the 'off' response characteristics.

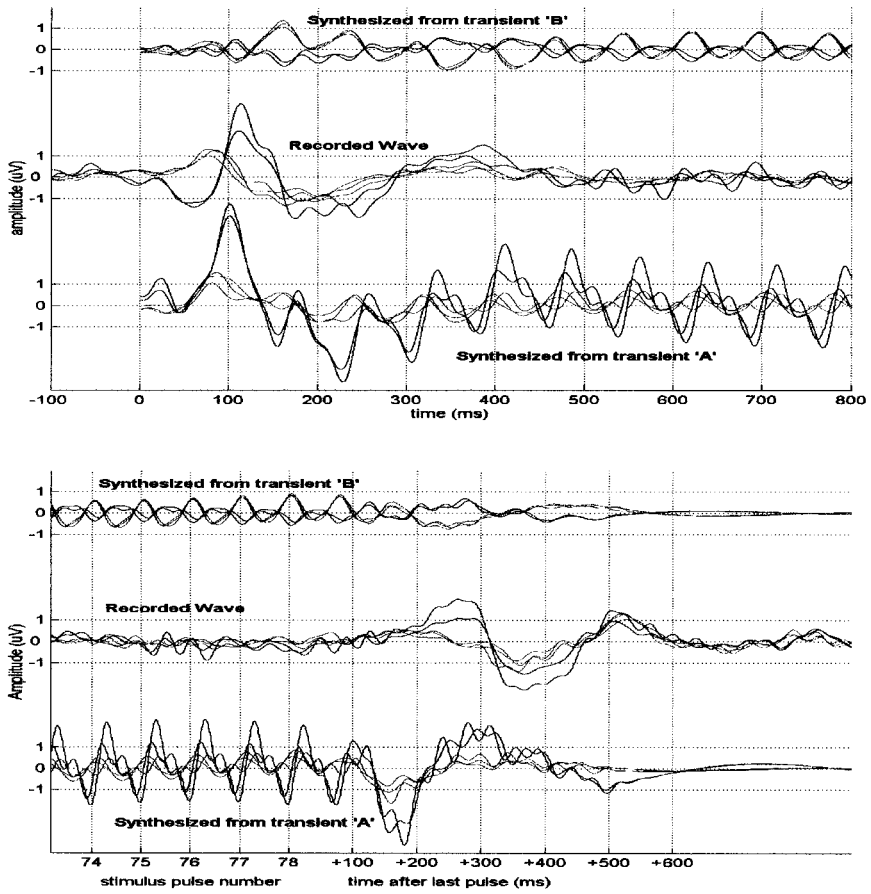


Figure 2.4 Synthesis of the onset (top panel) and offset (bottom panel) of 13 Hz response. The red leads are frontal (F3, Fz, F4) and the blue leads are occipital (O1, O2).

T^2 Analysis

It is evident from studying the filtered time-domain responses that the response to the stimuli in the train after first stimulus was not robust; in some cases it was impossible to detect whether an ongoing response at the stimulus rate was occurring at all, except in the case of visual stimulation. An extensive literature on the use of the steady-state response for audiological and neurological diagnosis has shown a number of signal processing

techniques to be more effective than simple time-domain averaging at the detection of the steady-state response (Valdes-Sosa, Bobes et al., 1987; Victor and Mast, 1991). I adapted one of the most effective detection techniques, the T^2 statistic, to the analysis of this data set.

Since the stimulus repeats at a known rate, either 1.5, 4, or 13 Hz, the response should be a periodic signal either at that rate or a multiple of that rate given that the response is stable. Fourier analysis allows the isolation of the response at a particular rate (in practice within a range of rates which depends on the length of the data sample and the window applied to the data sample). The result of the Fourier analysis for a particular frequency is a phasor which can be thought of as a vector projecting from the origin of the complex plane having an amplitude (the length of the vector) and a phase angle representing the timing of the periodic response with respect to the beginning of the analysis window. The test assumes (like most evoked potential analyses) that the response for each trial is a signal of identical amplitude and phase, contaminated by an independent noise component. The phasors for the set of trials in the analysis form a cluster of estimates of the true response within an elliptical confidence interval; if the confidence ellipse does not include the origin then a response is detected. Although the polar ($z = ae^{i\theta}$) interpretation of the Fourier component is convenient for interpretation, for analytical purposes the Cartesian representation ($z=x+iy$, where x,y represent the cosine and sine amplitudes respectively) is used to decompose the response vector into two independent components. The pair of components are analyzed using the single sample multivariate version of the t-test, Hotelling's T^2 statistic, where

$$t^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \, \mathbf{S}^{-1} \, (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

and:

$$\mathbf{S} = \frac{\mathbf{A}_1 + \mathbf{A}_2}{N_1 + N_2 - 2}$$

$$\mathbf{A}_1 = \sum_{i=1}^{N_1} (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)'$$

$$\mathbf{A}_2 = \sum_{i=1}^{N_2} (\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)'$$

which is evaluated using the F distribution and

$$\mathrm{F}_{(2,n-2)} \sim \mathrm{t}^2 \frac{n-2}{2(n-1)} \; .$$

In practice this analysis method can reliably detect responses when the stimulus is presented 14 dB above threshold using 200 seconds of data (Picton et. al., 1987). Victor and Mast (1991) developed a more effective version of the T^2 statistic, the T^2_{circ} statistic. This version assumes that the variances of the real and the complex parts of the Fourier component are equal and have zero covariance; the confidence region in this case is a circle rather than an ellipse. For sets of trials numbering greater than 64 the two methods' performance was indistinguishable and as all of our analyses are on data sets of 100 trials I used the T^2 statistic.

In the usual application of the T^2 test to steady-state data, the stimulus would be presented continuously and the data sliced into arbitrary epochs consisting of single periods at the stimulus rate for analysis. This would allow for responses only at the stimulus rate and harmonics of this rate, and would allow for only a single T^2 value to be calculated for each set of data for each frequency evaluated. In this experiment, I

presented 120 bursts of 6 s of stimulation (100 trials after artifact rejection) which allowed us to evaluate the T^2 statistic at all frequencies and at various times through the course of the trial, allowing the characterization of the development of the SSR. Consequently, I calculated the T^2 value for all frequencies between 4 and 90 Hz for 100 ms time windows beginning at 400 ms pre-stimulus onset and moving in 10 ms steps to 2500 ms after stimulus offset, in the manner of a spectrogram. The 100 points of data from each trial was windowed with a hamming window and zero-extended to 1000 points in order to get a 1 Hz resolution from the FFT. The results of this analysis are shown in Figure 2.5. This figure represents the grand average of the analyses conducted for each individual subject. The presence of an ongoing response for all stimulation rates for all modalities can be easily detected. The colours of the graph are scaled such that blue indicates a non-significant response while yellow or red indicates significance (unadjusted for multiple tests). The similarity of the responses to each of the 9 pulses at 1.5 Hz can be seen, although the maximum value of T^2 is higher for the first stimulus than for succeeding stimuli in all three modalities (T^2_{max}=75 vs. 63 for visual, 16 vs. 11 for somatosensory, and 37 vs. 15 for auditory). For the auditory and somatosensory stimulation, each stimulus evokes a response in the 40 Hz range, indicating the possible presence of a MLR component for each response. The response at upper harmonics does not fall off monotonically with increasing frequency; rather certain 'bands' show enhanced responses and these bands differ with stimulus modality. This pattern is most apparent at 4 Hz but is consistent at all three stimulus rates. The enhanced frequency band covers 40-60 Hz for auditory stimulation, and 13-60 Hz for somatosensory

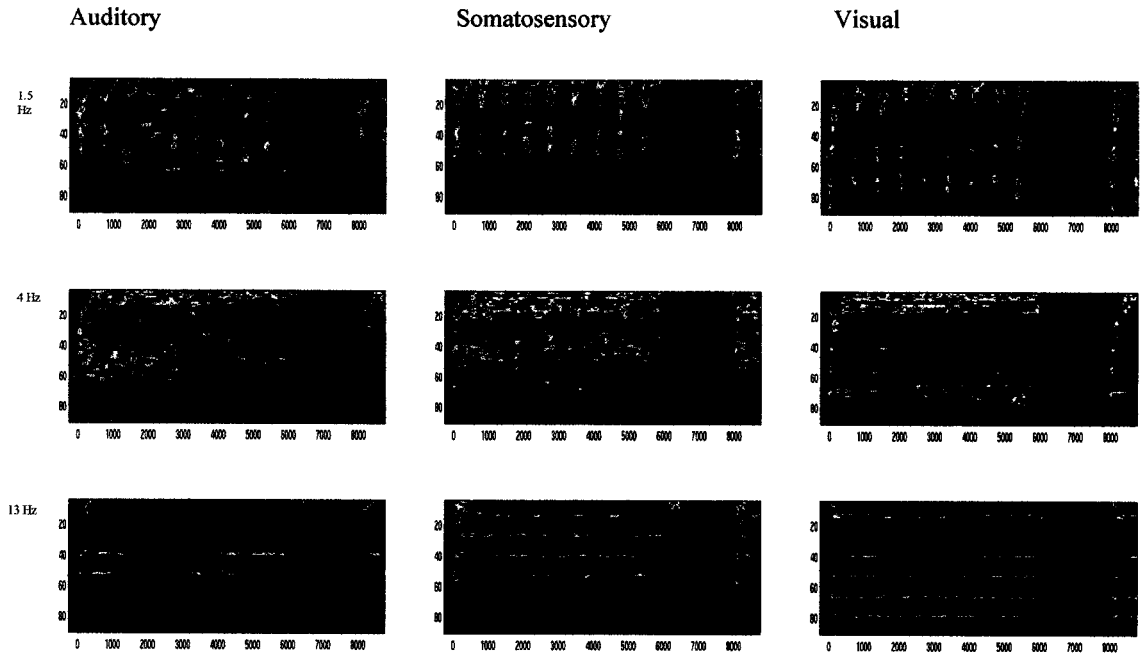


Figure 2.5 t^2 analysis of data. Each graph, with time as the abscissa and frequency as the ordinate, shows the analysis of a train of data from the onset of the first stimulus at t = 0 and the onset of the last stimulus in the train at t = 6000ms, through the 'off period extending from t = 6000ms to t = 8000ms, to shortly after the beginning of the next train at t = 8000ms. The magnitude of the t^2 value is indicated by the colour at that time/frequency point, with blue being non-significant and red being highly significant. In the top panels where the stimulus rate is 1.5 Hz responses to the individual stimuli occurring every 666 ms can be seen. In the bottom row of panels, where the stimulation is 13 Hz, responses only at the harmonics of the stimulation rate can be seen.

stimulation. The pattern is especially obvious for the 13 Hz auditory stimulation, where no significant response is present at 13 Hz or 26 Hz, but robust responses are present at 39 and 52 Hz. After about one second of stimulation, the response fades and returns after about 2 seconds.

Discussion

One goal of Experiment 1 was to examine the transition of the AEP waveform evoked by trains of repetitive stimulation over stimulus rates varying from 1.5 to 13 Hz. In

addition I assessed whether changes in transient responses recorded at low stimulus frequencies could account for the steady-state waveform observed at higher stimulus frequencies. A second goal was to develop signal processing tools for use in later studies utilizing a similar stimulus paradigm.

Effects of Repetition Rate

A transient 'on' response was seen at each stimulation rate, but the 'off' response was prominent mainly at 4 Hz and 13 Hz. This result is in general agreement with Bullock's observation of a "high frequency" off response developing in the visual system above about 2 Hz (Bullock et al., 1994). In my study 'off' responses were seen in all sensory modalities although the form was modality dependent. Small auditory and visual 'off' responses were detected at about 100 ms after the first omitted stimulus; in all modalities a larger and more complex off waveform peaked near 300 ms after the first omitted stimulus with gradual recovery thereafter. It was also apparent that at least in the auditory and somatosensory systems responses to stimuli following the 'on' response were different in form from the 'on' reaction at 1.5 Hz stimulation. This might be understood in terms of the effects of the interstimulus interval (ISI) since it is well known that the amplitude of the transient N1 response decreases with interstimulus interval (ISI). In the procedure of this study the stimulus evoking the 'on' response was preceded by a long ISI (>2.0 sec) and other stimuli in the train by shorter ISIs which varied with stimulus frequency. Of course this begs the question of why ISI matters. There are several possible explanations, including refractory periods of neurons generating these

responses or linear summation of overlapping components causing a cancellation of the observed response. I evaluated the latter explanation by simulations.

Simulations

Galambos (1981) reported a 40 Hz peak in the frequency/amplitude characteristic and reconstructed it from a 'transient' recorded at 10 Hz. My construction from 13 Hz auditory stimulation resembled in broad outline their results, although the amplitude of the simulated wave was larger than the recorded wave. Reconstructions from 1.5 Hz and 4 Hz ongoing transients did not do as well, although a 40 Hz peak was seen in the former. The disparity between these reconstructions suggests that a simple linear summation of transients at any frequency will not work. Rather, the rate for collecting the 'transient' needs to be above a certain transition range which may be between 4 and 13 Hz. Each of the various components of the response has its own recovery cycle characteristics and thus might change differently over the first few stimulations and contribute more or less to the fully developed SSR.

In this respect it is noteworthy that the amplitude of the N1 or the magnetically recorded counterpart the N1m has been shown to be dependent on stimulation rate, with longer interstimulus intervals producing larger responses over the range of one to several seconds (Pantev et al., 1993). In the stimulus rate range between 1 Hz and 10 Hz, the N1m has been shown to essentially disappear once the stimulation rate reaches 2 Hz, with a component at 50 ms gaining prominence at higher rates (Carver et. al. 2002). Some of

the earlier response components have also been shown to have rate dependent amplitude and latency characteristics. There is evidence that the Pa, at around 25 ms latency, remains stable in amplitude for stimulation rates ranging from 0.5 Hz to 15 Hz (Erwin and Buchwald, 1986a; McFarland, Vivion et al., 1975; Goldstein, Rodman et al., 1972). The P1 potential, occurring at 50 ms post-stimulus, has been reported to decrease in amplitude as stimulation rate increases above 1 Hz and may be absent for stimulation rates above 5 Hz (Erwin and Buchwald, 1986a). However, other authors have reported that the P50 amplitude is not reduced at rates up to 16 Hz (McFarland, Vivion et al., 1975) and that there is no difference in amplitude between 0.66 Hz and 2 Hz stimulation (Onitsuka, Ninomiya et al., 2000); however this study did find a significant suppression of the amplitude of the simultaneously magnetically recorded P50m. These authors concluded that this dissociation between the P50 and the P50m indicated that the P50 recorded by EEG might be an overlapping potential generated by multiple sources with different recovery cycles, as the EEG recorded the activity of both radial and tangential sources and the MEG recorded only tangential sources.

I also tried to reconstruct on and off responses by summation; this process was less successful. It is clear that the 'off' response is not well modeled by summation of components in the 1.5 Hz transient. This may indicate that the 'off' response is a different process, possibly the result of the release of inhibition built up by repetitive stimulation. Similarly, it appears that the response to the onset of steady-state stimulation consists of a separate 'on' component containing N1/P2 components that only occur for the first stimulus in the train. The model able to predict both the onset response and the

ongoing response would likely need to include a set of overlapping responses each with its own recovery cycle characteristics. If so, the composition of the response at a particular latency would depend on the stimulus history extending back at least 300 ms and possibly as long as 10 s.

T^2 analyses

These analyses revealed more texture than apparent in time domain averages. An ongoing response was revealed even when this was not apparent in the time domain averages. The analysis of the 4 Hz stimulation, with relatively closely spaced harmonics, was able to show which harmonics responded more powerfully to the stimulus. The auditory and somatosensory stimulus evoked activity in frequency bands similar to the general shape of the frequency/amplitude characteristic measured more conventionally. In addition, the method was able to reveal more information on the dynamics of the formation of a SSR than conventional analysis. The 40-Hz component of the auditory SSR waxed and waned during the stimulation interval, more so than in the somatosensory and visual cases.

Conclusion

Although this chapter only provides a descriptive overview of the outcome of the experiment, there are still some notable results. The T^2 analysis provides the ability to

enhance the detection of and better characterize ongoing responses where the low signal-to-noise ratio precludes analysis of time-domain averages. This is especially useful when the stimulus consists of brief trains of steady-state stimulation and the dynamics of the response across the train are of interest. In this stimulation paradigm (a variant of which is used in Experiment 3 to be reported in Chapter 4) a large number of discrete stimuli are experienced, but because they are organized into trains, the number of trials available for averaging is not large.

Another finding was that although the present results do not directly refute the hypothesis that the SSR might be produced by the linear summation of transients, it is clear from the simulations that the source transient must be produced at higher stimulus rates where the N1/P2 component is no longer visible, as this component is not a factor in the production of the SSR. In the next experiment I use this information to produce source transients using stimulation rates swept from 10 Hz to 50 Hz. The question addressed was whether linear summation of these source transients would give an adequate account of the frequency/amplitude characteristic over this frequency range.

Chapter 3

Experiment 2
LINEAR AND NON-LINEAR PROPERTIES OF THE AUDITORY STEADY-STATE RESPONSE INVESTIGATED BY TEMPORAL DECONVOLUTION

Introduction

Auditory stimulation at steady-state rates provides a number of advantages over transient stimulation. Although the response to each individual stimulus in the steady-state stimulus train is quite small, the stimuli are presented rapidly so that many responses can be collected in a short period, allowing for a much better signal to noise ratio in the resulting average than for transient responses collected in the same period. The SSR collected can then be further improved by filtering since the response must occur at the stimulation rate or harmonics of the stimulation rate. This is of course true for transient stimulation as well, but for 40 Hz steady state stimulation the response can be high pass filtered at nearly 40 Hz, removing almost all uncorrelated noise in the EEG that was not already removed by the averaging process and which typically has most of its energy concentrated below 20 Hz.

Once the SSR has been processed, the interpretation of the response is much simpler than the interpretation of transient responses. For 40 Hz stimulation, where the majority of the energy in the response is concentrated at 40 Hz with relatively small amounts at harmonics of this rate, the response can be simplified to a single amplitude and phase value for each electrode site. This ease in interpretation enables applications which involve rapid on-line processing of the collected EEG, such as monitoring patient awareness during anesthesia (Gilron, Plourde et al., 1998) Using steady-state stimulation, it is also possible to collect responses from several channels simultaneously when each channel is modulated at a different rate in the steady-state range (John, Lins et al., 1998). This allows, for example, several carrier frequencies to be applied simultaneously, speeding up data collection for tonotopic mapping. However, for many of the

advantages of steady-state stimulation to be realized, an acceptable account of its generation must be developed.

Most analyses of steady-state data assume it is generated by a single generator or two bilateral generators located in primary auditory cortex, and that these generators may be the same ones that produce components of the transient response. The prevailing theories for SSR generation were discussed in Chapter 1. The linear summation hypothesis of Galambos et al. (1981) proposed that the middle latency components of the transient, which are prominent at stimulation rates in the 6-10 Hz range, continue to be produced more or less identically at higher stimulation rates. The linear summation of these overlapping potentials is thought to completely explain the recorded SSR at rates in at least the 10-50 Hz range. However, various phenomena are left unexplained by this account, including the observation that the tonotopic organization of the magnetically recorded Pa and the SSR differ (Pantev, Roberts et al., 1996) and the observation that the superimposition of transient-induced middle latency responses (MLRs) fails to perfectly synthesize the SSR at rates other than 40 Hz (Santarelli and Conti, 1999).

The deconvolution technique developed by Gutschalk et al. (1999) and described in detail below allows for a more direct test of the linear summation hypothesis. By taking SSRs recorded at various stimulation rates (six rates between 32 and 53 Hz in their study) and applying a transformation to this data, a 'source transient' which best explains all of the SSRs is produced, given that linear summation of this source transient produced the observed SSRs. This source transient can then be used to simulate the observed SSRs again, allowing a determination of the goodness of fit of the model at the different rates. Gutschalk et al. reported that the simulation from a source generated in this way explained over 95% of the variance in the recorded SSRs for every subject. Dipole analysis of the source waveform showed two temporally overlapping

sources, one exhibiting a N19m-P30m-N40m pattern, and the other a N24m-P36m-N46m pattern. The source transient was quite similar to the ~10 Hz recorded middle latency response on an individual subject basis.

