# The Neural Basis of Speech Perception

by

David Poeppel
S.B., Cognitive Science
Massachusetts Institute of Technology, 1990

Submitted to the Department of Brain and Cognitive Sciences
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
in Cognitive Neuroscience
at the
Massachusetts Institute of Technology

September 1995

Signature of author ____________________
Department of Brain and Cognitive Sciences
July 28, 1995

Certified by ____________________
Kenneth Wexler
Professor, Thesis Supervisor

Certified by ____________________
David Caplan
Associate Professor, Thesis Co-Supervisor

Accepted by ____________________
Emilio Bizzi
Chair, Department of Brain & Cognitive Sciences

# The Neural Basis of Speech Perception

by

David Poeppel

Submitted to the Department of Brain & Cognitive Sciences
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Cognitive Neuroscience
at the
Massachusetts Institute of Technology
September 1995

## Abstract

The standard view of the functional neuroanatomy of speech processing attributes the neural basis of speech perception to the left posterior superior temporal cortex, in line with the doctrine that a number of areas in the left hemisphere are specialized for the representation and processing of speech and language. The research presented in this dissertation reevaluates this model. By systematically distinguishing between the neural basis for speech perception proper (construed as one interface with the language processing system and hence a subroutine of auditory language comprehension) and the neural basis for other aspects of language representation and processing, it is demonstrated that both the left and right superior temporal gyri (non-primary auditory cortex) are implicated in the construction of speech sounds. Evidence that speaks to the question of the neural basis of speech perception is drawn from three sources: the deficit-lesion correlation literature on an auditory verbal agnosia called pure word deafness (Chapter 2), the literature on brain activity measurements using metabolic functional imaging methods such as positron emission tomography (Chapter 3), and experimental results derived from the new electromagnetic brain imaging technique magnetoencephalography (Chapter 4). Motivated by the observations from the deficit-lesion and the neuroimaging literature, the thesis develops the model that speech perception is mediated by the posterior superior temporal gyri in both hemispheres. Two hypotheses about the speech perception system are experimentally investigated. The first hypothesis is that one can reliably detect left and right supratemporal cortical activation generated by speech sounds, and the second that the execution of task demands differentially affects the two hemispheres. These two questions were evaluated in three magnetoencephalography studies recording the auditory evoked neuromagnetic field generated by tones, vowels, and syllables. In passive listening conditions, the evoked response data as well as the source localization data revealed no systematic differences between left and right auditory association cortices, supporting the model that both areas contribute to the processing of speech sounds. The comparison between a task-driven and a passive condition showed that the right and left auditory association cortices responded differently to attentional requirements: compared to a passive listening condition, the attended condition yielded simultaneous response amplitude increases in the left and decreases in the right in the same supratemporal cortical area. This research concludes that both left and right posterior superior temporal gyri play significant roles in the processing of speech sounds, although the precise nature of the computations remains unspecified.

Thesis Supervisor: Kenneth Wexler
Title: Professor

Thesis Co-Supervisor: David Caplan
Title: Associate Professor

# Acknowledgments

I would like to acknowledge the contributions of Kevin Broihier, David Caplan, Albert Galaburda, Zoubin Ghahramani, Alec Marantz, David Pesetsky, Colin Phillips, Steve Pinker, Ken Wexler, and Elron Yellin.

It was a privilege to have as classmates and graduate school friends Sergey Avrutin, Daphne Bavelier, Emanuela Bricolo, Kevin Broihier, Jenny Ganger, Zoubin Ghahramani, John Houde, Gavin Huntley-Fenner, John Kim, Adee Matan, Kelly Olguin, Sandeep Prasada, Rose Roberts, Flip Sabes, Cristina Sorrentino, Annie Senghas, William Snyder, Josh Stern, Fei Xu, and many others.

Thanks are also due to Pat Claffey, Ellie Bonsaint, Leonardo Topa, and Robin Nahmias for providing important aspects of the logistical framework required to successfully exist in E10.

As for my family, one couldn't wish for a better one. I was (and am) insufferable, but my wife Amy and my son Alexander continue to be the best family anyone could hope for. My sisters Julie and Lili always help me keep things in perspective. My parents guided me along the way, each in their own unique manner. And my extended family, including my parents-in-law and sisters-in-law, were always encouraging and supportive. My thanks to all of them.

To Amy and Alexander

# Table of Contents

## List of Figures

## List of Tables

# CHAPTER ONE

## BRAIN AND LANGUAGE: THE STANDARD MODEL AND ITS PROBLEMS

Where and how are speech and language represented and processed in the human brain? Do specific brain areas mediate parts of language processing like *sentence comprehension* or *auditory word recognition*? Are there specialized brain regions for processing the *sounds of language* or *word structure* or *compositional meaning*? Although these are old questions, in the last few years the problem of localizing the components of language in the human brain has generated considerable new excitement. One important reason is that a number of new technologies, collectively known as *functional neuroimaging methods*, allow for a refreshing new approach to the problem.