The experiment reported in this chapter uses a similar technique for the reconstruction of a 'source transient' with a number of significant modifications. Rather than concatenating SSRs recorded at an integral number of fixed rates, a stimulus with a continuously varying interstimulus interval is used, similar to the maximum-length sequence technique (Eysholdt and Schreiner, 1982; Picton, Champagne et al., 1992; Burkard, Shi et al., 1990a, b). The stimulus in the auditory case sounds something like a siren, with tone pips coming increasingly faster until a maximum is reached and then the process is reversed back to a minimum. This cycle is presented continuously to the subject for a period of time and the cycles are averaged. As the study is done in EEG rather than MEG, radial components of the evoked field can be studied.

Temporal Deconvolution

The linear summation model of the SSR assumes that if the transient response (that is, the response to a single isolated stimulus) were known, the observed steady-state response could be generated from the summation of overlapping transients occurring at the stimulus rate. The temporal deconvolution method constructs a 'source transient' from the measured steady-state data by applying the conditions of the linear summation model and some matrix algebra.

Assume for purposes of explanation that the transient response is known, and that stimuli occur at t = -8ms, t = -3ms, and t = 0ms with a 1ms sampling rate. Referring to Figure 3.1, I will

denote the steady-state response $s(t)$ and the transient response $b_n(t)$, with the subscript n used to

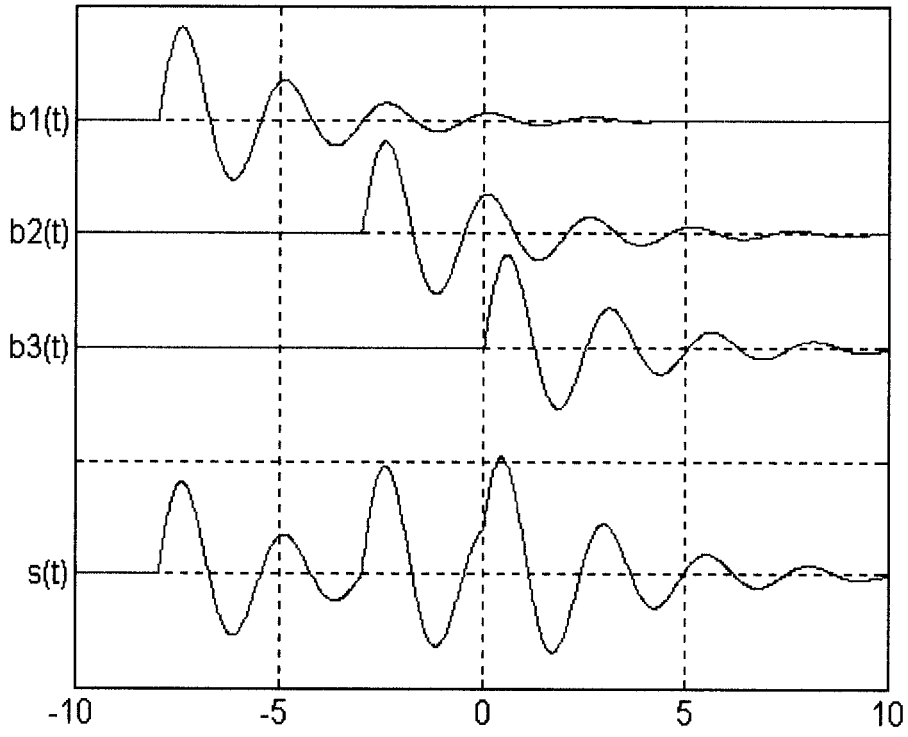


Figure 3.1 Demonstration of convolution of source transients into steady-state. The three identical 'source transients' b1(t), b2(t) and b3(t) occurring at t = -7, t = -3 and t = 0 respectively are summed algebraically into s(t); an identical waveform could be produced by the discrete convolution of the source waveform with an appropriately constructed modulation matrix. This example demonstrates a waveform produced by transients with a successively decreasing inter-stimulus interval, similar to the second half of the stimulus used in this study; a conventional steady-state is produced using a constant inter-stimulus interval.

index the transient. The figure shows the three overlapping identical transients above their summation. It can be seen that the point in the steady-state response at t=0 is the sum of the first, fourth and seventh point of the transient:

$$s(0) = b_1(7) + b_2(3) + b_3(0).$$

If the steady-state response *s(t)* is denoted as a matrix S where each row vector is the time domain response for one channel, and the transient response is denoted similarly by the matrix B (the transient response is different for each channel), the above convolution of transient responses can be written as a matrix multiplication

$$S = B\,M \qquad (1)$$

with M denoting a modulation matrix set up to perform the convolution above. The first column in M is set up to select the appropriate points from B for summation into the first point in S, the second column is set up to construct the second point in S, and so on. An example modulation matrix M from the stimulation shown in Figure 3.1 would have the following values:

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

This 12x12 matrix could reconstruct the 12 points in *s(t)* after t=0 from the 12 point transient response b(t). In general, an *n*-point steady state response could be constructed from an *m*-point transient response using an *n* x *m* modulation matrix, if the stimulation times and the transient

response are known. More importantly, with a given modulation matrix M the transient response B can be estimated by applying the pseudoinverse matrix M^{-1} to both sides of equation (1)

$$F = B M \qquad (2)$$

$$F M^{-1} = B M M^{-1}$$

$$F M^{-1} = B \qquad (3)$$

Thus the transient response can be estimated by postmultiplication of the observed steady-state response by the pseudoinverse of the modulation matrix. This process is referred to as deconvolution. If the linear summation model is correct then the calculated transient response (for convenience, called herein the "source transient" or ST) will be the best estimate of a transient that would linearly summate to produce the observed response. If this process is simulated in the absence of noise the deconvolution will be perfect, exactly regenerating the transient wave used to generate a simulated steady-state wave. Also note that the estimate of the transient, B, can be used with the modulation matrix M and equation (1) to regenerate an estimate of what the steady-state wave should look like if the linear summation hypothesis were true. This process is referred to as reconvolution or reconstruction. The extent to which the actual steady-state differs from the reconvoluted steady-state indicates how well the linear summation model fits the data.

The process of estimating the source transient by deconvolution and then reconstructing the steady-state response by reconvolution requires that the steady-state waveform be generated from a transient that overlaps by a number of different amounts; stimulation at a constant rate would not provide sufficient data for deconvolution. The stimulation paradigm used in this

experiment provides an ideal situation for testing the deconvolution process. The stimuli overlap from 100ms to ~16ms over the eight second sweep epoch, with overlap increasing in the first half of the epoch and decreasing in the second. A modulation matrix was created from the inter-stimulus intervals contained in this epoch, in order to reconstruct a 130 ms representation of the transient response. The process of calculating the pseudoinverse of this matrix needs only be performed once for each stimulation paradigm.

Materials and Methods

Subjects

Four male and six female subjects aged 19 to 39 participated. Three of the subjects were paid $20 each for their participation. One subject was a volunteer who participated without remuneration. The remaining subjects participated in order to receive course credit in Introductory Psychology. All subjects gave their written consent. Experimental procedures were in conformance with the Declaration of Helsinki and were approved by the University Ethics Committee.

Training Environment and Auditory Stimuli

All sessions were carried out in an electrically shielded room acoustically treated to reduce ambient noise. The auditory stimuli were made from white noise bursts, each burst being 10 ms long including a 2 ms $\cos^2$ rise and fall envelope. There were 3 stimulation conditions presented to each subject in the same order. Condition 1 (1.5 Hz transient) consisted of bursts

presented at a constant interval of 666ms (1.5 Hz) for 20 minutes giving a total of 1800 transient stimuli. Condition 2 (sweep) consisted of a 10-50-10 Hz sweep over 8 seconds (see Panel D of Figure 3.5 for a graphical representation of the instantaneous stimulation frequency over the 8 second sweep period), with 150 such sweeps being presented with no pause between successive sweeps. Condition 3 (SSR) consisted of noise bursts presented at a constant interval of 25 ms (40 Hz) for 5 minutes for a total of 12000 stimuli. All stimuli were generated by a Tucker Davis DA-2 converter, sampling rate 100kHz, and delivered through a speaker (Radio Shack 40-4064) placed 150cm in front of the subject approximately at eye level. Stimulus intensity was set to 76 dB SPL. Subjects were encouraged to read a newspaper or book during the stimulation, and were not instructed to pay any particular attention to the stimulus.

EEG Recording and Analysis of Data

The EEG was recorded from 32 electrodes placed according to the international 10-20 system (Quickcap, silver-chloride electrodes), referenced to CZ with a ground at AFZ. Infra-ocular leads IO1 and IO2 and mastoid leads M1 and M2 were applied using paste-on disposable leads (Kendall Q-Trace Gold 550 Ag/AgCl ECG electrodes cut to 1 cm x 2 cm). Scalp electrodes were filled with Electro-Gel and impedances were reduced to < 10000 Ω by abrasion with a blunt sterile needle before recording. The sites for the paste-on leads were prepared by rubbing with an alcohol impregnated cotton pad and impedances at these sites generally reached levels < 20000 ohms. The signals were recorded continuously through each stimulus run using a Neuroscan Synamps digitizing at 1 kHz and filtering DC to 200 Hz. The continuously recorded data were processed off-line.

In the following description of signal processing methods, it is convenient to discuss Condition 1 (1.5 Hz transient stimulation) and then Condition 3 (the 40 Hz SSR), followed by Condition 2 (10-50 Hz sweep).

Condition 1: Transient Stimulation Signal Processing

Prior to processing, the continuously collected data from the transient stimulation condition were epoched from 100 ms pre-stimulus to 400 ms post-stimulus onset The 1800 epochs were written to a binary file and the total variance (power) in each lead for each epoch was calculated. A threshold for total variance was then calculated such that 1440 epochs (80% of collected epochs) with no lead exceeding the variance threshold were available (John, Dimitrijevic et al., 2001). These accepted epochs were subject to further analysis and the remaining artifact contaminated epochs were rejected from further analysis.

An average of the accepted trials was calculated from the 1440 epochs for each subject. The group average (n=11) of these data was exported to BESA for source analysis. Two sources, one in each hemisphere, were modeled by procedures described later in the Method Section. These sources were subsequently used as spatial filters to generate source waveforms in all stimulus conditions (1-3).

In Condition 1, three waveforms were derived for each subject. These were (1) the transient response at the FZ lead with respect to a common average reference, (2) the global field power calculated also from the common average referenced data, and (3) two source waveforms, one for each hemisphere, using spatial filters calculated from the dipole analysis of the group average of this data set. The MLR complex was extracted from these waveforms after filtering using a 20-150 Hz band-pass zero-phase filter.

Condition 3: 40 Hz SSR Signal Processing

Prior to processing, the continuously collected 40 Hz data were epoched from 75ms pre-stimulus to 100 ms post-stimulus onset. Epochs therefore included seven full cycles of 40 Hz. The 12000 epochs were written to a binary file and the total variance (power) in each lead for each epoch was calculated. A threshold for total variance was then calculated such that 7500 epochs with no lead exceeding the variance threshold were available; these accepted epochs were subject to further analysis and the remaining artifact contaminated epochs were rejected from further analysis. An average was then created from the 7500 accepted epochs for each subject and re-referenced to a common average reference.

Two source waveforms were created from the averages for each subject by using the dipole forward solutions performed on the 1.5 Hz data as spatial filters, one source waveform for the source in each hemisphere. The 175 point waveform spanning the 175ms epoch was multiplied by a Hamming window and an FFT was calculated giving the phase and amplitude of the 40 Hz component.

Condition 2: 10-50 Hz Sweep Signal Processing

In the sweep condition, the stimulus takes 8 seconds to go through a complete cycle from 10 Hz repetition rate up to 50 Hz and back to 10 Hz. An averaged response of the entire period is required for the deconvolution process but the standard procedure of epoching and artifact rejection would not work in this case as the likelihood of finding even a few 8 second epochs free from stimulus artifact is small. In order to circumvent the problem, the 8 second long average was reconstructed from a series of shorter 100 ms averages which overlapped each other

by 50% in time. Each short average was then multiplied by a 100 point triangular (bartlett) window and then added to the 8 second "long" average after time shifting by the appropriate amount. This window function ensured an overall unity gain for the overlap-add long average. Each shorter average was made by averaging together the 120 epochs (80% of available epochs) with the lowest variance, so that each of the shorter averages was constructed from an identical number of trials. The long average was then filtered using a 20-150 Hz band-pass zero-phase filter. From this average spatially filtered source waveforms (one for each hemisphere) were generated for each subject in the same manner as for the SSR.

Temporal Deconvolution

The temporal deconvolution procedure was applied to the sweep data (Condition 2). A deconvolution matrix M was constructed according to the method outlined in Gutschalk et al (1998) in order to create a source transient assuming a zero response before the stimulus and after 130 ms. The pseudoinverse of this matrix M^{-1} was calculated in Matlab using Singular Value Decomposition. Using this matrix, a 33 channel 'source transient' (ST) was calculated from the long average as well as source transients for the spatially filtered waveforms. Each of these source waveforms was used to reconstruct the original long average under the assumption that linear superposition of this source transient could explain the original waveform. The adequacy of the temporal deconvolution model was assessed as a function of the stimulation rate by determining the residual variance of the model by squaring the difference between the reconstructed waveform and the actual waveform. The source transient was also used to reconstruct a 40 Hz SSR to compare the amplitude and phase of the actual SSR recorded in Condition 3 with the SSR predicted by the source transient.

Source Analysis

The group average (n=11) of the condition 1 (1.5 Hz transient) response was exported to BESA for source analysis. This waveform was filtered 20-150Hz to enhance the MLR components. The digitized electrode positions from one of the subjects was used to provide the electrode coordinates for BESA. A window from 20-50 ms post stimulus onset was used for the fitting procedure. Two single dipoles were placed in opposite hemispheres, with one forced to be symmetric to the other and both were allowed to vary freely in orientation. The inverse solution from this fit was exported from BESA and used to generate source waveforms from the various recorded waveforms.

Results

The group average for the 1.5 Hz condition 1 transient response (called herein the MLR), filtered 20-150 Hz, is shown in Figure 3.2 Panel A for all electrodes. This waveform shows negative going peaks in the frontal leads at 16, 39 and 64 ms (39 ms peak is labeled), and positive going peaks at 29 and 55 ms (both labeled). The positive peaks match the typical timings for the Pa, and Pb, and the first two negative peaks match the Na and the Nb (Picton, Hillyard et al., 1974). A slight additional delay (for example the Pa is usually cited as a 25 ms peak whereas the peak here occurs at 29 ms) might be due to the 2 ms rise time of the stimulus. The corresponding peaks can also be seen in the global field power waveform which is presented

in Panel D. A map of the field distribution for the 29 ms peak (Figure 3.2 Panel C) shows a pattern consistent with bilateral sources in auditory cortex.

The dipole fit on the MLR waveform from Panel A was performed across the 20-55 ms window encompassing both the Pa and the Pb, using one dipole in each hemisphere forced to be symmetric but allowed to vary independently in orientation. The localization results are shown in Figure 3.2 (Panel F) with goodness-of-fit (Panel E) and GFP (Panel D). The best fit dipole locations were consistent with sources in auditory cortex. The best goodness-of-fit (GOF) of 92.8% occurred at 30 ms, but the other peaks in the 16-80 ms range also had GOF values exceeding 80%, indicating that these response peaks were also represented by sources in these locations. The inverse solution for this fit was exported from BESA and used to spatially filter the data collected in all conditions. This had the function of performing a data reduction, allowing the 32 channel data set to be reduced to 2 channels.

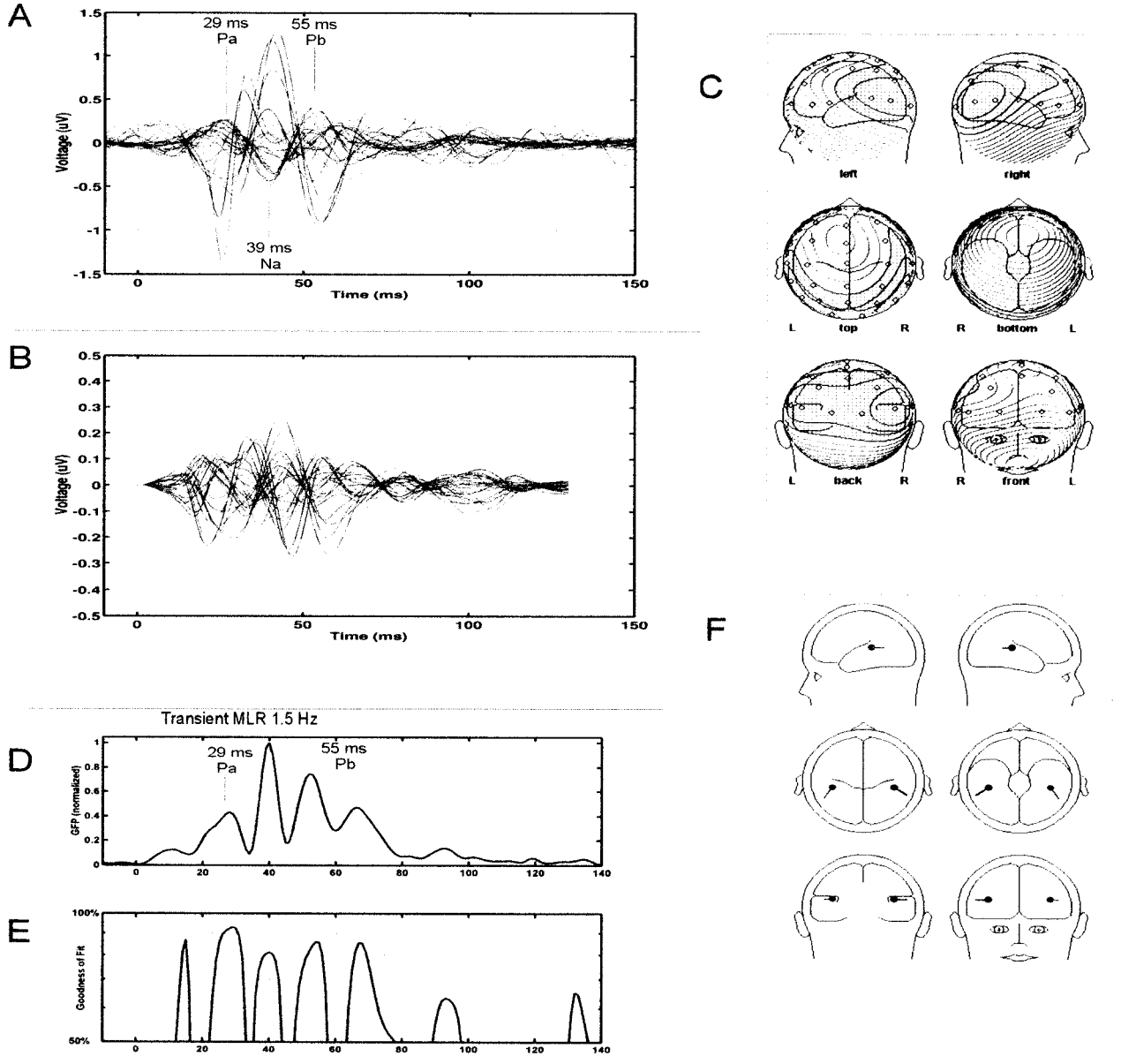


Figure 3.2 Panel A shows the group average (n=11) of the response to the condition 1 (transient) stimulation, for all 33 leads using a common average reference band-pass filtered 20-150 Hz to accentuate the middle latency components. Panel B shows the group average (n=11) of the deconvolved 'source transient' for all 33 leads using a common average reference band-pass filtered 20-150 Hz. Note the scale change between panels, the source transient has a smaller amplitude. Panel C shows the field pattern at the 29 ms point for the Panel A waveform. Results of dipole source localization procedure, fit across the 20-50 ms time window of the condition 1 transient. Left and right hemisphere dipoles were forced to be symmetric, with orientation allowed to vary individually for each dipole. Panel A shows the source locations. The global field power for the condition 1 transient is shown in panel B. The goodness of fit (GOF, log scale) is shown in panel C; the maximum GOF of 92.8% occurred at 29 ms.

The eight second long averages from the sweep stimulus (condition 2) were deconvolved to produce a 'source transient' (ST) waveform for each of the 11 subjects. The 32 channel representation of the group average of the STs is shown in Figure 3.2 Panel B. This pattern is superficially similar to the MLR but smaller in amplitude. The GOF for this waveform using the model generated from the MLR source fit was 91.8% at the 30 ms point, indicating similar generators for the EEG data of Figure 3.2A and the reconstructed MLR from the sweep data (Figure 3.2B).

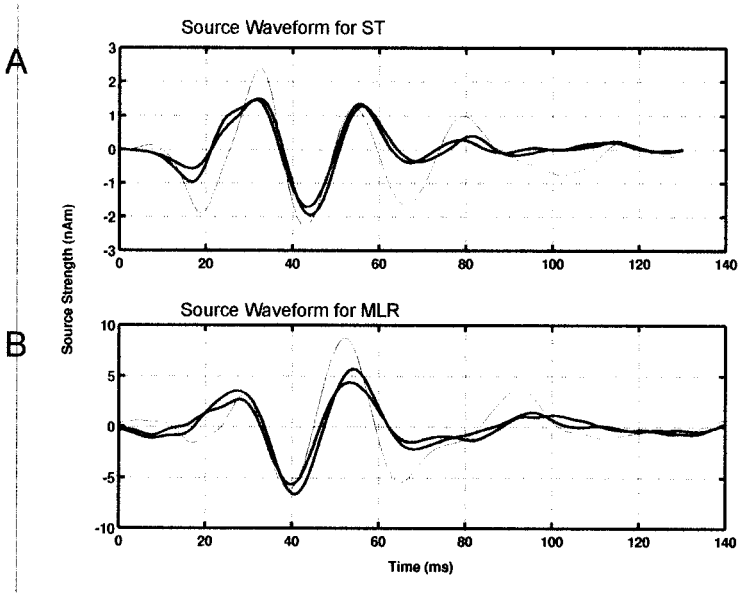


Figure 3.3 Panel B shows the source waveforms for the left hemisphere (blue trace) and right (green trace) sources fit to the 20-50 ms time window of the condition 1 group average. A negative peak at 40 ms and positive peaks at 28 ms and 55 ms are observed. In panel A the same spatial filter is used to create source waveforms from the 'source transient' group average. The amplitude of the 'source transient' is smaller (note scale change) and the first positive peak appears later at 32 ms and the negative peak at 43 ms also appears delayed from the condition 1 wave. The final positive peak at 55 ms appears aligned with the condition 1 waveform. An additional negative peak at 19 ms appears in this waveform. The red trace is the FZ waveform with respect to a common average reference, arbitrarily rescaled to show the relation between the source waveform and a raw trace.

Using the source fit as a spatial filter produced the source waveforms shown in Figure 3.3, with the upper panel showing the ST source waveforms and the lower panel showing the MLR source waveforms, separately for each hemisphere. The red trace is the EEG at Fz rescaled to show the similarity of the source waveforms to the original recording. There is little evidence for hemispheric differences in either set of source waveforms. The deflections were P32, N44 and P56 for the ST sources and P27, N40 and P55 for the MLR sources. The comparison

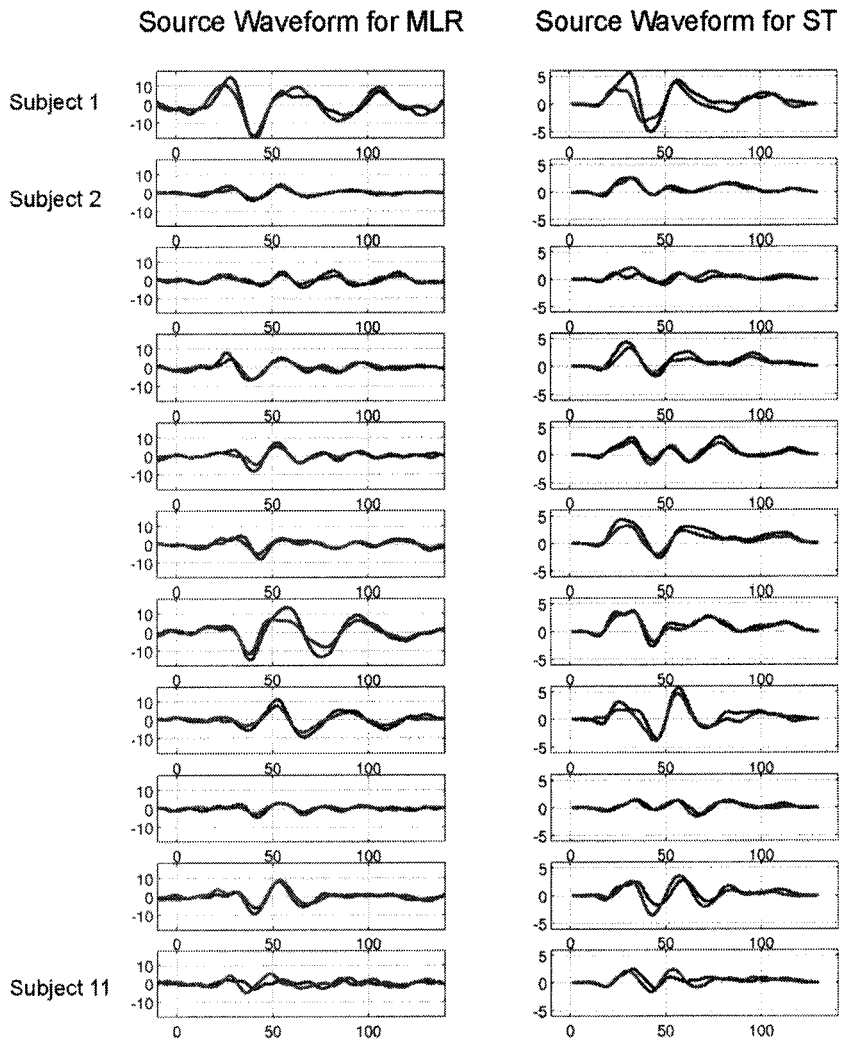


Figure 3.4 A comparison similar to Figure 3.3 is shown for each individual subject. The blue and green traces show the left and right hemisphere source waveforms respectively. The left panel of each pair shows the condition 1 source waveforms in the left panel and the source transient source waveforms in the right panel.

between the ST source waveforms and the MLR source waveforms for each individual subject is shown in Figure 3.4. There was qualitatively good agreement between the two source waveforms for each subject, although the amplitude of the waveforms was variable between subjects.

Figure 3.5 demonstrates the result of the reconstruction operation on the eight second long averages of condition 2 (10-50 Hz sweep). Panel A shows the group average (n=11) of the recorded sweep data filtered 20-150 Hz and spatially filtered, with the instantaneous stimulation frequency shown in Panel D for reference. A slight amplitude enhancement in the 2-3 second and 5-6 second time period corresponds to a stimulation rate of 30-45 Hz. In Panel B, the group average ST is used to reconstruct the same waveform by multiplication of the ST by the modulation matrix. This reconstruction also shows a clear amplitude enhancement in the 30-45 Hz area, demonstrating how the linear superimposition of the ST can produce the 40 Hz amplitude effect. The agreement between the reconstruction in Panel B and the recording in Panel A was calculated individually for each subject by squaring the difference between the recorded and reconstructed waveforms; the group average of these error waveforms is shown in Panel C of Figure 3.5. No systematic relation between stimulation rate and error is evident, showing that the linear superimposition model is equally effective at all rates from 10-50 Hz.

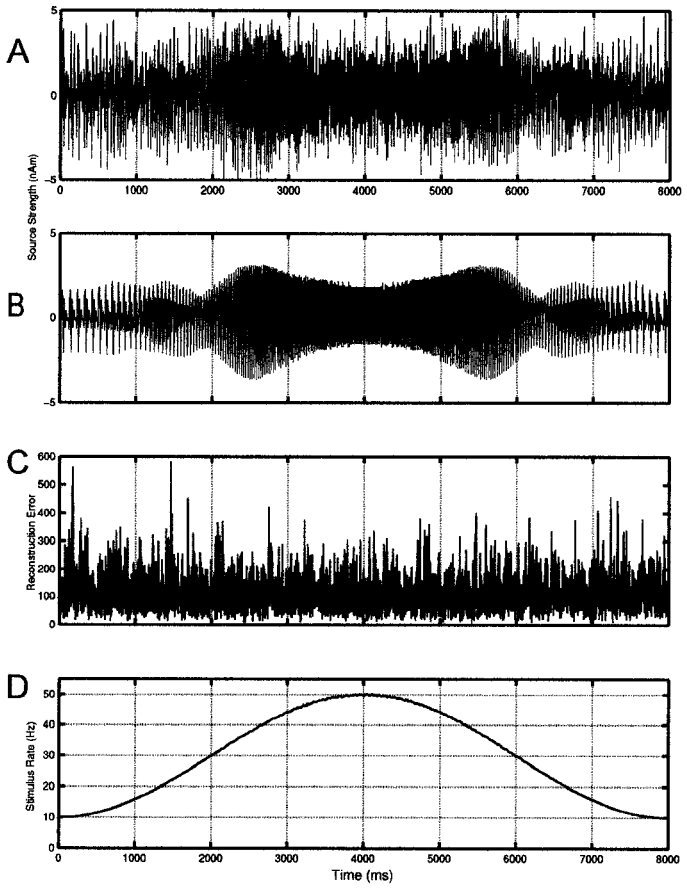


Figure 3.5 Panel A shows the source waveforms generated from the group average (n=11) for the condition 2 (10-50 sweep) stimulus over the entire 8 second sweep period. Panel B shows the reconstructed version of the panel A waveform generated by the convolution of the group average source transient. Amplitude enhancement across the 2-3 second and 5-6 second time interval, corresponding to 30-45 Hz stimulation, is evident. For each individual subject, the difference between the recorded and reconstructed waveforms (squared) was averaged together to show the reconstruction error across the sweep period (panel C). There appears to be no systematic variation in error with stimulation frequency. In panels A-C, the blue and green traces show the left and right hemisphere source waveforms respectively. Panel D shows the instantaneous stimulation rate across the 8 second sweep period.

The effectiveness of the superimposition model at 40 Hz is shown in Figure 3.6. For each subject a segment of 40 Hz response was synthesized from the subject's ST created by the deconvolution of his or her 10-50 Hz sweep response. The synthesis was accomplished by adding overlapping STs spaced at 25 ms intervals in a fashion similar to that shown in Figure

3.1. A 175 ms segment of this signal was extracted from the stabilized synthesized response, windowed (hanning window), and analyzed with a FFT. The phase and the amplitude of the fundamental at 40 Hz was recorded for each subject. The same FFT analysis was performed on the 175 ms windowed segment of recorded 40 Hz response from the condition 3 average (40-Hz continuous SSR). Figure 3.6 Panel A shows the relationship between the synthesized and recorded amplitudes; the linear relation R=0.8966 was fit by the least-squares line

```
SYNTH_AMPL = 0.9945*(RECORDED_AMPLITUDE) + 7.77
```

indicating a direct relation between the actual and synthesized amplitudes. For the phase (in radians) of the fundamental, shown in Panel B, r=0.8909,

```
SYNTH_PHASE = 0.8645*(RECORDED_PHASE) -0.3552,
```

again indicating a direct relation although not as strongly as for the amplitude.

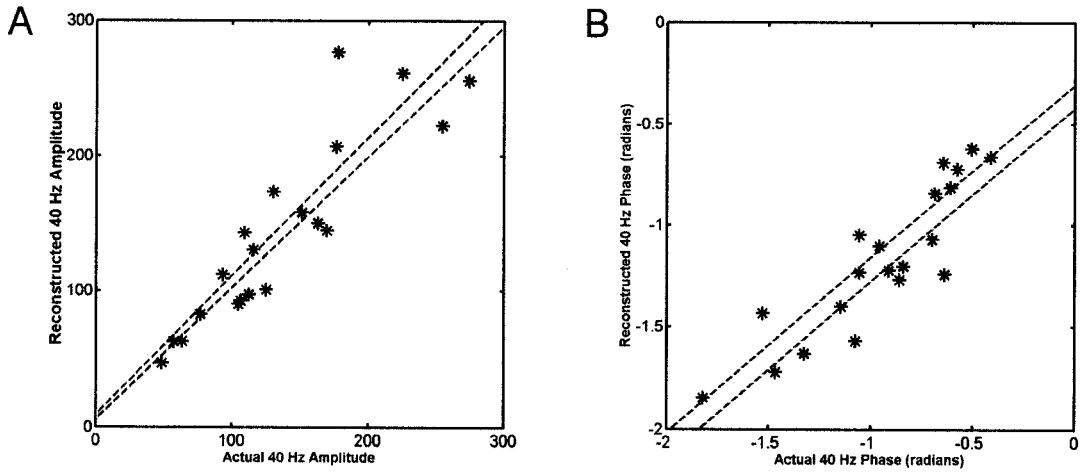


Figure 3.6 For each individual subject, the source transient was used to synthesize the 40 Hz steady-state response source waveforms and the amplitude and phase of the fundamental (40 Hz) FFT component was recorded. A similar FFT analysis was performed on the actual recorded steady-state response (condition 3). The slope of the least-squares fit line (0.99 for amplitude and 0.86 for phase) indicates a direct correspondence of actual and reconstructed 40 Hz response.

Discussion

The overall results are consistent with the linear superimposition hypothesis. First, the source transient, which is a sort of 'average' of the 240 separate transients occurring across the 8 second sweep, shows no systematic error in reconstructing the sweep result. This indicates that no particular range of sweep rate is produced by a larger or smaller source transient than any other range. Discrepancies between the reconstructed and actual sweep at particular frequencies are predicted by the idea that the 40 Hz amplitude enhancement is due to a resonance phenomenon, but no such discrepancies were found. In fact the linear reconstruction shows an obviously enhanced response in the 30-45 Hz range that must be due to linear summation.

Second, a source transient deconvolved from the sweep data which were collected with a continuously varying interstimulus interval successfully predicted the amplitude and phase of the 40 Hz response generated by a stimulus presented at a constant interstimulus interval. This shows that a gradually building resonant response is not necessary to generate the 40 Hz amplitude peak. In fact, the increase in amplitude of the response when the stimulus is sweeping rapidly through the 30-45 Hz region can be seen in both the recorded and reconstructed sweep waveforms, even though resonant brain activity could not have contributed to the reconstructed sweep waveform.

It is noteworthy that the waveforms of the source transient and the MLR appeared to be similar. Each showed a similar pattern of positive/negative/positive deflections at ~30, 42 and 55 ms; the slight delay from the usual Pa/Pb times of 25/50 ms is probably due to the rise time of the stimulus which was a noise burst rather than a click. These similarities would be expected by a linear summation hypothesis. However, there was an amplitude difference between the source

transient and the MLR, suggesting that the MLR components are still partially refractory at the 10 Hz plus rate. Alternatively, the auditory gamma band response (Pantev, Makeig et al., 1991), a transient 40 Hz oscillatory response which occurs at low stimulation rates, may augment the MLR at these rates.

The success of the linear superimposition model in describing the behavior of the SSR over the 10-50 Hz stimulus frequency range does not imply that the components of the MLR waveform arise from the same neural generators. There is evidence that components of the MLR waveform can be functionally dissociated. For example, the 30 ms Pa component has been reported to be less sensitive to rate (Erwin and Buchwald, 1986b) than the 50 ms Pb, although in our study both the 32 and 55 ms positivities (which seem to correspond to these two components) were much smaller than their transient counterparts when evoked at 40 Hz. It is worth noting that the early positivity in the ST was reduced to 41% of its amplitude in the MLR, whereas the late positivity in the ST was reduced to 22% of its MLR amplitude, which is consistent with the idea that the earlier component is less sensitive to rate. The Pb also disappears during sleep whereas the Pa does not (Erwin and Buchwald, 1986a). Because the source transient components deconvolved from the sweep data in our study mirrored the behavior of the different MLR components, studies that treat the SSR as a unitary component explained by a single source generator may be in error. The MEG deconvolution study of Gutschalk, Mase, et al. (1999) supports this interpretation, as they found that the source transient was modeled by two equivalent current dipoles in each hemisphere with overlapping activation periods. Studies that attempt to manipulate the SSR in various ways might benefit from the ability to separate the two components composing it, at the cost of a more complex stimulus consisting of at least several stimulation rates or a continuously changing ISI as in this study, and

also the requirement for a more complex analysis procedure. Online real-time analyses of the SSR probably would not benefit from this added complexity, nor would auditory threshold determinations.

One limitation of the linear superimposition model not addressed in the present study concerns its ability to predict the start-up of the 40-Hz SSR. Linear summation would predict a stable SSR waveform after about 4 stimuli (100 ms), since the source transient returns to baseline after this interval. However, the amplitude and phase of the SSR are known to take up to 300 ms to stabilize at their respective asymptotes after the onset of a train of stimuli (this finding confirmed by Figure 4.4 in the next Chapter). It is possible that one or more of the MLR components may be desynchronized or suppressed by the low-frequency transient (N1/P2) occurring during this interval, reappearing when low-frequency transient activity has subsided. Alternatively, Experiment 1 found that properties of the transient response changed at frequencies above about 4 Hz. This is range at which membrane dynamics change, including lengthening of postsynaptic potentials that allow temporal integration of inputs (Buonomano and Merzenich, 1998b). Generators of the MRL may be recruited into a network by these dynamics. If so, the present findings indicate that this behavior appears to remains stable over the range of stimulus frequencies from 10-50 Hz.

In the present study we modeled the sources of the MLR waveform with single equivalent dipoles in each hemisphere. Least squares fits localized these sources to the auditory cortex in broad agreement with intracortical recordings in surgical patients (Godey, Schwartz et al., 2001) and magnetic source imaging studies (Yvert, Crouzeix et al., 2001) which have localized generators for the Pa wave and the 40-Hz SSR, respectively, to region of Heschl's gyrus (the primary auditory cortex, or auditory core in the model of Hackett et al.). Although the spatial

resolution of our EEG recordings is not sufficient for accurate co-registration on neuroanatomy measured by MRI, the spatial resolution of EEG is sufficient to distinguish the cortical sources of the SSR from those of the low-frequency transient components such as the N100 (Pantev, Elbert et al., 1993) and P2 (Pantev, Oostenveld et al., 1998; Shahin, Bosnyak et al., in press) which likely reflect activity in secondary (belt/parabelt) auditory cortex (Pantev, Bertrand et al., 1995). The study presented in the next chapter makes use of this property to investigate the dynamics of the SSR when functional activity in the auditory cortex is remodeled by experience. SSRs evoked by brief trains of steady-state stimulation are distinguished from low-frequency transient responses by filtering, allowing the effect of training on an auditory discrimination task to be studied in multiple cortical areas simultaneously.

Chapter 4

Experiment 3
DISTRIBUTED AUDITORY CORTICAL REPRESENTATIONS ARE MODIFIED BY TRAINING AT PITCH DISCRIMINATION WITH 40-HZ AMPLITUDE MODULATED TONES

Editorial Note:
The experiment reported in this chapter has been submitted for publication and is under editorial review. References, Figures, and Figure Captions are appended at the end of the chapter, with guidelines for figure insertion included in the text.

Distributed Auditory Cortical Representations are Modified by Training at Pitch Discrimination with 40-Hz Amplitude Modulated Tones

Daniel J. Bosnyak[1], Robert A. Eaton[2], and Larry E. Roberts[1]
[1]Department of Psychology and [2]Department of Physics and Astronomy, McMaster University, Hamilton, Ontario, Canada L8S 4K1

We trained nonmusician subjects to discriminate small increases in pitch from a 2-kHz standard stimulus using 40-hz amplitude modulated pure tones. This method permitted separation of transient N1, N1c, and P2 auditory evoked potentials whose sources are known to localize to the secondary auditory cortex (AII) from the 40-Hz steady state response (SSR) whose sources localize to primary auditory cortex (AI). Training enhanced the amplitude of N1c (right hemisphere, latency 155 ms) and P2 (latency 172 ms) responses evoked by the trained stimuli. N1 amplitude did not change although its latency (107 ms) decreased. Bivariate t-statistics in a moving window revealed a patchy modulation of the SSR that commenced near the P2 and shortened in phase after training (amplitude unchanged). No 40-Hz activity was detected in control subjects when P2 responses were evoked by unmodulated tones, indicating that the SSR was a separate brain event. Source localizations of the N1, N1c, P2, and SSR were spatially differentiable and in agreement with previous findings. These results suggest that training at pitch discrimination enhanced cortical representations for the trained stimuli in distributed regions of AII including the right hemispheric N1c which may reflect neurons specialized for processing of spectral pitch. Neurons representing the trained frequencies in AI did not appear to increase substantially in number (SSR) although temporal properties of the representation were modified. Enhanced P2 and N1c responses may reflect new tunings on AII neurons whose establishment and expression are gated by converging input from other regions of the brain.