The more general question - Which parts of the brain mediate which parts of behavior? - is of course not new. Developing a *functional neuroanatomy* of the human brain is an established scientific problem at the center of trying to understand the relation between brain and behavior (Teuber, 1978). The fundamental assumption underlying functional localization is that the components of behavior, ranging from sensory and motor aspects of behavior to cognitive faculties like memory and language, are represented and processed more or less independently and can be attributed to particular areas of the brain. Although it is intuitively compelling to believe that such different psychological processes as visual perception and memory, language and motor control are as independent materially as they are distinct introspectively, it needed to be established that distinct mental processes are in fact mediated by different parts of the central nervous system. With respect to sensory and motor systems, the correspondence between the functional decomposition and cerebral functional specialization has proved to be an exceedingly successful notion. We now have detailed models of where and how visual information is represented, where the motor commands to move one's legs or eyes are executed, which part of the brain is required in order to be able to hear sound, and so on (Gazzaniga, 1995; Kandel, Schwartz, and Jessel, 1991; Kolb and Wishaw 1985).

The fact that it has been possible to use animal models has constituted a major advantage in the study of sensation and motor control. In contrast, the functional localization of cognitive systems has been less successful, one important reason being the absence of animal models. For instance, although we know the location and basic structure of the human primary auditory cortex, we do not know where and how speech perception proper takes place. Although primary auditory cortex is a necessary component, it is not sufficient to mediate the construction of speech sounds - other cortical areas are necessary as well. But which other areas are necessary in addition to the primary projection areas is unknown, although numerous candidates are up for discussion. In short, while for many cognitive functions it has been possible to draw from detailed animal models, this is obviously not possible for certain cognitive systems, including speech and language processing.

Until quite recently there was only one major method to investigate which part of the brain processes which cognitive operation, the deficit-lesion-correlation approach: one studies patients who have specific cognitive deficits due to brain injury and attempts to correlate their behavioral dysfunction with the site of their brain lesion. This approach has yielded most of what we know about the functional anatomy of language. But the approach has a number of disadvantages, including that one is limited to investigating compromised (diseased or injured) nervous tissue. Therefore the harnessing of new medical technologies for behavioral research has been met with great enthusiasm. The functional neuroimaging methods make it possible to study the intact brain of a person while he or she is engaged in performing some specified cognitive task. These methods have captured the scientific and popular imaginations, and it is now common to come across reports of 'visualizing the brain at work' or 'seeing the mind' using new technologies (Orrison et al., 1995; Posner & Raichle, 1994).

Neuroimaging research includes structural and functional imaging. Both approaches are being applied to issues of speech and language processing. *Structural neuroimaging research* uses methods like computerized tomography (CT) and magnetic resonance imaging (MRI) to visualize the macroscopic structure of the brain in as much spatial detail as possible. If one is trying to map cognitive functions onto anatomical structures by correlating behavioral deficits with an anatomical abnormalities (the standard deficit-lesion correlation approach of neuropsychology mentioned briefly above), structural imaging provides the anatomical information.[1]

*Functional neuroimaging research* can be categorized into two kinds of imaging strategies, metabolic imaging and electromagnetic imaging. In metabolic imaging one takes advantage of the fact that there is a close relationship between increased neuronal activity in a brain area and particular metabolic-physiological indices like blood flow, blood volume, glucose consumption, or the oxygenation of hemoglobin. The metabolically based functional neuroimaging methods include positron emission tomography (PET), single

---

[1] The occasional attempts to correlate directly structural information with cognition have generated some very provocative results. For example, a recent MRI study by Schlaug et al. (1995) suggests that musicians with perfect pitch (the ability to identify and label a presented sound of a given frequency) differ from musicians and non-musicians without perfect pitch in the relative hemispheric asymmetry of the planum temporale. Those subjects with perfect pitch showed a very pronounced asymmetry between the left and right plana, with the left being considerably larger. It is unclear how to interpret this finding, but, one way or another, one must account for the correlation between the gross anatomy and the curious behavioral phenomenon. Another example of this research strategy is the work of Elena Plante and her colleagues. They correlate abnormal patterns of temporal lobe asymmetries with specific impairments of the speech and language system (Plante et al, 1991a, 1991b).

photon emission computerized tomography (SPECT), and functional magnetic resonance imaging (fMRI). The electromagnetically based functional neuroimaging methods include the older and well-established method of electroencephalography (EEG) and the newer magnetoencephalography (MEG). Both electromagnetic methods measure - at the scalp - the electrical or magnetic activity that is characteristic of large groups of neurons.