INTRODUCTION

It is now well established that the frequency tuning of neurons in the mammalian auditory cortex is not hardwired after early development but can be altered in the adult brain by experience with behaviorally significant acoustic signals (Buonomano and Merzenich, 1998). Plastic modification induced by aversive conditioning in adult guinea pigs has been documented for neurons in primary (AI) and secondary (AII) regions of the auditory cortex as well as in the medial, dorsal, and ventral divisions of the auditory thalamus (Edeline, 1999). When brain regions are contrasted within the same training procedure, tone-evoked plasticity is expressed more commonly by neurons in AII (96%) than by neurons in AI (63%; Diamond and Weinberger, 1984). A primate study by Recanzone et al. (1993) found that discrimination training for small changes in spectral pitch enhanced the cortical territory representing the trained frequencies in AI of owl monkeys by a factor exceeding 5. The sharpness of tuning and temporal response properties of multiunit recordings were also modified for the trained frequencies in this study. In the somatosensory system, decreased variability in the temporal response properties of neurons has been reported to be the best predictor of discriminative performance (Recanzone et al., 1992).

Neural plasticity of the magnitude seen in these animal studies suggests that remodeling of the human auditory cortex by behavioral training should be expressed in auditory evoked potentials (AEPs) which reflect the activities of networks of neurons in the brain. Among the earliest AEP components expressed in the auditory cortex are N19 and P30 "middle latency" responses which have been localized by source modeling (Godey et al., 2001; Scherg and von Cramon, 1986; Yvert et al., 2001) and by intracortical measurements (Celesia, 1976; Godey et al, 2001; Liégeois-Chauvel et al., 1993) to the tonotopically organized anteromedial region of Heschl's gyrus (AI) and are abolished by ischemic infarcts invading this region (Scherg and von Cramon, 1986). Because the N19-P30 waveform has a wave period near 25 ms, it survives signal averaging and is expressed in "steady state responses" (SSRs) when acoustic stimuli are amplitude modulated near 40 Hz (Galambos et al., 1981; Gutschalk et al., 1999; Hari et al., 1989). Neuromagnetic localizations of the cortical sources of the 40-Hz SSR appear to overlap those of the N19-P30 waveform in Heschl's gyrus (Engelien et al., 2000; Gutschalk et al., 1999; Pantev et al., 1996a; Schneider et al., 2002) although evidence for a contribution of nonlinear processes to SSR generation exists and this

question is currently debated (Conti et al., 1999; Ross et al., 2002; Santarelli et al., 1999). In contrast, the cortical generators of N1 and P2 "transient" responses of the AEP (with latencies of ~100 ms and ~180 ms, respectively) are spatially differentiable from those of the 40-Hz SSR and reflect activations centered in the belt and parabelt regions of AII based on magnetic (Pantev et al., 1993; Engelien et al., 2000; Houtilainen et al., 1998) and electrical (Pantev et al., 1995) source imaging and on intracortical mapping (Godey et al., 2001; Liégeois-Chauvel et al., 1993; Scherg & von Cramon, 1986). N1 and P2 sources are also differentiable from one another, with P2 sources centering anterior and medial with respect to those of the N1 which localize to region of the planum temporale (Hari et al., 1987; Joutsiniemi et al., 1989; Pantev et al., 1996b; Shahin et al, 2002). Because transient N1, P2, and 40-Hz SSR responses appear to tap neural representations in different regions of the auditory cortex, it should be possible to use these responses to investigate how neural activity in these regions is affected when the tuning of auditory neurons is modified by behavioral training.

The experiment reported here followed this approach. We trained nonmusician subjects to discriminate small increases in pitch from a standard stimulus of 2 kHz using 40-Hz amplitude modulated pure tones. This stimulus procedure (Eaton and Roberts, 1999) allowed us to separate N1, P2, and other transient components of the AEP whose sources localize to AII from the 40-Hz SSR whose cortical sources reside more specifically in AI. Our goal was to begin to describe the network behavior that underlies remodeling of human auditory cortex by experience.

MATERIALS AND METHODS

Subjects

Eight subjects (six males) aged 25 to 30 years participated in 18 sessions of discrimination training and testing. All were graduate students at McMaster University (5 right-handed). None had received formal musical training or played a musical instrument. Subjects were paid $150 for their participation. Subjects gave their written consent following procedures approved by the university ethics committee in conformance with the Declaration of Helsinki.

Training Environment and Auditory Stimuli

All sessions were carried out in an electrically shielded and acoustically dampened room. Auditory stimuli consisted of 10 ms sinusoidal tone pips (onset and offset windowed with a 2 ms cosine^2 function) of different carrier frequencies presented at 40 Hz for a duration of 1 sec. For convenience, we will refer to these stimuli as 40-Hz amplitude modulated (AM) pure tones (see Figure 1A for the spectrum and time domain waveform of the stimulus at 2 kHz). The stimuli were generated by a Tucker Davis sound generator and delivered through Noisebuster stereo headphones (Noise Cancellation Technologies Inc., Model NB-EX) which attenuated background noise by about 15 dB at frequencies below 500 Hz. The intensity of individual 1-sec pulse trains was varied randomly between 57-60 dB above each subject's measured threshold throughout discrimination training and testing in order to ensure that subjects used pitch and not intensity as the basis for their discriminative choices.

Figure 1 about here

Experimental procedure

Subjects participated in 18 experimental sessions of the following types which were given over approximately 20 days in the order indicated below.

(1) Preliminary Session. The experiment commenced with a preliminary session in which auditory thresholds and frequency discrimination ability were measured using 40-Hz AM tones. Hearing thresholds were determined at 2 kHz for each subject using six cycles of a staircase procedure. Frequency discrimination ability was then evaluated using the staircase method described by Levitt (1971). On each trial, a standard stimulus (S1) and a comparison stimulus (S2) were presented separated by an interval of 0.5 s. Subjects indicated by a button press whether the two tones were of the same frequency (50% of the trials) or different frequencies. Subjects were not informed of the correctness of their decisions. Forty trials were presented for each of 9 S1 frequencies between 1.8 kHz and 2.2 kHz in a single test that lasted approximately 25 minutes. S2 frequencies differed initially from the S1 by 60 Hz and were adjusted up or down according to the subject's performance. The purpose of discrimination assessment was to select the S2 stimuli that would be used for subsequent discrimination testing and training for each subject.

The preliminary session also familiarized subjects with 40-Hz AM tones they would encounter during test and training sessions.

(2) Test Sessions. Two test sessions were administered, one given the day following the preliminary session and the second about 18 days later after the training series (see below) had been completed. Test sessions provided a fine-grained assessment of discrimination ability before and after discrimination training. Each test session consisted of three blocks each containing 360 trials requiring same/different frequency judgments without knowledge of results. The three blocks differed with regard to the set of stimuli used. In one block the standard stimulus (S1) was 2.0 kHz while the comparison frequencies (S2) varied from 2.0 to 2.1 kHz. Because this stimulus set was used later for discrimination training, we refer to it as the "trained set". The remaining blocks evaluated "control" stimulus sets which employed either 1.8 kHz or 2.2 kHz as the S1 stimulus (S2 stimuli were 0-100 Hz higher). The order of assignment of stimulus sets to the three blocks varied between subjects but was the same for each subject before and after training. Control sets allowed us to determine whether changes in behavioral performance and brain activity detected after training related specifically to the trained carrier frequency, which would be expected if discriminative learning had occurred.

For each test block the procedure was as follows (similar to the SDH procedure described by Jesteadt and Bilge, 1974). On each trial subjects listened to the S1 stimulus followed 0.5 s later by an S2 stimulus. The S2 was either the same frequency as the S1 stimulus (50% of the trials) or one of 6 different comparison tones which were higher in frequency than S1. The lowest of the 6 different comparison tones was always 2 Hz higher than the standard, while the highest comparison tone was usually 60 Hz higher than the standard. The four comparison stimuli between these two extremes were chosen by the experimenter on an individual basis such that subjects were likely to detect two of them at least 50 percent of the time and the other two less than 50 percent of the time. On each S1/S2 trial the subjects indicated "same" or "different" by a button press; the next trial commenced 1 s later. Subjects were instructed to base their response choices on a change in pitch and not stimulus intensity. Test sessions lasted about 90 minutes (30 minutes for each block evaluating a single stimulus set) and used identical comparison frequencies before and after training so that performance between the test sessions could be compared.

(3) Training Sessions. The procedure for the training sessions was identical to the test sessions, except for the following differences. Only the stimulus set with the 2.0 kHz S1 was trained. In addition, feedback was given about the correctness of discriminative decisions by two LEDs placed 1 m in front of the subject at eye level. If the subject's response was correct, a green LED was illuminated; if the response was incorrect, a red LED lit up. The LED stayed on for 500 ms. Each training session contained 480 trials and lasted about one half hour. The S2 stimulus was 2.0 kHz on 240 trials ("same") and one of six 6 higher frequencies on the other 240 trials ("different"). Subjects received a total of 15 daily training sessions with a one day pause on the weekend.

An adaptive procedure was applied between the training sessions when the subjects performed without error on more than one comparison frequency. When this happened the highest comparison frequency was removed and replaced with a new comparison frequency in the region where the probability of a correct detection was near 0.5. The comparison frequencies of 2 Hz and 60 Hz were excepted from this procedure and kept constant for all subjects.

(4) Retention Session. The retention session took place 7 weeks after the second test session. The procedure for the retention session was identical to that of a training session. Two of the eight subjects were not available for the retention session.

Analysis of Behavioral Data

Behavioral performance was evaluated at each comparison frequency for each subject. At each comparison frequency which differed from S1 ($\Delta f>0$) the probability of a "hit" [P(H)] was calculated by dividing the number of "different" responses (hits) by the number of stimulus presentations. Next, the probability of a "false alarm" [P(FA)] was calculated as the proportion of trials on which the subjects responded "different" when the S2 frequency equaled the S1 frequency of 2 kHz ($\Delta f=0$). From these two measures a performance score (**P**) was calculated for each comparison frequency according to the formula $\mathbf{P} = [P(H) - P(FA)] / 1 - P(FA)$. The measure **P** corrects P(H) at each comparison frequency for the tendency of the subject to commit false alarms and reaches 1.0 (the maximum value attainable) at Δfs where the subject makes no errors (Green and Swets, 1966). Psychophysical functions were constructed for each subject, test session, and stimulus set by plotting **P** against all $\Delta fs>0$ and fitting the curve with a logistic. The discrimination threshold was defined as the value

of Δf corresponding to $\mathbf{P} = 0.5$.

The discrimination performance of each subject was also evaluated by *d'* using the procedure of Dember and Warm (1979). This metric was calculated for each comparison tone ($\Delta f>0$) by subtracting a z-score calculated for P(H) from a z-score calculated for P(FA). Values of *d'* were calculated for training sessions 1-3 grouped together and for training sessions 13-15 grouped together using comparison tones common to both sessions.

Electrophysiological Recording

The electroencephalogram (EEG) was recorded during the two test sessions and during training sessions 3 and 13. For the first two subjects a 19-channel recording was taken (10/20 system, Electrocap, tin electrodes) and for the remaining 6 subjects a 64-channel recording was made (NeuroMedical QuickCap, Ag/AgCl electrodes). Electrode sites were abraded with a blunt sterile needle and covered with Electro-Gel to lower skin impedance to <10 kOhm. The EEG was sampled at 500 Hz with a DC amplifier (NeuroScan Synamps) and recorded using Cz as the reference electrode and AFz as ground. Data were re-referenced off line to a common average prior to signal processing.

Analysis of EEG Data.

EEG data were epoched from 400 ms before stimulus onset to 400 ms after stimulus offset. Epochs were baselined for the interval 50 ms prior to stimulus onset and were linear detrended. Individual epochs were passed through a 60 Hz notch filter and sorted in order of total variance (energy) for artifact rejection. Epochs with the largest variance were rejected until 80% of the trials remained, following the procedure of John et al. (2001). When a trial was rejected, the data of the entire epoch were discarded.

EEG responses to the S1 stimulus only were analyzed. These stimuli were processed by the subjects in attention, were presented most frequently in the experiment (360 presentations for the S1s of each stimulus set in the test sessions and 480 presentations of the trained S1 during training sessions), and were uncontaminated by the preparation of behavioral responses. Transient responses of the EEG were analyzed using the 19 electrodes (10/20 array) that were common to every subject (n=8) and recording. Source modeling of the transient responses and analyses of the SSR in 100 ms moving windows (see below) were carried out for the 6 subjects for whom 64 channel recordings were taken. Signal processing procedures for transient and steady state responses were as follows.

(1) Transient Responses. Epochs were averaged for each subject and EEG recording session. Figure 1B (middle trace) shows this average for the second test session (Fz electrode, high pass filtered at 1 Hz) where a 40 Hz oscillation can be seen to be riding on a slower transient waveform. The averaged data for each subject were filtered 1-15 Hz forward and backward (zero phase shift) with a 6^{th} order Chebyshev filter to remove the 40 Hz component, thereby setting into relief the transient waveform with prominent P1, N1, and P2 components (upper trace, Figure 1B). Spherical spline maps of current source density were generated at the amplitude maximum of each component to show scalp topography (64 channel subjects only). The amplitude and latency of P1, N1 and P2 components, and a fourth component (the N1c, not identified in Figure 1B) which showed properties of interest, were determined by a computer algorithm that searched electrodes containing their amplitude maxima for amplitude peaks occurring within latency windows determined from the grand averaged data. P1 amplitude was recorded as the maximum voltage occurring in the Fz electrode between 40 ms and 110 ms after stimulus onset. N1 amplitude was recorded as the most negative voltage occurring at Fz between 90 ms and 120 ms after stimulus onset, and P2 amplitude as the maximum voltage occurring between 120 ms and 200 ms at this electrode. The N1c was defined as the minimum voltage occurring between 120 ms and 180 ms at electrodes T7 and T8 in accordance with the radial orientation of this brain event in each hemisphere (Woods, 1995). The difference between N1 and P2 amplitude (P2 minus N1) was also calculated for each subject. This metric provided a conservative estimate of P2 amplitude by removing possible contributions arising from changes in the overlapping N1 waveform.

(2) 40-Hz Steady State Response. The lower trace of Figure 1B shows the time-domain average of the 40-Hz SSR extracted from the trace seen in the middle panel by band pass filtering (30-50 Hz, zero phase shift, 8^{th} order Butterworth). In principle, changes induced in the 40-Hz SSR by discrimination training could be expressed tonically throughout the S1 stimulus or be confined to more restricted epochs of the stimulation period. In addition, training could modify the number of neurons activated by the S1 stimulus which would be expected to affect SSR amplitude, or the temporal properties of the neural representation which could influence SSR phase. In order to evaluate

these possible effects, a single trial analysis was conducted in which a Fourier transformation was applied within a hamming window 100 ms wide that was moved across the EEG (from 400 ms prior to S1 onset to 400 ms after S1 offset) in 10 ms time steps. Within each window, separately for each subject, trial, and test session, the 40 Hz component was represented as a vector in a polar plot where SSR amplitude was given by vector length and SSR phase by the angle θ in as depicted in Figure 1C. In this representation confidence limits circling the vector endpoints do not include the origin when an SSR is present (see the example of Picton et al., 1987). The likelihood of this outcome under the null hypothesis is distributed as Hotelling's bivariate T^2 (Valdes-Sosa, Bobes et al., 1987; Victor and Mast, 1991). Rejection of the null hypothesis implies the presence of a 40-Hz SSR of some phase and amplitude in the test session.

This technique gave an assessment of the SSR on each test session administered before and after discrimination training. In order to evaluate before/after training effects on the SSR, we utilized a two-sample version of the T^2 test to contrast the two test sessions (Timm, 1975). This test is conceptually similar to the single sample case and is calculated as:

$$t^2 = \frac{N_1 N_2}{N_1 + N_2} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \, \mathbf{S}^{-1} \, (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$$

where:

$$\mathbf{S} = \frac{\mathbf{A}_1 + \mathbf{A}_2}{N_1 + N_2 - 2}$$

$$\mathbf{A}_1 = \sum_{i=1}^{N_1} (\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)(\mathbf{x}_{1i} - \bar{\mathbf{x}}_1)'$$

$$\mathbf{A}_2 = \sum_{i=1}^{N_2} (\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)(\mathbf{x}_{2i} - \bar{\mathbf{x}}_2)'$$

Because the number of T^2 statistics generated for each subject and session was large and not independent for overlapping 100 ms windows, critical values of T^2 for statistical evaluation of before/after differences were determined by Monte Carlo simulations conducted separately for each subject as described below.

Significant before/after differences identified by T^2 established that some aspect of the SSR had been modified by training. However, further evaluation was necessary to identify which aspect of the response had changed. For this purpose we calculated for each subject, test session, and 100 ms window the mean phase of the vectors (SSR phase) and mean vector length (the resultant, SSR amplitude) in order to identify which of these measures contributed to before/after differences in the 40-Hz SSR. Also calculated were (1) phase coherence by the method of Picton et al. (2001), and (2) absolute vector length for each test session. Comparison of these measures between test sessions allowed determination of whether before/after differences in SSR amplitude measured as the mean vector were a consequence of a decrease SSR phase variability around its central tendency or an overall increase in vector length independent of phase. Absolute vector lengths were normalized with respect to the maximum length observed for each subject before contrasting before/after differences.

Amplitude modulation of the discriminative stimuli permitted identification of a response at the modulation frequency (the 40-Hz SSR) whose cortical sources have been found to localize to the region of Heschl's gyrus in AI. However, transient responses (whose cortical sources are distributed in AII) could in principle contain a 40 Hz spectral component of the transient waveform that is potentially confusable with the 40-Hz SSR. To evaluate this possibility, we applied the T^2 method to evaluate 40 Hz activity in a separate control group of 8 subjects (undergraduate students paid $8/hr) who performed the 2.0 kHz discrimination task for a single session without feedback for correctness, using unmodulated S1 and S2 stimuli. This condition gave an estimate of 40 Hz energy present in transient AEPs evoked by acoustic stimulation when the 40-Hz SSR was absent.

Source Modeling

Source analysis of the average-referenced AEP field patterns (N1, N1c, P2, and the 40-Hz SSR) was carried out using BESA 2000 (MEGIS GmbH, Munich, Germany). Analyses were conducted separately for each stimulus set and test session using the group averaged data. Two regional sources were used to describe the cortical generators for each AEP component (one source in each hemisphere, constrained to localize symmetrically following Scherg and von Cramon, 1986). Sources were determined at the peak of the AEP waveform (root mean squared transformed) within the same latency windows used for analysis of the EEG. Medial/lateral (x), anterior/posterior (y), and inferior/superior (z) coordinates of each regional source were recorded together with dipole moment. The residual variance of the source model averaged 1.4%, 3.5%, 4.8%, and 1.8% for the N1, N1c, P2 and SSR respectively (2.7% overall), with no fit exceeding 7.0% residual variance. It should be noted that regional sources determined by BESA use three orthogonal vectors (one in each plane)

to describe cortical activations contributing to AEPs. These vectors were investigated further as described in the results section, to provide information on the relative contribution of tangential and radial vectors to N1 and N1c transient responses and SSR waveforms.

Statistical Evaluation

Changes in behavioral performance and in transient AEPs induced by discrimination training were evaluated by repeated measures analyses of variance (ANOVAs). Analyses applied to the two test sessions included the variables before/after training and stimulus set (S1 stimuli of the trained set and the two control sets). Pre-planned contrasts were evaluated by conventional *t*-tests and post-hoc contrasts with the Least Significant Difference test. Peak amplitude and latency were analyzed for the AEPs, and, for behavioral performance, the metrics ***P***, ***d'***, discrimination threshold, and slope of the psychophysical functions determined for each stimulus set. All probabilities are two-tailed unless otherwise stated.

Monte Carlo methods were used to evaluate the 40-Hz SSR. The presence of an SSR for each subject and test session was not in doubt; T^2 for the 40-Hz Fourier component exceeded 45 in all subjects and 100 ms moving windows. In order to contrast the test sessions for training effects, we generated the distribution of T^2 under the null hypothesis for each subject and stimulus set using the procedure of Manly (1991). For each moving window 144 trials were taken at random from the maximum of 288 trials that were available after artifact rejection in the first test session (before training), and a further 144 trials were taken the trials available in the second test session (after training). These 288 trials were used to calculate T^2 when no difference was expected between before/after measurements. This constituted one simulation. One thousand of these simulations were performed for each stimulus set to approximate the distribution of T^2 under the null hypothesis. Although these simulations were conducted separately for each subject, the results across subjects were similar, and we found that a critical value of $T^2 = 8.0$ created a rejection region of $p < 0.01$ for all subjects considered singly. In order to determine a critical value to apply to a T^2 map of a group of subjects, we combined one randomly selected "null hypothesis" map from each subject into a group mean map, and repeated this process 1000 times to generate a distribution for this map under the null hypothesis. In this case a critical value of $T^2 = 6.0$ was found to depict $p < 0.01$. The SSR was evaluated at several electrode sites but the response in the 40 Hz region was maximal at Fz and only the results for this electrode are reported.

RESULTS

Behavioral Performance

Behavioral performance (**P**) on the trained stimulus set is shown over the 15 sessions of training in Figure 2A, where performance on the opening and closing test sessions is also depicted. Performance improved rapidly from the opening test session and then more gradually thereafter. A significant main effect of training sessions ($F_{(14,92)} = 4.72$, $p < 0.001$) was found, as were significant preplanned contrasts between training sessions 1 and 15 ($t_{(7)} = 3.86$, $p = .006$) and 3 and 13 ($t_{(7)} = 2.68$, $p = 0.03$) which corroborated gradual improvement throughout the training series. Performance on the test sessions given before and after training is contrasted for the trained stimulus set and the two control sets in Figure 2B. Main effects were found for before/after ($F_{(1,7)} = 19.27$, $p = 0.003$) and for stimulus set ($F_{(2,14)} = 7.32$, $p = 0.006$) and as well as an interaction of these variables ($F_{(2,14)} = 22.01$, $p < 0.001$). Performance improved after training on all three stimulus sets, but moreso for the 2.0 kHz set ($t_{(7)} = 6.90$, $p < 0.001$) than for the 1.8 kHz ($t_{(7)} = 2.66$, $p = 0.04$) and 2.2 kHz ($t_{(7)} = 2.72$, $p = 0.03$) untrained stimuli.

Figure 2 about here

Training effects were corroborated by *d'* and by psychophysical functions calculated for each subject. When averaged over subjects *d'* increased from 0.99 at the outset of training (sessions 1-3 collapsed) to 1.59 at the end of training (sessions 13-15 collapsed), giving $t_{(7)} = 4.69$, $p = 0.002$. Psychophysical functions are shown for each stimulus set in Figure 2C. Discrimination thresholds (Δf at **P** = 0.5) decreased from 20.3 Hz, 20.2 Hz, and 16.7 Hz prior to training for the 1.8, 2.0, and 2.2 kHz sets, respectively, to 9.3 Hz for the trained 2.0 kHz set and to 16.0 Hz and 11.6 Hz for the 1.8 and 2.2 kHz sets, respectively. These results gave rise to a main effect of before/after ($F_{(1,7)} = 7.624$, p = 0.028) and to an interaction with stimulus set ($F_{(2,14)} = 6.289$, $p = .011$) which was attributable to before/after differences appearing for the trained stimuli ($t_{(7)} = 2.99$, $p = 0.02$) but not for either of the

control sets. When the threshold of discrimination at 2.0 kHz was divided by stimulus frequency after training ($\Delta f/f$), a ratio of 0.46% was found which is similar to ratios reported by discrimination studies using unmodulated tones (He et al., 1998). The slope of the psychophysical function after training was steepest for the 2.0 kHz trained stimulus set and shallowest for the 1.8 kHz control set (Figure 2C), but differences in slope among the stimulus sets did not reach significance.

Six subjects returned for a retention test on the 2.0 kHz stimulus set two months after their last test session. Performance at retention (**P** = 0.63) was lower than on the last training session (**P** = 0.75, $t_{(5)} = -2.80$, $p = 0.038$) but remained better than in the first test block (**P** = 0.32, $t_{(5)} = 3.76$, $p = 0.013$).

Transient AEPs

N1 and P2 transient responses evoked by the S1 reached their amplitude maxima at frontal electrodes with a polarity reversal at occipital sites. Time domain averages at the frontal electrode (Fz) and global field power (root mean square of all electrodes) are shown for the trained 2.0 kHz S1 in Figure 3A where N1 and P2 components are identified (pretraining latencies of 116 ms and 172 ms respectively). The early occurring P1 (pretraining latency 57 ms) is also identified in these traces, and scalp topographies are shown for the N1 and P2 at their posttraining amplitude maxima. These results show that discrimination training resulted in an enhancement of P2 amplitude. When referred to the prestimulus baseline, P2 amplitude increased from 0.65 μV before training to 1.46 μV after training ($t_{(7)} = 6.03$, $p < .001$), corresponding to an increase of 124% for the group as a whole. Enhancement of the P2 was especially prominent in global field power (inset, Figure 3A). On the other hand, N1 and P1 amplitude tended to decrease after training, but these effects did not reach significance.

Figure 3 about here

P2 amplitude is shown before and after training for each stimulus set in Figure 3B, referenced in this case to the peak of the N1 (P2-N1 amplitude) in order to remove influences attributable to variability in the N1. Analysis of variance revealed a main effect of before/after ($F_{(1,7)} = 6.7$, $p = 0.036$) but the interaction of before/after with stimulus set was not significant. When the stimulus sets were examined separately, before/after differences were found to be significant only for the trained 2 kHz stimulus set ($t_{(7)} = 4.26$, $p = 0.008$). However, differences for the control sets were in the direction of training and suggested partial generalization of P2 enhancement to the untrained stimuli. Correlations were calculated between before/after differences in P2-N1 amplitude and the behavioral measure **P** for the trained stimulus set alone, and when the three stimulus sets were combined. These correlations were positive but none reached significance.

Acquisition of the enhanced P2 (referenced to the prestimulus baseline) is shown in Figure 3C which includes training sessions 3 and 13 as well as the opening and closing test sessions. A main effect of sessions was found for this measure ($F_{(3, 21)} = 4.15$, $p = 0.019$) which was attributable to increases in P2 amplitude occurring on the 13th session of training and on the closing test session compared to session 3 and pretraining performance ($p < 0.015$ or better). For purposes of comparison, Figure 3C also depicts changes observed in the amplitude of P1 and N1 responses referenced to their prestimulus baselines. Main effects of sessions did not reach significance for either measure.

Figure 3D depicts changes occurring over sessions in a fourth AEP component, identified over the right hemisphere (electrode T8) as the surface-negative N1c in accordance with properties described by Woods (1995). This component was distinguishable from the N1 and P2 by its radial orientation, by its latency (155 ms) falling between that of these two AEPs, and by its preferential expression in the right hemisphere. Discrimination training enhanced this component of the AEP between the two test sessions for the trained S1 stimulus ($t_{(7)} = 3.81$, $p = 0.007$), gradually over the training series (see Figure 3C; main effect of sessions $F_{(3, 21)} = 4.05$, $p = 0.02$). Although before/after differences were largest for the trained stimulus set, enhancement generalized as well to the 2.2 kHz control set where before/after differences reached significance ($t_{(7)} = 2.89$, p = 0.023, see inset of Figure 3D). Analysis of variance applied to the two test sessions confirmed these findings by revealing a main effect of before/after ($F_{(1, 7)} = 9.52$, $p = 0.018$) whereas effects involving stimulus set did not reach significance. We also searched for an N1c occurring in the left hemisphere in the test sessions (electrode T7, not shown in Figure 3D). An enhanced polarity-inverted response was observed after training at a peak latency (155 ms) that corresponded with the amplitude maximum of the N1c recorded in the right hemisphere. However, this polarity inverted enhancement did not reach significance ($t = -0.84$) nor were before/after

differences detected at any other time point in the T7 trace of the left hemisphere.

We also examined the effect of discrimination training on the latency of the P1, N1, N1c, and P2 responses evoked by the trained S1. N1 latency decreased from 116 ms in the first test session to 107 ms in the closing test session, $t_{(7)} = 7.94$, $p < .001$. This effect was seen in every subject and is apparent in Figure 3A (time domain traces and global field power). P1 and P2 latency, and N1c latency in the right hemisphere, did not change with discrimination training when measured to their amplitude maxima, although the leading edge of the P2 and N1c waveforms tended to commence earlier after training compared to their pretraining baselines (see Figures 3A and 3D).

Steady State Response

A time domain trace of the 40-Hz SSR evoked by the 2.0 kHz S1 after training is depicted in the lower trace of Figure 1B at its amplitude maximum (Fz electrode). Neither responding at this electrode nor SSR global field power differed between test sessions administered before and after training when calculated over the 1s S1 period. However, fine grained dynamics were revealed by T^2 when 100 ms windows were moved across the 40 Hz waveform at Fz in 10 ms time steps as described in Figure 1C. Figure 4A gives the results for a representative subject. The two polar plots shown in Figure 4A (right side) contain vectors representing SSR amplitude and phase for each of the 288 accepted test trials in a single 100 ms window before (upper, test 1) and after (lower, test 2) discrimination training. Although phase covers 360 degrees and is variable across single trials, the mean vector (resultant, shown as the red arrow) is shifted from the origin in both polar plots, indicating that a 40-Hz SSR is present. Spectral plots of T^2 are shown to the left of Figure 4A and indicate that a 40-Hz SSR was present throughout the stimulation period before (upper plot) and after (middle plot) training (all $T^2 > 45$). The lower spectral plot in Figure 4A shows the T^2 difference between the two test sessions before and after training for this subject, scaled for Monte Carlo significance at $T^2 = 8.0$, $p < .01$. Before/after differences reached significance particularly in the first half of the S1 stimulation period with patches of significance appearing subsequently.

Figure 4 about here

Similar findings were obtained for all subjects to which this analysis was applied. The results are collapsed across subjects in Figure 4B where the before/after T^2 difference is shown for the trained S1 (2.0 kHz) as well as for the S1s of the untrained 1.8 kHz and 2.2 kHz stimulus sets. Time-domain traces of the 40-Hz SSR evoked by the trained S1 before and after training are superimposed above the T^2 difference map for the 2.0 kHz stimulus. Inspection of the T^2 data shows that significant before/after differences were observed in the SSR evoked by the trained S1 particularly in the time interval 150-225 ms after S1 onset, with brief epochs of significance appearing thereafter. Integration of the T^2 statistic over the time interval 50-400 ms at 40 Hz found that before/after differences were stronger for the trained 2.0 kHz S1 than for the untrained 1.8 kHz S1 ($t_{(5)} = 2.29$, $p = 0.035$, one-tailed test) while differences between the 2.0 kHz and 2.2 kHz S1 stimuli were not significant. These results indicate that generalization occurred from training on the 2.0 kHz set to the 2.2 kHz control set, but not to the 1.8 kHz control set.

Augmentation of the SSR within the interval 150-225 ms raises the question of whether the T^2 results shown in Figure 4B might alternatively be attributed to a 40 Hz spectral component of the transient P2 which was also augmented in the vicinity of this time window. To assess this hypothesis, we evaluated 40 Hz activity in the absence of the SSR when N1 and P2 transient responses were evoked by unmodulated 2.0 kHz tones. The results are shown in Figure 4C where the N1/P2 waveform evoked by the unmodulated tone is superimposed on 40 Hz activity evaluated by T^2 at the same scaling used for the upper two T^2 maps of Figure 4A. 40 Hz activity was detected between 30-50 ms where middle latency responses with a wave period near 25 ms would be expected to appear, but no 40 Hz activity was seen within the latency window encompassing N1 and P2 transient responses. These findings indicate that T^2 differences observed for the 2.0 kHz amplitude-modulated S1 (Figure 4B) are not likely to be attributable to a high-frequency component of the enhanced P2 transient response, because no such component was detected in the transient response in the unmodulated control condition. Rather, the two responses appeared to be separate brain events.

Changes in the SSR induced by training and detected by T^2 could be generated by changes in the amplitude or phase of the SSR, or both. In order to address this question, we first calculated the mean amplitude and phase of the SSR for each subject and 100 ms window during the S1 stimulus. The results are shown in Figure 4D for the group as a whole where each measure is aligned to the transient N1/P2

waveform obtained before and after training (top panel). For convenience, the T^2 difference plot for the trained S1 is repeated at the bottom of the panel of Figure 4D; the shaded area covers the time interval of the maximum T^2 difference. Inspection of the phase data, shown as the phase difference between the stimulus and the SSR, for the first test session (blue trace) shows that SSR phase shortened gradually commencing about 100 ms after S1 onset and continuing until asymptote near 400 ms. After training (red trace) SSR phase difference advanced by about 0.3 radians (4.8% of the wave period of the SSR) with respect to pretraining performance within this time interval, commencing near but persisting beyond the leading edge of the P2 waveform. Phase advances tended to recur subsequently during the S1 interval, coinciding with significant differences in the T^2 difference map. On the other hand, before/after differences in SSR amplitude depicted as mean vector length were less apparent during the S1 (Figure 4D, second panel), although a small enhancement is seen during the interval 100-200 ms after stimulus onset. Supplementary analyses not presented in the figure showed that this enhancement was closely paralleled by an increase in phase coherence with no change in absolute vector length, suggesting that it was secondary to a reduction of phase variability around its central tendency during this interval. Multiple regression applied to T^2 differences recorded for the group during the S1 yielded $R = 0.455$ ($F_{(2,96)} = 15.5, p < 0.00001$) to which before/after differences in phase contributed ($t_{(96)} = 5.00$, $p < 0.00001$) but differences in mean vector length did not ($t_{(96)} = -0.54$, $p = 0.41$). These findings indicate that discrimination training modified the temporal properties of the 40-Hz SSR but had little effect on the absolute amplitude of this response. A brain movie showing phase and amplitude dynamics of the mean vector for a representative subject throughout the S1 can be viewed at www.psychology.mcmaster.ca/hnplab.

Source Analyses

The spatial coordinates of regional sources modeled from the grand averaged data for each AEP (N1, N1c, P2, and SSR) were evaluated by analyses of variance collapsing first over before/after test sessions (to examine effects of stimulus set) and then over stimulus sets (to examine effects of before/after). No effects of stimulus set or before/after were found, except for the sources of the P2 which shifted to be more inferior when training had been completed (z coordinate, $F_{(3,12)} = 13.53$, $p = 0.0007$). Main effects attributable to AEP were found in both of these analyses. When the six localizations determined for each AEP (three stimulus sets before and after training) were collapsed into a single data set, main effects of AEP were significant for the medial lateral (x) coordinate ($F_{(3,15)} = 25.97$, $p < 0.00001$), anterior-posterior (y) coordinate ($F_{(3,15)} = 9.22$, $p = 0.001$), and inferior-superior (z) coordinate ($F_{(3,15)} = 24.95$, $p < 0.00001$). The modeled sources for each AEP are co-registered on the average brain of BESA 2000 in Figure 5 in order to visualize their relative positions. Post-hoc contrasts showed that cortical sources underlying the N1 and N1c were centered lateral with respect to those of the P2 in the region of the auditory cortex ($p < 0.01$ or better, axial view), while sources of the SSR were medial with respect to P2, N1, and N1c sources ($p < 0.03$ or better). P2 sources were also centered anterior with respect to sources of the N1, N1c, and SSR ($p < 0.05$ or better), and superior with respect to these sources (minimum $p < 0.0001$) notwithstanding the inferior shift of P2 sources after training. These results which confirm SSR sources medial to those of the N1 and P2 are consistent with previous studies which have localized SSR generators to the region of Heschl's Gyrus (Engelien et al, 2000; Gutschalk et al., 1999; Pantev et al., 1993, 1996a; Schneider et al., 2002; Shahin et al., 2003). The relative positions observed for P2 and N1 sources are also in agreement with previous findings which have localized the generators of these responses to the secondary regions of the auditory cortex in the superior temporal gyrus (Scherg and von Cramon, 1986; Pantev et al., 1996b; Picton et al., 1999) including for the P2 sites anterior to the auditory core (Hari et al., 1987; Joutsiniemi et al., 1989; Pantev et al., 1996b).

Figure 5 about here

Dipole moment was also contrasted for each AEP before and after training, using the three stimulus sets as the unit of observation. This analysis revealed a main effect of before/after ($F_{(1,4)} = 12.83$, $p = 0.023$) and an interaction of before/after with AEP ($F_{(3,12)} = 6.331$, $p = 0.008$). Both of these effects were attributable to enhanced dipole moments occurring for the P2 in each stimulus set after training ($F_{(1,4)} = 18.06$, $p = 0.013$) compared to the other AEPs. Dipole moment was not significantly enhanced after training for any other component in either hemisphere. However, subsequent analyses showed that the regional source fitted to the N1 field pattern contained a

radially-oriented vector that was augmented after discrimination training only in the right hemisphere, with an amplitude peak near 148 ms when the N1 source model was applied to the N1c time interval. This suggests that dipole moment calculated for the regional source fitted to the N1c field pattern contained contributions arising from the temporally overlapping N1 that obscured changes in radially-oriented N1c activity. We also examined the contribution of the three orthogonal vectors of the SSR regional source to the SSR waveform after discrimination training, following the report of Scherg and von Cramon (1986). The regional model accounted for 97.9% of the observed field pattern when the three vectors were included. Goodness of fit decreased to 93.3% when only a single tangential source was used to model the field pattern, whereas a single radial source accounted for only 6.3% of the variability in the recorded field pattern. These findings indicate that activity modeled by the tangential vector was the principal contributor to the SSR waveform.

DISCUSSION

We trained nonmusician subjects to discriminate small increases in the pitch of a 2.0 kHz standard stimulus, using 40-Hz AM modulated pure tones as the discriminative stimuli. Amplitude modulation allowed us to separate the 40-Hz auditory SSR whose generators localize to the region of Heschl's gyrus in AI from transient responses of the AEP (N1, N1c, P2) which reflect the behavior of differentiable populations of neurons distributed widely in AII. Discrimination improvement was accompanied by enhancement of the P2 (latency 172 ms) and of the N1c (in the right hemisphere, latency 155 ms), indicating that more neurons were activated in AII after training on the discrimination task. The 40-Hz SSR, on the other hand, gave a different picture of cortical dynamics. Overall, there was no overall amplitude enhancement of the SSR; instead we observed a shortening of phase within a latency window coinciding with the onset of the P2 with brief advances in phase reappearing subsequently during the S1. These findings suggest that training at pitch discrimination did not expand the cortical representation for the 2.0 kHz S1 in AI. Rather, the temporal properties of neurons contributing to the SSR were modified as a consequence of experience on the task or by inputs arising from other regions of the brain.

Our evidence for P2 enhancement by training at pitch discrimination extends early findings reported by Eaton and Roberts (1999) and corroborates the results of Tremblay et al. (2001) who trained nonmusician subjects to discriminate temporal features of speech signals. More recently, Atienza et al. (2002) found an enhancement of the P2 when subjects were trained to detect pitch deviants in a short stream of pitch stimuli. In each of these studies P2 amplitude increased by a factor exceeding 1.5 when measured from the amplitude peak of the N1 which did not change with training in any study. These findings indicate that the P2 brain event is sensitive to neuroplastic change. Heretofore this component of the AEP has received little attention in studies of auditory perception, perhaps because in the absence of a training manipulation the P2 shows more limited dynamics.

Enhancement of the N1c by acoustic training has not previously been reported. The expression of the N1c in the right hemisphere in our study where subjects were processing pitch cues is consistent with functional and anatomical evidence for specialization of auditory neurons in this hemisphere for processing of spectral information. Compared to homologous auditory neurons in the left hemisphere, neurons in the right hemisphere are characterized by higher synaptic densities, more closely spaced cortical columns, and comparatively less myelination which are features that may favor spectral integration of acoustic signals (Zatorre and Belin, 2001). Woods (1995) noted that less is known about the N1c component of the AEP compared to other components, but that with binaural stimulation it is observed preferentially in the right hemisphere. A key to expression and enhancement of the right-sided N1c may be the presence of multiple auditory objects in a stimulus sequence that must be distinguished by their spectral properties in order for the subject to comply with task requirements.

In contrast to the P2 and N1c, the N1 AEP (latency 107 ms) was not amplified by discrimination training in our study or in the aforementioned EEG studies of acoustic discrimination. However, enhancement of its magnetic counterpart the N1m by training at pitch discrimination has been reported by Menning et al. (2000). It should be noted that an augmented P2 brain event commencing within the N1 response window would subtract from N1 amplitude in electrical recordings but not in magnetic ones owing to the insensitivity of magnetic sensors to radial currents contributing to the P2. Although N1 amplitude was not modified, N1 latency diminished by 9 ms after training in our study. Competition among neuroplastic synapses favoring fast inputs could generate a latency shift of this magnitude (Song et al., 2000) as could an overlapping of AEP components. In this respect it may be noteworthy that N1c and P2 responses tended to commence earlier after training within a time interval

coinciding with the onset of the N1 (see Figures 3A and 3D). When we modeled the N1 field pattern with a regional source, a radial component appeared in the right hemisphere with an early onset latency that could have reflected a contribution arising from the N1c.

The cortical sources that we modeled for the P2 and N1c were consistent with previous studies that have differentiated these sources residing in AII from those of the 40-Hz SSR which localize more medially to Heschl's gyrus in AI (Schneider et al., 2002; Pantev et al., 1993; Shahin et al., 2003). However, the changes we observed in the SSR after training did not include amplitude enhancements which were expected on the basis of the study of Recanzone et al. (1993) in owl monkeys where an expansion of the tonotopic representation for the trained frequencies was seen. Rather, our results are more in line with those of Kilgard et al. (2001) who found that behavioral conditioning with multiple frequencies tended to preserve segregated tonotopic representations in AI. Several variables may account for the different findings among these studies including the training procedures used, their duration, and the methods used to measure cortical reorganization. With respect to the latter variable, it should be noted that our results do not appear to be attributable to insensitivity of the SSR to the anatomy or functional organization of Heschl's gyrus. Schneider et al. (2002) recently found that the N19-P30 source waveform underlying the SSR was augmented by 102% in musician compared to nonmusician subjects, when extracted by deconvolution from AM rates near 39 Hz. The SSR source waveform also correlated highly (r = .87) with the volume of gray matter in the anteromedial portion of Heschl's gyrus well as with musical aptitude (r = .71). In our study temporal modulation of the SSR generalized more to the untrained 2.2 kHz S1 than to the untrained 1.8 kHz S1, perhaps because subjects were trained to detect only increases from 2.0 kHz and thus experienced carrier frequencies in the 2.0 to 2.1 kHz range and no frequencies lower than this range. Although behavioral performance did not differ significantly between the two control sets, behavioral performance was consistently better on the 2.2 kHz set as assessed by **P**, d', discrimination thresholds, and the slope of psychophysical functions obtained after discrimination training.

Modification of distributed auditory cortical representations in the present study raises the question of how remodeling was achieved and expressed in the AEP. Epipial and intracortical measurements taken by Sukov and Barth (1998) of middle latency responses in rat auditory cortex indicate that positive-going surface potentials are generated by depolarization of pyramidal neurons in neocortical layers III-VI while surface negativities reflect depolarization of apical dendrites in the upper neocortical laminae. Laminar analyses reported for the cat are in agreement with these findings (see activations A and C respectively, of Mitzdorf, 1985, p. 80; Mitzdorf, 1994). If this interpretation of scalp-recorded potentials is accepted for the P2 and N1c components of the human AEP, our results imply that more pyramidal neurons were depolarizing in AII after training on the discrimination task than before training commenced. Modulation of the neocortical mantle by the basal forebrain (nucleus basalis magnocellularis, NBM) is one possible source of these enhancements. This structure, which has been implicated in neuroplastic remodeling by many researchers (e.g. Dykes, 1997; Edeline, 1999; Weinberger et al., 1990; Wenk, 1997), contains large cholinergic and GABAergic neurons that project to targets in the neocortex in a broadly-tuned corticotopic arrangement (Jiménez-Capdeville et al., 1997). Because GABAergic fibres synapse on inhibitory interneurons (Freund and Meskenaite, 1992), coactivation of cholinergic and GABAergic pathways acts synergistically to increase the sensitivity of pyramidal cells to their afferent inputs, shortening response latency by a magnitude similar to that which we observed in SSR phase after training (Metherate and Ashe, 1993) and strengthening synaptic connections on auditory neurons by a Hebbian correlation rule (Kilgard and Merzenich, 1998; Metherate and Weinberger, 1990; Cruikshank and Weinberger, 1996). When measured by slow dc potentials this modulatory influence has an onset latency resembling that of the auditory N1/P2 complex (Pirch, 1993; Pirch et al, 1983) as do top-down signals from prefrontal cortex which may converge on auditory neurons and serve an additional teaching function (Tomita et al., 1999). Although strengthening of modulatory inputs by conditioning (Rigdon and Pirch, 1986; Tomita et al., 1999) may itself account for the augmented P2, evidence summarized by Dykes (1997) indicates that receptive fields for the experienced stimuli are likely to be established on AII neurons and to contribute to progressive improvements in behavioral performance such as those observed in our study. Network behavior of this nature would be expected to influence neuroplastic remodeling of sensory modalities in addition to audition, although not necessarily at the same latencies observed in the auditory case.

Shahin et al. (2003) recently reported that P2 responses evoked by musical tones in violinists and pianists were larger than those observed in nonmusician subjects, as were right-sided N1c's.

These results could have been predicted from the present findings owing to the different training histories of musicians and nonmusicians with respect to tones of musical timbre. On the other hand, our findings with regard to the effects of training on the 40-Hz SSR suggest a dissociation of transient and SSR components of the AEP, with neuroplastic transient responses expressing as amplitude enhancements in training studies and in musicians but 40-Hz SSR enhancements in musicians only (Schneider et al., 2002) where they may be an anatomical marker for musical skill. However, we cannot exclude the possibility that other training experiences may modify SSR amplitude and its anatomical substrate, depending on the type of training that is given, its duration, and when it is delivered in the course of brain development.

ACKNOWLEDGEMENTS

This research was supported by grants from the Canadian Institutes of Health Research and the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Atienza M, Cantero JL, Dominguez-Marin E (2002) The time course of neural changes underlying auditory perceptual learning. Learn Mem 9:138-150.

Buonomano DV, Merzenich MM (1998) Cortical plasticity: From synapses to maps. Annu Rev Neurosci 21:149-186.

Celesia GG (1976) Organization of auditory cortical areas in man. Brain 99:403-414.

Cruikshank SJ, Weinberger NM (1996) Receptive-field plasticity in the adult auditory cortex induced by Hebbian Covariance. J Neurosci 16:861-875.

Dember WN, Warm JS (1979) Psychology of perception. New York: Holt Rinehart Winston.

Diamond DM, Weinberger NM (1984) Physiological plasticity of single neurons in auditory cortex of the cat during acquisition of the pupillary conditioned response: II. Secondary field (AII). Behav Neurosci 98:189-210.

Dykes RW (1997) Mechanisms controlling neuronal plasticity in somatosensory cortex. Can J Physiol Pharmacol 75: 535-545.

Eaton RA, Roberts LE (1999). Effect of spectral frequency discrimination on auditory transient and steady state responses in humans. Soc Neurosci Abstr 29:156.15.

Edeline J (1999) Learning-induced physiological plasticity in the thalamo-cortical sensory systems: A critical evaluation of receptive field plasticity, map changes and their potential mechanisms. Prog Neurobiol 57: 165-224.

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

Freund TF, Meskenaite V (1992) γ-Aminobutyric acid-containing basal forebrain neurons innervate inhibitory interneurons in the neocortex. Proc Natl Acad Sci USA 89:738-742.

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

Godey B, Schwartz D, de Graaf JB, Chauvel P, Liégeois-Chauvel C (2001) Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. Clin Neurophysiol 112:1850-1859.

Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley

Gutschalk A, Mase R, Roth R, Ille N, Rupp A, Hähnel S, Picton TW, Scherg M (1999) Deconvolution of 40 Hz steady-state fields reveals two overlapping source activities of the human auditory cortex. Clin Neurophysiol 110: 856-868.

Hari R, Pelizzone M, Makeka JP, Hallstrom J, Leinonen L, Lounasmaa OV (1987) Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. Audiology, 26:31-43.

He N, Dubno JR, Mills JH (1998) Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. J Acoust Soc Am. 103:553-65.

Jesteadt W, Bilger RC (1974). Intensity and frequency discrimination in one- and two-interval paradigms. J Acoust Soc Am 55: 1266-76

Jiménez-Capdeville ME, Dykes RW, Myasnikov AA (1997) Differential control of cortical activity by the basal forebrain in rats: A role for both cholinergic and inhibitory influences. J Comp Neurol 381:53-67.

Joutsiniemi SL, Hari R, Vilkman V (1989) Cerebral magnetic responses to noise bursts and pauses of different durations. Audiology 28:325-333.

Kilgard MP, Merzenich MM (1998) Cortical map reorganization enabled by nucleus basalis activity. Science 279:1714-1718.

Kilgard MP, Pandya PK, Vazquez J, Gehi A, Schreiner CE, Merzenich MM (2001) Sensory input directs spatial and temporal plasticity in primary auditory cortex. J Neurophysiol 86:326-338.

Liégeois-Chauvel C, Musolino A, Badier JM, Marquis P, Chauvel P (1993) Evoked potentials recorded from the auditory cortex in man: evaluation and topography of the middle latency components. Electroenceph Clin Neurophysiol 92:204-214.

Levitt, H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467

Manly, BFJ (1991) Randomization and Monte Carlo Methods in Biology. New York: Chapman and Hall.

Menning H, Roberts LE, Pantev C (2000) Plastic changes in the auditory cortex induced by intensive frequency discrimination training. NeuroReport 11:817-822.

Metherate R, Ashe JH (1993) Nucleus basalis stimulation facilitates thalamocortical synaptic transmission in the rat auditory cortex. Synapse 14:132-143.

Metherate R, Weinberger NM (1990) Cholinergic modulation of responses to single tones produces tone-specific receptive field alterations in cat auditory cortex. Synapse 6:133-145.

Mitzdorf U (1985). Current source-density method and application in cat cerebral cortex: Investigation of evoked potentials and EEG phenomena. Physiol Rev 65:37-100.

Mitzdorf U (1994). Properties of cortical generators of event-related potentials. Pharmacopsychiatry, 27:49-52.

Pantev C, Elbert T, Makeig S, Hampson S, Eulitz C, Hoke M (1993) Relationship of transient and steady-state auditory evoked fields. Electroenceph Clin Neurophysiol 88: 389-396.

Pantev C, Roberts LE, Elbert T, Ross B, Wienbruch C (1996a) Tonotopic organization of the sources of human auditory steady-state responses. Hearing Research 101:62-74.

Pantev C, Eulitz C, Hampson S, Ross B, Roberts LE (1996b) The auditory evoked "Off" response: Sources and comparison with the "On" and the "Sustained" responses. Ear & Hearing 17:255-265.

Pantev C, Oostenveld R, Engelien A, Ross B, Roberts LE, Hoke M (1998) Increased auditory cortical representation in musicians. Nature 392:811-814.

Picton TW, Vajsar J, Rodriguez R, Campbell KB (1987) Reliability estimates for steady-state evoked potentials. Electroencephalogr Clin Neurophysiol. 68:119-31.

Picton TW, Alain A, Woods DL, John MS, Scherg M, Valdes-Sosa P, Bosch-Bayard J, Trujillo NJ (1999) Intracerebral sources of human auditory-evoked potentials. Audiol Neurootol 4:64-79.

Pirch JH, Corbus MJ, Rigdon GC (1983) Single-unit and slow potential responses from rat frontal cortex during associative conditioning. Exp Neurol 82:118-130.

Pirch JH (1993) Basal forebrain and frontal cortex neuron responses during visual discrimination in the rat. Brain Res Bulletin 31:73-83.

Recanzone GH, Merzenich MM, Schreiner CE (1992) Changes in the distributed temporal response properties of SI cortical neurons reflect improvements in performance on a temporally based tactile discrimination task. J. Neurophysiol 67:1071-1091.

Recanzone GH, Schreiner CE, Merzenich MM (1993) Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. J Neurosci 13:87-103.

Rigdon GC, Pirch JH (1986) Nucleus basalis involvement in conditioned neuronal responses in the rat frontal cortex. J Neurosci 6(9): 2535-2542.

Scherg M, von Cramon D (1986) Evoked dipole source potentials of the human auditory cortex. Electroenceph Clin Neurophysiol 65:344-360.

Schneider P, Scherg M, Dosch HG, Specht HJ, Gutschalk A, Rupp A (2002) Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. Nat Neurosci 5:688-694.

Shahin A, Bosnyak DJ, Trainor, DJ, and Roberts LE (2003) Enhancement of Neuroplastic P2 and N1c Auditory Evoked Potentials in Musicians. In Press, J. Neurosci

Song S, Miller KD, Abbott LF (2000) Competitive Hebbian learning through spike-timing-dependent synaptic plasticity. Nature Neurosci 3:919-926.

Sukov W, Barth DS (1998) Three-dimensional analysis of spontaneous and thalamically evoked gamma oscillations in auditory cortex. J Neurophysiol 79:2875-2884.

Timm NH (1975) Multivariate Analysis with applications in education and psychology. Monterey, California: Brooks/Cole.

Tomita H, Ohbayashi M, Nakahara, K, Hasegawa I, Miyashita, Y (1999). Top-down signal from prefrontal cortex in executive control of memory retrieval. Nature 401:699-703.

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

Valdes-Sosa MJ, Bobes MA, Perez-Abalo MC, Perera M, Carballo JA, Valdes-Sosa P (1987) Comparison of auditory-evoked potential detection methods using signal detection theory. Audiology.

26:166-78

Victor JD, Mast J (1991) A new statistic for steady-state evoked potentials Electroencephalogr Clin Neurophysiol 78: 378-88

Wenk GL (1997) The nucleus basalis magnocellularis cholinergic system: One hundred years of progress. Neurobiol Learn Mem 67:85-95.

Weinberger NM, Ashe JH, Metherate R, McKenna T.M., Diamond D.M., Bakin JS, Lennartz RC, Cassady JM (1990) Neural adaptive information processing: A preliminary model of receptive-field plasticity in auditory cortex during Pavlovian conditioning. In M Gabriel and J Moore (Eds.) Learning and computational neuroscience: Foundations of adaptive networks (pp. 91-138). Cambridge, MA: MIT Press.

Woods DL (1995) The component structure of the N1 wave of the human auditory evoked potential. In: Perspectives of Event-Related Potential Research (G. Karmos, M. Molnar, V Csepe, I. Czigler, JE Desmedt, Editors). Electroenceph Clin Neurophysiol Suppl. 44:102-109.

Yvert B, Crouzeix A, Bertrand O, Seither-Preisle A, Pantev C (2001). Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. Cerebral Cortex 11:411-423.

Zatorre RJ, Belin P (2001) Spectral and temporal processing in human auditory cortex. Cerebral Cortex 11:946-953.

FIGURE LEGENDS

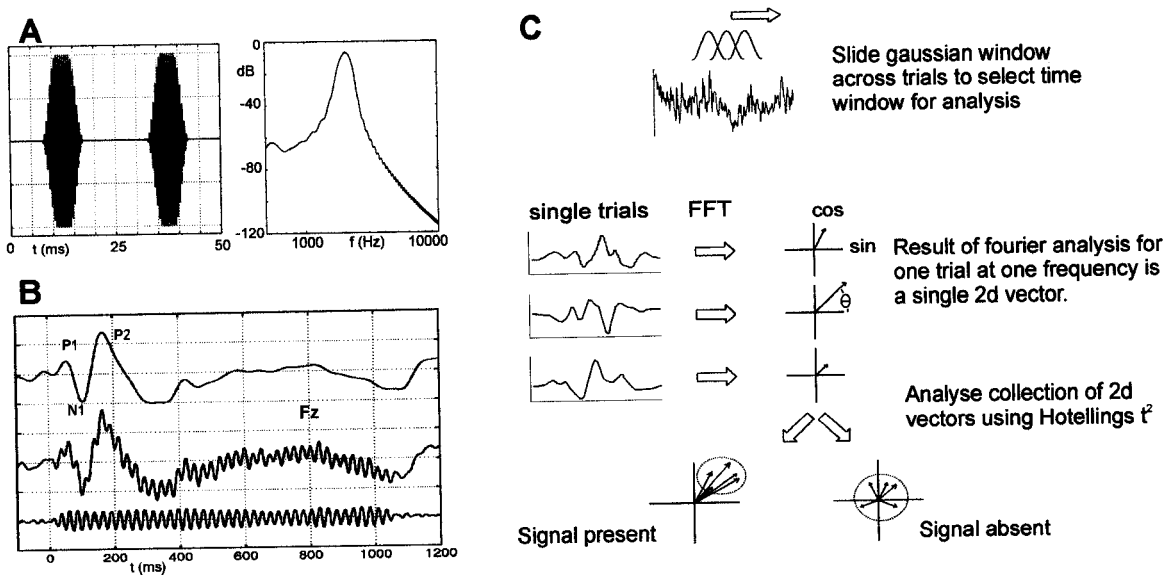
Figure 1. **A**). Waveform and spectrum of the 40-Hz AM stimulus at 2.0 kHz (the S1 stimulus of the trained stimulus set). **B**) Auditory evoked potential elicited by the 2.0 kHz stimulus of Figure 1A on test trials administered after 15 sessions of training for pitch discrimination. Middle trace: Auditory evoked potential (high pass filtered at 1 Hz) shows the 40-Hz steady state response riding on a low frequency transient waveform. Upper trace: P1, N1, and P2 transient responses are set into relief by low-pass filtering at 15 Hz to remove the 40-Hz component. Lower trace: Band pass filtering (30-50 Hz) singles out the 40-Hz steady state response. Each step on the ordinate is 1 microvolt (each trace referenced to zero). **C**) Single trial T^2 analysis of the 40-Hz steady-state response. SSR amplitude on each trial is represented by vector length and phase by the angle θ. Vector end points do not include the origin when a steady state response is present.

Figure 2. Behavioral performance. **A**). The performance measure **P** (hit rate corrected for false alarms) is plotted over 15 sessions of discrimination training and on the opening (Test 1) and closing (Test 2) test sessions. Data are for the trained 2.0 kHz stimulus set. **B**). Performance on test blocks administered before and after discrimination training is contrasted for the trained stimulus set (2.0 kHz) and for control sets (1.8 kHz and 2.2 kHz) above and below the trained stimuli. **C**). Psychophysical functions before and after training on the three stimulus sets.

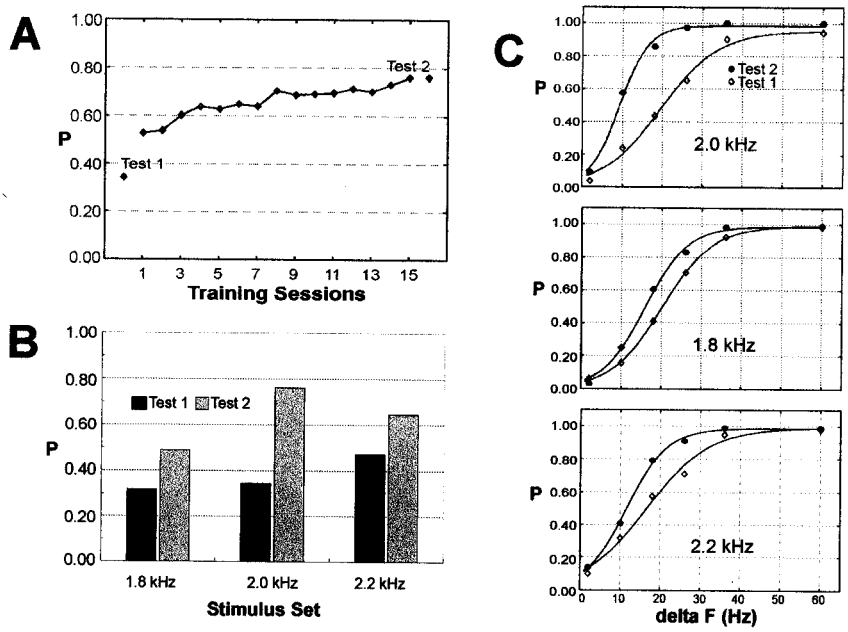
Figure 3. Effect of discrimination training on transient AEPs. **A**) Augmentation of the P2 evoked by the trained S1 stimulus. P1, N1, and P2 AEPs are identified and shown for the opening (blue, Test 1) and closing (red, Test 2) test sessions. Scalp topographies are given for the N1 and P2 AEPs. Global field power is shown in the inset. **B**). P2-N1 amplitude is shown on test sessions before and after training for each stimulus set. The 2.0 kHz set was trained. **C**). Acquisition of P2 enhancement and N1c enhancement over discrimination training. Changes observed for the P1 and N1 are also shown. Response amplitude is referenced to the prestimulus baseline and depicted for Test Session 1, Training Sessions 3 and 13, and Test Session 2 (labeled Sessions 1-4 respectively). **D**). Augmentation of the N1c in the right hemisphere by discrimination training. Insets shows scalp topography (current source density) for the trained S1 stimulus (upper left) and performance on Test Sessions 1 and 2 for each stimulus set (lower right).

Figure 4. Effect of discrimination training on the 40-Hz steady-state response. **A**). T^2 analysis applied to a single subject. Polar plots in the right panels depict the 40-Hz SSR in a single 100-ms window for the trained S1 on test blocks given before (upper, test 1) and after (lower, test 2) discrimination training. Each vector is a single trial. The spray of vectors is concentrated in the lower right quadrants indicating that an SSR is present (the red arrows are mean vectors). Corresponding T^2 plots are shown in the left panels of the figure (upper panel test 1, middle panel test 2). The lower left panel shows the T^2 difference map obtained for this subject, scaled such that red and above represent significant before/after differences in the SSR ($p < 0.01$). t^2 differences were more prominent in the first half of the stimulation period. **B**). t^2 difference maps are shown for the group as a whole for the trained S1 (2.0 kHz) and control (1.8 kHz and 2.2 kHz) S1 stimuli. Yellow and above denotes $p < 0.01$ or better. Time domain traces of the 40-Hz SSR are superimposed above the 2.0 kHz map (blue before training, red after). Before/after t^2 differences were observed for the trained S1 in the interval 100-225 ms after stimulus onset with brief periods of significance reappearing thereafter. Generalization to the untrained 2.2 kHz S1 but not to the 1.8 kHz S1 is seen. **C**). Evaluation of 40-Hz activity in control subjects tested for discrimination using unmodulated stimuli. The N1/P2 transient response evoked by the unmodulated 2.0 kHz S1 is superimposed on the t^2 map. No 40-Hz activity was observed to accompany the N1/P2 transient waveform evoked by this stimulus. **D**). Changes in SSR amplitude (mean vector) and phase during the S1 interval on test trials before (test 1) and after (test 2) training. For convenience the transient N1/P2 waveform evoked by the trained S1 (top) and the group t^2 difference map (bottom) are aligned to these data. The shaded area covers the region of the largest t^2 effect in the t^2 difference map. A phase advance occurring in the interval 100-225 ms after training was the principal source of the t^2 difference in this interval and in subsequent intervals during the S1. Periods of significance in the t^2 difference map over the duration of the S1 correlated with changes in SSR phase ($p < 0.0001$) but not with changes in SSR amplitude measured as the mean vector ($p > 0.40$).

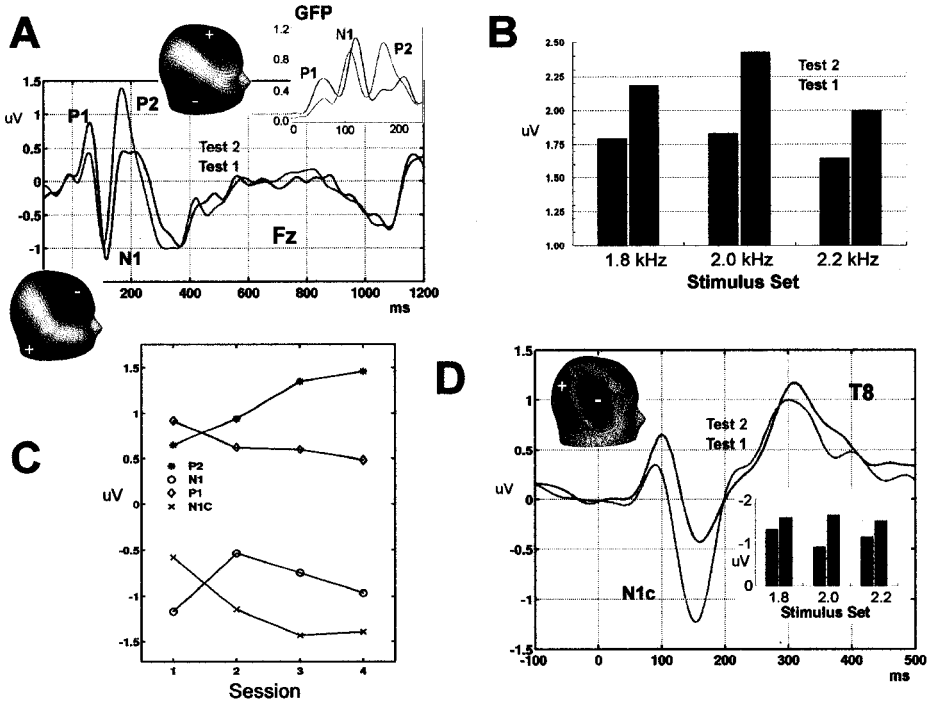
Figure 5. Localizations of the cortical sources of N1, N1c, and P2 transient responses and the 40-Hz SSR determined from the grand averaged data. The coordinate system is shown in the inset.



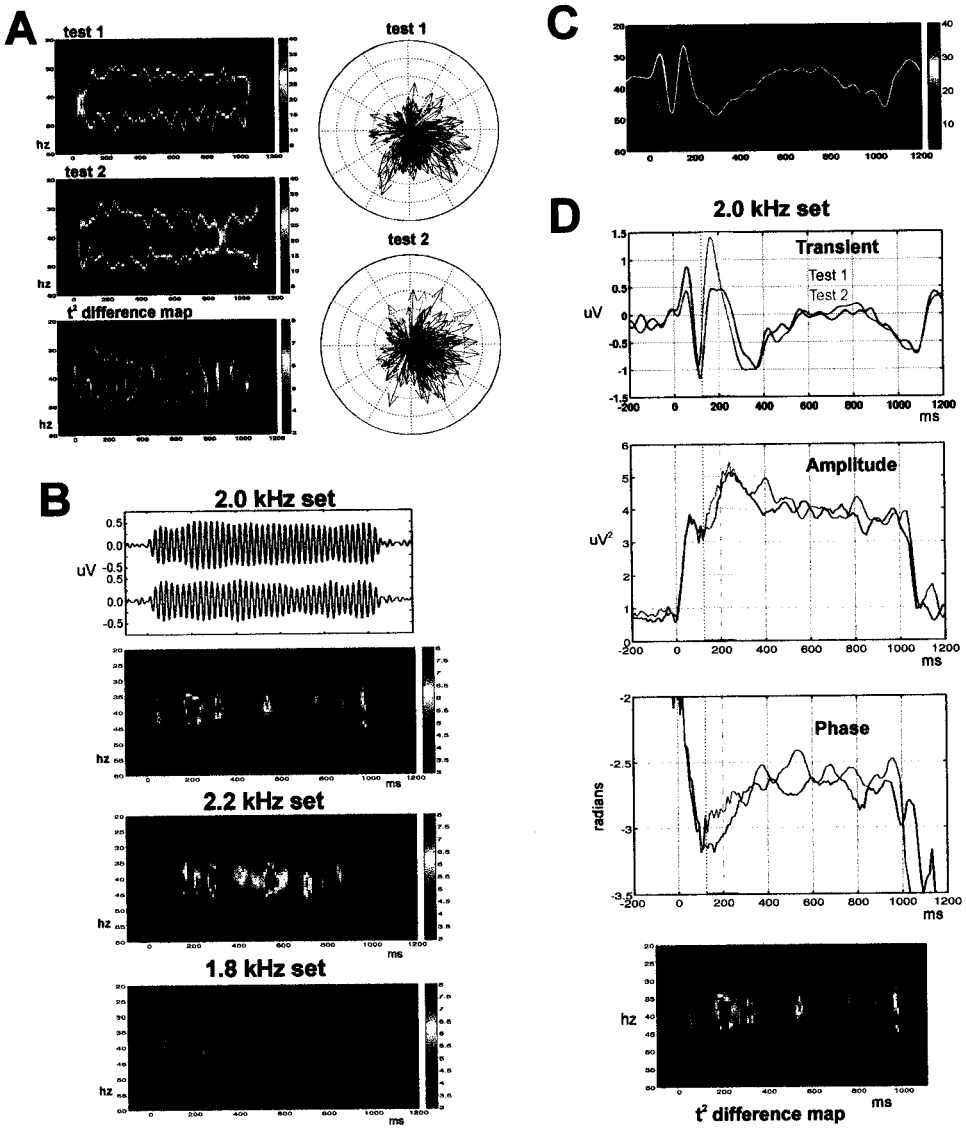
Chapter 4 Figure 1



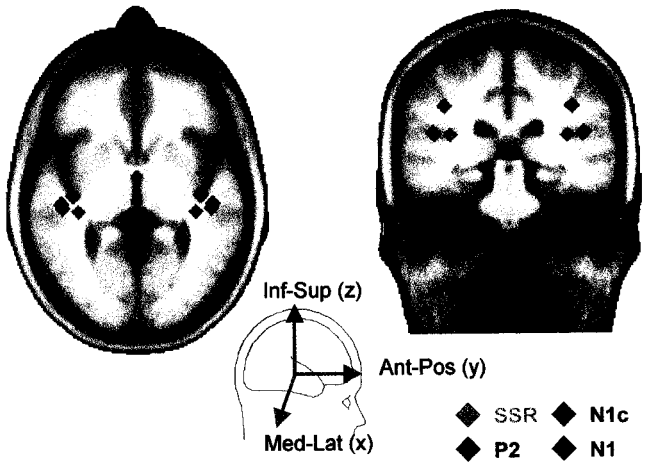
Chapter 4 Figure 2



Chapter 4 Figure 3



Chapter 4 Figure 4



Chapter 4 Figure 5

Chapter 5

SUMMARY AND CONCLUSION

When a sound is presented, a cascade of neural activity occurs within the auditory cortex as information regarding the stimulus is extracted and processed by multiple regions of the brain. It is by means of such processing that the brain encodes and represents dynamic sensory environments. The activity of populations of neurons involved in stimulus processing gives rise to auditory evoked potentials (AEPs) which can be recorded in the electroencephalogram (EEG) from electrodes placed on the scalp. This thesis has investigated the mechanisms and dynamics of the auditory steady state response, which is an AEP observed when acoustic stimuli are presented at rapid repetition rates.

There are two reasons in particular for focussing on the steady-state response (SSR). First, the SSR is a practical tool for investigating auditory processing. Data are gathered quickly and the signal-to-noise ratio can be enhanced by signal processing in the frequency domain. Second, unlike most other components of the AEP, the SSR appears to be generated by neural sources in the region of Heschl's gyrus where the tonotopic maps of primary auditory cortex are found. Hence study of the SSR can provide a picture of neural activity occuring in this region when auditory stimulus processing occurs. If a steady-state stimulus is presented in brief bursts, transient responses localizing to the secondary (belt/ parabelt) auditory cortex can also be extracted and studied. In chapter 4 of this thesis I used this property of the SSR to explore how neural activity is modified in distributed regions of the auditory cortex when subjects are trained at pitch discrimination. Experiments reported in preceding chapters described the

development of the SSR as stimulus frequency increases from the transient to the steady state range and enquired into the mechanism of its generation.

Mechanism of the SSR

Advantageous properties of steady state responding have been exploited by researchers for some time without a full understanding of the mechanism of its generation. Debate on this topic has focused on the question of whether the peak in steady state amplitude at 40 Hz is due to some possibly ubiquitous resonant property of cortical circuits at or near this rate, or is more simply accounted for by superimposition of transient responses that are somewhat rate-invariant.

The experiments in this thesis were founded on the idea of studying neuroplastic changes in auditory cortex noninvasively in humans using AEPs recorded in the EEG. The AEP contains various components which have been associated with anatomically and functionally differentiable areas of auditory cortex, although this description has been complicated somewhat by findings showing that some of these components may actually consist of several overlapping potentials (Onitsuka, Ninomiya et al., 2000; Gutschalk, Mase et al., 1999). The rate at which the auditory stimuli are presented changes the overall response, as each of the individual potentials responds to increasing stimulus rate with different recovery cycle characteristics (Erwin and Buchwald, 1986a). At higher stimulus rates, the later response components cease to occur at all, and the response transitions from transient to steady-state responding.

The first set of experiments (Chapter 2) investigated this transition. The paradigm used here, presenting the stimuli in trains with long silent periods between them,

facilitated the simultaneous study of a number of characteristics of the response. These characteristics were the response to the initial onset of the train (the 'on' response), the offset of the train (the "off" response), and the ongoing response to the individual stimuli in the train. The results show how the 'on' and 'off' responses change with rate, with the 'off' response becoming more robust as the rate increases. Simulations of the 40-Hz frequency/amplitude characteristic from the data demonstrated that the "on" response probably does not participate in response generation at rates above 1 or 2 Hz, since the continued presence of this response would lead to a peak in the frequency amplitude characteristic at 10 Hz which is not actually seen. However, the simulations showed a noticeable peak in the response at 40 Hz for transients generated at all three rates. I also described in Chapter 2 signal processing procedures that were developed to analyze the steady state response. These procedures which make use of the T^2 analysis are able to characterize the ongoing response in the time and frequency domains even when the response is not obvious in a time-domain trace.

In Chapter 3, a more complex model for simulating a test of the frequency/amplitude characteristic was developed. This model was based on the temporal deconvolution procedure described by Gutshalk et al (1999). This experiment, however, used a swept stimulus designed to more directly elicit a response that was useful for testing the linear summation hypothesis. Using this procedure the deconvolution could be applied directly to the recorded waveform, rather than to an artificially constructed concatenation of several waveforms as in the original study. Additionally, this experiment utilized EEG recording whereas all previous descriptions of the deconvolution procedure were recorded via MEG. The deconvolved source waveforms in Chapter 3 resembled a

transient with only Pa and Pb components, and when used for reconvolution, were able to reconstruct a response similar to the recorded response with a peak in response amplitude near the 40 Hz stimulation rate. The reconstruction did not differ systematically from the recorded wave over any particular stimulus rate range, and this did not expose any non-linear region. This demonstrated that a gradually building resonant response was not necessary to produce the 40 Hz amplitude peak, leaving the linear summation of source transients as the simplest explanation for this amplitude peak.

It should be noted that the source transient is not a perfect match for the measured data; in particular it may underestimate the size of the Pa and b components generated at lower rates (i.e. < 20 Hz) and also may not perfectly simulate the onset period of the SSR when presenting bursts of SSR stimuli.

Dynamics of the SSR

The study of learning of perceptual tasks can provide two types of information regarding how the brain represents its sensory inputs (Wright, 2001). First, it supplies general information on the properties of learning itself, such as how performance improves with time and so on. Second, study of perceptual learning can provide information about the mechanism underlying the improved performance on the particular task. In the case of simple auditory discrimination, one might suppose that the mechanism for improvement in the particular task would be in some way related to the mechanism that performs the task. For frequency discrimination, which is understood as occurring via a place coding on the basilar membrane and extending to the neocortical tonotopic map, this supposition leads to the proposition that extensive improvement at

frequency discrimination induced by training would bring about a rearrangement of the tonotopic map. Further, one might suppose that an improvement in discrimination ability, presumably requiring more processing power to accomplish, would lead to an expansion of the cortical area devoted to the function. Frequency discrimination training at a particular frequency allows a direct test of this idea, as one would expect to find a larger area of the tonotopic map devoted to the particular trained frequency, either at the expense of other frequencies or not.

The predicted expansion of the representation for the trained frequency has in fact been found, specifically in the case of frequency discrimination training in owl monkeys where cortical area devoted to the trained frequency expanded by a factor of 5 or more (Recanzone, Schreiner et al., 1993). It has also been found in musicians, who have received extensive training at a number of auditory tasks including frequency discrimination and who show a 130% increase in gray matter volume in Heschl's gyrus compared to nonmusican controls (Schneider, Scherg et al., 2002). The somatosensory system has also shown similar enlargements in portions of the somatotopic map due to training (Recanzone, Merzenich et al., 1992) demonstrating that this type of plastic remodeling may be a universal property of basic sensory systems (for detailed reviews of the animal studies, see (Edeline, 1999; Buonomano and Merzenich, 1998a)).

Cortical reorganization on the scale observed in animal experiments suggests that reorganization should also be expressed in human AEPs. In Chapter 4 I tested this hypothesis by training human subjects in a frequency discrimination task. The stimuli used were bursts of steady state stimuli, and the techniques used in the previous chapters are used to analyze the results. The transient portion of the response, extracted by

filtering, showed a robust increase in the magnitude of the P2 and the right-hemisphere N1c components. These changes indicated a recruitment of additional neurons in AII to the response. The P2 response change is consistent with other findings in the literature (Tremblay, Kraus et al., 2001; Atienza, Cantero et al., 2002; Eaton and Roberts, 1999); the enhanced N1c after training is a new finding that has been reported only in this paper and another of our studies (Shahin, Bosnyak et al., in press) which found enhanced P2 and N1c AEPs in musicians compared to nonmusicians. Changes in AI were measured by extracting the SSR portion of the waveform. Simple time-domain analysis of the waveforms was not powerful enough to indicate any change in the amplitude or phase of the SSR. However, the T^2 analysis was able to show that the SSR was changed due to training for short periods during the presentation of the S1 stimulus, most notably during a period 150-225 ms post S1 onset coincident with the rising edge of the P2 but extending beyond this brain event. Further investigation showed that this change was due to a change in phase of the SSR during this window and not a change in amplitude. This suggests that the number of AI neurons tuned to the trained 2 kHz stimulus was not increased, but rather that the temporal response properties of neurons in this area were modified possibly by a dynamic change in inputs arising from other brain areas.

Future Directions

The description of the SSR as the result of a source transient leaves open a number of experimental possibilities. The two components of the source transient likely correspond to the Pa and Pb components of the AEP. This correspondence could be verified by a manipulation that dissociates the Pa and Pb components, for example by

manipulating the subject's sleep state during recording, which should differentially affect the two components. Such a demonstration would increase confidence in the assertion that the source transient and the middle latency components are generated by the same process.

The frequency discrimination training described in Chapter 4 was able to produce a change in the AEP response coming from secondary auditory cortex but did not produce a measurable increase in the amplitude of the ongoing SSR that would be expected if an expansion in cortical area for processing the training frequency had been induced. However, musicians have been shown to have enhanced SSR amplitudes which is likely due to long term training-induced changes. Although the nature of this change is being investigated via longitudinal and cross-sectional studies of the training of young musicians (Trainor, Shahin et al., in press), producing such a change in the SSR in laboratory based training procedure is still a goal. The failure to produce such a change here might be due to inadequate training volume, but increasing the length of the training program from the 18 sessions presented here is problematic. A more fruitful approach would be to change the nature of the training to one which might be more likely to produce a change in the SSR.

Ideally, experiments following this latter approach should take guidance from principles that describe remodeling of auditory cortical representations by experience. Kilgard and his colleagues have begun to investigate the rules governing synaptic changes in the primary auditory cortex of the rat when auditory stimuli are paired with stimulation of the basal forebrain (a reward area) by indwelling electrodes. Within the limits of their preparation and the parameters studied, map expansions and increases in

the size of the receptive fields of AI neurons were favored by auditory signals consisting of a single spectral frequency, whereas neural representations tended to segregate (smaller receptive fields and little change in cortical territory representing a trained frequency) when multiple frequencies were used. These effects also interacted with temporal modulation of the signal. In general, faster AM rates led to expansion of receptive fields and enhanced the frequency-following characteristics of AI neurons. However, within a given AM rate, adding spectral variability to the signal (multiple tones) favored segregation rather than receptive field expansion. These findings suggest that because subjects in the study of Chapter 4 typically experienced 10 different spectral frequencies between 2.0 kHz and 2.1 kHz within a training session, expansion of the tonotopic representation in AI may have been restricted by competitive interactions such that only temporal properties of the representation were modified and expressed in SSR phase. If so, my results may represent only one of a range of adaptations that may be possible and reflected in the SSR.

A more complete understanding of the dynamics of the SSR could be provided by future studies designed to take into account a principle of competitive interactions. As a starting point for such study, SSR dynamics could be compared between procedures in which subjects are trained at acoustic discriminations using single versus multiple carrier frequencies. If the animal data provide a reliable guide, amplitude enhancements are more likely to be observed in the single carrier case. One suitable method may be to train subjects to detect amplitude enhancement of a single steady-state pulses in a train of pulses, using either single or multiple carrier frequencies. A different approach could be to look for enhancement of SSR amplitude when subjects are trained to detect the onset

of 40 Hz amplitude-modulation (nonzero modulation depth) using only a single carrier frequency. These are but two examples of procedures that may reveal dynamics in the SSR (and in cortical representations in AI) extending to amplitude as well as phase. The value of such studies would be twofold. First, they would help us to understand the properties of the SSR and the mechanisms underlying its generation. Second, they would provide information about the rules that govern experience-induced plasticity in the human auditory system.

REFERENCES

ADAMS JC. Ascending projections to the inferior colliculus. *J Comp Neurol,* 183: 519-38, 1979.

AOYAGI M, KIREN T, KIM Y, SUZUKI Y, FUSE T, and KOIKE Y. Optimal modulation frequency for amplitude-modulation following response in young children during sleep. *Hear Res,* 65: 253-61, 1993.

ATIENZA M, CANTERO JL, and DOMINGUEZ-MARIN E. The time course of neural changes underlying auditory perceptual learning. *Learn Mem,* 9: 138-50, 2002.

AZZENA GB, CONTI G, SANTARELLI R, OTTAVIANI F, PALUDETTI G, and MAURIZI M. Generation of human auditory steady-state responses (ssrs). I: Stimulus rate effects. *Hear Res,* 83: 1-8, 1995.

BOURK TR, MIELCARZ JP, and NORRIS BE. Tonotopic organization of the anteroventral cochlear nucleus of the cat. *Hear Res,* 4: 215-41, 1981.

BULLOCK TH, KARAMURSEL S, ACHIMOWICZ JZ, MCCLUNE MC, and BASAR-EROGLU C. Dynamic properties of human visual evoked and omitted stimulus potentials. *Electroencephalogr Clin Neurophysiol,* 91: 42-53, 1994.

BUONOMANO DV. Decoding temporal information: A model based on short-term synaptic plasticity. *J Neurosci,* 20: 1129-41, 2000.

BUONOMANO DV and MERZENICH MM. Cortical plasticity: From synapses to maps. *Annu Rev Neurosci,* 21: 149-86, 1998a.

BUONOMANO DV and MERZENICH MM. Net interaction between different forms of short-term synaptic plasticity and slow-ipsps in the hippocampus and auditory cortex. *J Neurophysiol,* 80: 1765-74, 1998b.

BURKARD R, SHI Y, and HECOX KE. Brain-stem auditory-evoked responses elicited by maximum length sequences: Effect of simultaneous masking noise. *J Acoust Soc Am,* 87: 1665-72, 1990a.

BURKARD R, SHI Y, and HECOX KE. A comparison of maximum length and legendre sequences for the derivation of brain-stem auditory-evoked responses at rapid rates of stimulation. *J Acoust Soc Am,* 87: 1656-64, 1990b.

CARVER FW, FUCHS A, JANTZEN KJ, KELSO JA. Spatiotemporal analysis of the neuromagnetic response to rhythmic auditory stimulation: rate dependence and transient to steady-state transition. *Clin Neurophysiol,* 113:1921-31, 2002.

CHATTERJEE M and ZWISLOCKI JJ. Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. *Hear Res,* 111: 65-75, 1997.

CLOPTON BM and WINFIELD JA. Tonotopic organization in the inferior colliculus of the rat. *Brain Res,* 56: 355-8, 1973.

CONTI G, SANTARELLI R, GRASSI C, OTTAVIANI F, and AZZENA GB. Auditory steady-state responses to click trains from the rat temporal cortex. *Clin Neurophysiol,* 110: 62-70, 1999.

COOPER NP, ROBERTSON D, and YATES GK. Cochlear nerve fiber responses to amplitude-modulated stimuli: Variations with spontaneous rate and other response characteristics. *J Neurophysiol,* 70: 370-86, 1993.

DEIBER MP, IBANEZ V, FISCHER C, PERRIN F, and MAUGUIERE F. Sequential mapping favours the hypothesis of distinct generators for na and pa middle latency auditory evoked potentials. *Electroencephalogr Clin Neurophysiol,* 71: 187-97, 1988.

DIETRICH V, NIESCHALK M, STOLL W, RAJAN R, and PANTEV C. Cortical reorganization in patients with high frequency cochlear hearing loss. *Hear Res,* 158: 95-101, 2001.

DIMYAN MA and WEINBERGER NM. Basal forebrain stimulation induces discriminative receptive field plasticity in the auditory cortex. *Behav Neurosci,* 113: 691-702, 1999.

DOBIE RA and WILSON MJ. Optimal ('wiener') digital filtering of auditory evoked potentials: Use of coherence estimates. *Electroencephalogr Clin Neurophysiol,* 77: 205-13, 1990.

EATON RA and ROBERTS LE. Effect of spectral frequency discrimination on auditory transient and steady state responses in humans. *Society for Neurosciences Abstracts,* 29: 156.15, 1999.

EDELINE JM. Learning-induced physiological plasticity in the thalamo-cortical sensory systems: A critical evaluation of receptive field plasticity, map changes and their potential mechanisms. *Prog Neurobiol,* 57: 165-224, 1999.

ELBERT T, STERR A, ROCKSTROH B, PANTEV C, MULLER MM, and TAUB E. Expansion of the tonotopic area in the auditory cortex of the blind. *J Neurosci,* 22: 9941-4, 2002.

ERWIN R and BUCHWALD JS. Midlatency auditory evoked responses: Differential effects of sleep in the human. *Electroencephalogr Clin Neurophysiol,* 65: 383-92, 1986a.

ERWIN RJ and BUCHWALD JS. Midlatency auditory evoked responses: Differential recovery cycle characteristics. *Electroencephalogr Clin Neurophysiol,* 64: 417-23, 1986b.

EYSHOLDT U and SCHREINER C. Maximum length sequences -- a fast method for measuring brain-stem-evoked responses. *Audiology,* 21: 242-50, 1982.

FOX K. A critical period for experience-dependent synaptic plasticity in rat barrel cortex. *J Neurosci,* 12: 1826-38, 1992.

FRANOWICZ MN and BARTH DS. Comparison of evoked potentials and high-frequency (gamma-band) oscillating potentials in rat auditory cortex. *J Neurophysiol,* 74: 96-112, 1995.

GALAMBOS R. Tactile and auditory stimuli repeated at high rates (30-50 per sec) produce similar event related potentials. *Ann N Y Acad Sci,* 388: 722-8, 1982.

GALAMBOS R, MAKEIG S, and TALMACHOFF PJ. A 40-hz auditory potential recorded from the human scalp. *Proc Natl Acad Sci U S A,* 78: 2643-7, 1981.

GILRON I, PLOURDE G, MARCANTONI W, and VARIN F. 40 hz auditory steady-state response and eeg spectral edge frequency during sufentanil anaesthesia. *Can J Anaesth,* 45: 115-21, 1998.

GODEY B, SCHWARTZ D, DE GRAAF JB, CHAUVEL P, and LIEGEOIS-CHAUVEL C. Neuromagnetic source localization of auditory evoked fields and intracerebral evoked potentials: A comparison of data in the same patients. *Clin Neurophysiol,* 112: 1850-9, 2001.

GOLDSTEIN R, RODMAN LB, and KARLOVICH RS. Effects of stimulus rate and number on the early components of the averaged electroencephalic response. *J Speech Hear Res,* 15: 559-66, 1972.

GRAY CM, KONIG P, ENGEL AK, and SINGER W. Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature,* 338: 334-7, 1989.

GRAY CM and SINGER W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc Natl Acad Sci U S A,* 86: 1698-702, 1989.

GUTSCHALK A, MASE R, ROTH R, ILLE N, RUPP A, HAHNEL S, PICTON TW, and SCHERG M. Deconvolution of 40 hz steady-state fields reveals two overlapping source activities of the human auditory cortex. *Clin Neurophysiol,* 110: 856-68, 1999.

HACKETT TA, PREUSS TM, and KAAS JH. Architectonic identification of the core region in auditory cortex of macaques, chimpanzees, and humans. *J Comp Neurol,* 441: 197-222, 2001.

HACKETT TA, STEPNIEWSKA I, and KAAS JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. *J Comp Neurol,* 394: 475-95, 1998.

HARI R, HAMALAINEN M, and JOUTSINIEMI SL. Neuromagnetic steady-state responses to auditory stimuli. *J Acoust Soc Am,* 86: 1033-9, 1989.

HELMHOLTZ HLF. *On the sensations of tone (2nd english ed.).* New York: Dover, 1877/1954.

HOWARD MA, 3RD, VOLKOV IO, ABBAS PJ, DAMASIO H, OLLENDIECK MC, and GRANNER MA. A chronic microelectrode investigation of the tonotopic organization of human auditory cortex. *Brain Res,* 724: 260-4, 1996.

HUBEL DH and WIESEL TN. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *J Physiol,* 206: 419-36, 1970.

IMIG TJ and MOREL A. Topographic and cytoarchitectonic organization of thalamic neurons related to their targets in low-, middle-, and high-frequency representations in cat auditory cortex. *J Comp Neurol,* 227: 511-39, 1984.

IMIG TJ, RUGGERO MA, KITZES LM, JAVEL E, and BRUGGE JF. Organization of auditory cortex in the owl monkey (aotus trivirgatus). *J Comp Neurol,* 171: 111-28, 1977.

JOHN MS, DIMITRIJEVIC A, and PICTON TW. Weighted averaging of steady-state responses. *Clin Neurophysiol,* 112: 555-62, 2001.

JOHN MS, LINS OG, BOUCHER BL, and PICTON TW. Multiple auditory steady-state responses (master): Stimulus and recording parameters. *Audiology,* 37: 59-82, 1998.

KAAS JH and HACKETT TA. Subdivisions of auditory cortex and levels of processing in primates. *Audiol Neurootol,* 3: 73-85, 1998.

KAAS JH and HACKETT TA. Subdivisions of auditory cortex and processing streams in primates. *Proc Natl Acad Sci U S A,* 97: 11793-9, 2000.

KAKIGI A, HIRAKAWA H, HAREL N, MOUNT RJ, and HARRISON RV. Tonotopic mapping in auditory cortex of the adult chinchilla with amikacin-induced cochlear lesions. *Audiology,* 39: 153-60, 2000.

KELLY JB, LISCUM A, VAN ADEL B, and ITO M. Projections from the superior olive and lateral lemniscus to tonotopic regions of the rat's inferior colliculus. *Hear Res,* 116: 43-54, 1998.

KILGARD MP, PANDYA PK, VAZQUEZ J, GEHI A, SCHREINER CE, and MERZENICH MM. Sensory input directs spatial and temporal plasticity in primary auditory cortex. *J Neurophysiol,* 86: 326-38, 2001.

LINS OG, PICTON PE, PICTON TW, CHAMPAGNE SC, and DURIEUX-SMITH A. Auditory steady-state responses to tones amplitude-modulated at 80-110 hz. *J Acoust Soc Am,* 97: 3051-63, 1995.

LIPPE WR. Recent developments in cochlear physiology. *Ear Hear,* 7: 233-9, 1986.

MAKELA JP, HAMALAINEN M, HARI R, and MCEVOY L. Whole-head mapping of middle-latency auditory evoked magnetic fields. *Electroencephalogr Clin Neurophysiol,* 92: 414-21, 1994.

MCFARLAND WH, VIVION MC, WOLF KE, and GOLDSTEIN R. Reexamination of effects of stimulus rate and number on the middle components of the averaged electroencephalic response. *Audiology,* 14: 456-65, 1975.

MERZENICH MM and BRUGGE JF. Representation of the cochlear partition of the superior temporal plane of the macaque monkey. *Brain Res,* 50: 275-96, 1973.

MOGILNER A, GROSSMAN JA, RIBARY U, JOLIOT M, VOLKMANN J, RAPAPORT D, BEASLEY RW, and LLINAS RR. Somatosensory cortical plasticity in adult humans revealed by magnetoencephalography. *Proc Natl Acad Sci U S A,* 90: 3593-7, 1993.

MOREL A, GARRAGHTY PE, and KAAS JH. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. *J Comp Neurol,* 335: 437-59, 1993.

ONITSUKA T, NINOMIYA H, SATO E, YAMAMOTO T, and TASHIRO N. The effect of interstimulus intervals and between-block rests on the auditory evoked potential and magnetic field: Is the auditory p50 in humans an overlapping potential? *Clin Neurophysiol,* 111: 237-45, 2000.

PANTEV C, BERTRAND O, EULITZ C, VERKINDT C, HAMPSON S, SCHUIERER G, and ELBERT T. Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalogr Clin Neurophysiol,* 94: 26-40, 1995.

PANTEV C, ELBERT T, MAKEIG S, HAMPSON S, EULITZ C, and HOKE M. Relationship of transient and steady-state auditory evoked fields. *Electroencephalogr Clin Neurophysiol,* 88: 389-96, 1993.

PANTEV C, HOKE M, LEHNERTZ K, LUTKENHONER B, ANOGIANAKIS G, and WITTKOWSKI W. Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. *Electroencephalogr Clin Neurophysiol,* 69: 160-70, 1988.

PANTEV C, MAKEIG S, HOKE M, GALAMBOS R, HAMPSON S, and GALLEN C. Human auditory evoked gamma-band magnetic fields. *Proc Natl Acad Sci U S A,* 88: 8996-9000, 1991.

PANTEV C, OOSTENVELD R, ENGELIEN A, ROSS B, ROBERTS LE, and HOKE M. Increased auditory cortical representation in musicians. *Nature,* 392: 811-4, 1998.

PANTEV C, ROBERTS LE, ELBERT T, ROSS B, and WIENBRUCH C. Tonotopic organization of the sources of human auditory steady-state responses. *Hear Res,* 101: 62-74, 1996.

PELIZZONE M, HARI R, MAKELA JP, HUTTUNEN J, AHLFORS S, and HAMALAINEN M. Cortical origin of middle-latency auditory evoked responses in man. *Neurosci Lett,* 82: 303-7, 1987.

PICTON TW, CHAMPAGNE SC, and KELLETT AJ. Human auditory evoked potentials recorded using maximum length sequences. *Electroencephalogr Clin Neurophysiol,* 84: 90-100, 1992.

PICTON TW, HILLYARD SA, KRAUSZ HI, and GALAMBOS R. Human auditory evoked potentials. I. Evaluation of components. *Electroencephalogr Clin Neurophysiol,* 36: 179-90, 1974.

PICTON TW, VAJSAR J, RODRIGUEZ R, and CAMPBELL KB. Reliability estimates for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol,* 68: 119-31, 1987.

RAJAN R and IRVINE DR. Absence of plasticity of the frequency map in dorsal cochlear nucleus of adult cats after unilateral partial cochlear lesions. *J Comp Neurol,* 399: 35-46, 1998.

RAUSCHECKER JP. Processing of complex sounds in the auditory cortex of cat, monkey, and man. *Acta Otolaryngol Suppl,* 532: 34-8, 1997.

RAUSCHECKER JP, TIAN B, PONS T, and MISHKIN M. Serial and parallel processing in rhesus monkey auditory cortex. *J Comp Neurol,* 382: 89-103, 1997.

RECANZONE GH, MERZENICH MM, JENKINS WM, GRAJSKI KA, and DINSE HR. Topographic reorganization of the hand representation in cortical area 3b owl monkeys trained in a frequency-discrimination task. *J Neurophysiol,* 67: 1031-56, 1992.

RECANZONE GH, SCHREINER CE, and MERZENICH MM. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. *J Neurosci,* 13: 87-103, 1993.

REGAN D. Comparison of transient and steady-state methods. *Ann N Y Acad Sci,* 388: 45-71, 1982.

REGAN D. *Human brain electrophysiology: Evoked potentials and evoked magnetic fields in science and medicine.* Amsterdam: Elsevier, 1989.

ROMANSKI LM, BATES JF, and GOLDMAN-RAKIC PS. Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J Comp Neurol,* 403: 141-57, 1999.

SACHS MB. Neural coding of complex sounds: Speech. *Annu Rev Physiol,* 46: 261-73, 1984.

SANES DH and CONSTANTINE-PATON M. The sharpening of frequency tuning curves requires patterned activity during development in the mouse, mus musculus. *J Neurosci,* 5: 1152-66, 1985.

SANTARELLI R and CONTI G. Generation of auditory steady-state responses: Linearity assessment. *Scand Audiol Suppl,* 51: 23-32, 1999.

SANTARELLI R, MAURIZI M, CONTI G, OTTAVIANI F, PALUDETTI G, and PETTOROSSI VE. Generation of human auditory steady-state responses (ssrs). Ii: Addition of responses to individual stimuli. *Hear Res,* 83: 9-18, 1995.

SCHNEIDER P, SCHERG M, DOSCH HG, SPECHT HJ, GUTSCHALK A, and RUPP A. Morphology of heschl's gyrus reflects enhanced activation in the auditory cortex of musicians. *Nat Neurosci,* 5: 688-94, 2002.

SCHREINER CE, READ HL, and SUTTER ML. Modular organization of frequency integration in primary auditory cortex. *Annu Rev Neurosci,* 23: 501-29, 2000.

SEKI S and EGGERMONT JJ. Changes in cat primary auditory cortex after minor-to-moderate pure-tone induced hearing loss. *Hear Res,* 173: 172-86, 2002.

SHAHIN S, BOSNYAK DJ, TRAINOR LJ, and ROBERTS LE. Enhancement of neuroplastic p2 and n1c auditory evoked potentials in musicians. *J Neurosci*, in press.

STAPELLS DR, GALAMBOS R, COSTELLO JA, and MAKEIG S. Inconsistency of auditory middle latency and steady-state responses in infants. *Electroencephalogr Clin Neurophysiol,* 71: 289-95, 1988.

STAPELLS DR, LINDEN D, SUFFIELD JB, HAMEL G, and PICTON TW. Human auditory steady state potentials. *Ear Hear,* 5: 105-13, 1984.

STAPELLS DR, MAKEIG S, and GALAMBOS R. Auditory steady-state responses: Threshold prediction using phase coherence. *Electroencephalogr Clin Neurophysiol,* 67: 260-70, 1987.

STAPELLS DR and PICTON TW. Technical aspects of brainstem evoked potential audiometry using tones. *Ear Hear,* 2: 20-9, 1981.

TANG Y and NORCIA AM. Improved processing of the steady-state evoked potential. *Electroencephalogr Clin Neurophysiol,* 88: 323-34, 1993.

TANG Y and NORCIA AM. An adaptive filter for steady-state evoked responses. *Electroencephalogr Clin Neurophysiol,* 96: 268-77, 1995.

TECCHIO F, BICCIOLO G, DE CAMPORA E, PASQUALETTI P, PIZZELLA V, INDOVINA I, CASSETTA E, ROMANI GL, and ROSSINI PM. Tonotopic cortical changes following stapes substitution in otosclerotic patients: A magnetoencephalographic study. *Hum Brain Mapp,* 10: 28-38, 2000.

TRAINOR LJ, SHAHIN S, and ROBERTS LE. Effects of musical training on auditory cortex in children. *Ann N Y Acad Sci*, in press.

TREMBLAY K, KRAUS N, MCGEE T, PONTON C, and OTIS B. Central auditory plasticity: Changes in the n1-p2 complex after speech-sound training. *Ear Hear,* 22: 79-90, 2001.

VALDES-SOSA MJ, BOBES MA, PEREZ-ABALO MC, PERERA M, CARBALLO JA, and VALDES-SOSA P. Comparison of auditory-evoked potential detection methods using signal detection theory. *Audiology,* 26: 166-78, 1987.

VICTOR JD and MAST J. A new statistic for steady-state evoked potentials. *Electroencephalogr Clin Neurophysiol,* 78: 378-88, 1991.

VON BEKESY G. *Experiments in hearing*. New York: McGraw-Hill, 1960.

WEINBERGER NM and BAKIN JS. Learning-induced physiological memory in adult primary auditory cortex: Receptive fields plasticity, model, and mechanisms. *Audiol Neurootol,* 3: 145-67, 1998.

WEINBERGER NM, JAVID R, and LEPAN B. Long-term retention of learning-induced receptive-field plasticity in the auditory cortex. *Proc Natl Acad Sci U S A,* 90: 2394-8, 1993.

WRIGHT BA. Why and how we study human learning on basic auditory tasks. *Audiol Neurootol,* 6: 207-10, 2001.

YAN J and EHRET G. Corticofugal reorganization of the midbrain tonotopic map in mice. *Neuroreport,* 12: 3313-6, 2001.

YVERT B, CROUZEIX A, BERTRAND O, SEITHER-PREISLER A, and PANTEV C. Multiple supratemporal sources of magnetic and electric auditory evoked middle latency components in humans. *Cereb Cortex,* 11: 411-23, 2001.

ZHANG LI, BAO S, and MERZENICH MM. Persistent and specific influences of early acoustic environments on primary auditory cortex. *Nat Neurosci,* 4: 1123-30, 2001.