Both structural and functional imaging methods are being used to investigate the neural basis of speech and language. The kinds of questions that are addressed with these methods and some of the state-of-the-art findings are discussed in more detail below. To orient the discussion, it is worth first highlighting what is held to be the 'standard' view. The standard model of where the brain processes language includes three major components: *Broca's area*, typically considered to be the pars opercularis and triangularis of the inferior frontal gyrus in left frontal lobe, is taken to mediate expressive language (language production), *Wernicke's area*, which includes *at least* the left posterior superior temporal gyrus extending from approximately lateral to Heschl's gyrus posteriorly, is taken to mediate receptive language (language comprehension), and the two are connected by a fiber bundle called the *arcuate fasciculus*. Precisely what the boundaries of these putative cortical language areas are is in serious dispute.[2] This model goes back to the mid-19th century and the findings of Broca (1861) and Wernicke (1874) but has been popularized most effectively in a series of influential papers by the neurologist Norman Geschwind (Geschwind 1965, 1970, 1979). It is the model one learns about in neurology,

---

[2] The precise areas implicated are usually left unspecified or called 'auditory association cortex' or 'Wernicke's area.' The anatomical region called Wernicke's area is itself rather variably defined and coarse: typically, Wernicke's area is said to extend along the superior temporal gyrus, from a point lateral and slightly posterior to Heschl's gyrus all the way posterior to the border with the angular gyrus. But the superior temporal gyrus is (at its posterior aspect) also coterminous with the supramarginal gyrus. Moreover, given its extent, Wernicke's area presumably includes a number of functionally distinct cortical fields.

neuroscience, and neuropsychology textbooks.[3] In its original form, this model does not hold up in the face of current theoretical considerations and empirical evidence.

Consider first some elegant new structural brain imaging research that speaks to this 'classical' view. Using MRI, the aphasiologist Nina Dronkers and her colleagues have done some careful reconstruction of lesions in patients with Broca's or Wernicke's aphasia (Dronkers et al. 1993, 1995). They describe the following findings: having a lesion in Broca's area does not necessarily lead to Broca's (agrammatic) aphasia. Similarly, having a lesion in Wernicke's area does not necessarily lead to Wernicke's aphasia. In fact, Dronkers and her colleagues show that the common lesion site among their patients with Broca's aphasia is in the anterior insular cortex and the common lesion site among their patients with Wernicke's aphasia is the middle temporal gyrus in the left hemisphere. These results suggest that the tight connection postulated between language processing and Broca's and Wernicke's areas does not hold up in its most direct form. Dronkers' careful correlation between linguistic deficits and brain injury locations reveals that the connections once held to be typical are much more tenuous and subject to extreme variation.

Further empirical considerations are also problematic. Specifically, although Broca's aphasics are most severely impaired in their linguistic productive capacities, they have demonstrable comprehension deficits as well (see Zurif, in press, for an extended discussion of this point). Similarly, even fluent Wernicke's aphasics can have considerable difficulty in language production. While it is clear that both Broca's and Wernicke's areas play important roles in the representation and processing of speech and language, it has also become clear that a model that includes no more than these brain areas and that

---

[3] For more detailed summaries of influential brain-language theories, including holist and localizationist theories, see Caplan (1992, 1994).

subscribes to a very coarse model of language representation and processing cannot capture the extremely complex relationship between language and the brain.

The assumption that language is most sensibly fractionated into comprehension and production with respect to the neural organization of language (rather than, say, particular language *representations*) has not held up, both on theoretical and on empirical grounds. Implicit in the model has been the characterization of language as a collection of over-learned behavioral responses: a language stimulus is registered, there is receptive processing (Wernicke's area is activated, an interpretation is computed), then the information is shuttled (along the arcuate fasciculus) to Broca's area - which is responsible for expressive language and hence more or less directly generates an output, for example a sentence. While this may seem like an uncharitable caricature of the scientific position, neural models of language and speech processing often implicitly subscribe to such a view. The research most often does not make any contact with conceptual and theoretical notions that we now take to be crucial for an explanation of linguistic processes in the framework of the computational theory of mind, including the concept of language as a *discrete combinatorial system*, the notion of an *underlying grammar*, *generativity*, *representation*, *derivation*, *feature*, *levels of representation within a linguistic system*, and so on (Chomsky, 1986).[4] Crucially, the conceptualization of language and language processing as no more than an input-output mapping - a behaviorally conditioned response of some sort, and therefore basically an instance of an extraordinarily complex 'reflex' - turns out to be inadequate. In a very methodical analysis of Skinner's Verbal Behavior, Chomsky (1959) demonstrated that a model that views language as an elaborate reflex chain is untenable on logical and empirical grounds in that it makes the wrong predictions for language acquisition and for the adult language system.

---

[4] Jackendoff (1994) and Pinker (1994) provide general introductions to language and cognitive science and discuss many of the foundational issues that provide some of the intellectual framework serving as background for this research.

Turning to functional neuroimaging, have the new methods supported the classical model or refuted it? These results require a more articulated interpretation. PET and SPECT as well as fMRI studies have in fact implicated both Broca's and Wernicke's areas in verbal processing. But two features of the new results are noteworthy. First, these methods have also implicated a number of other areas in addition to Broca's and Wernicke's areas. Functional neuroimaging studies illustrate vividly that language processing is most likely carried out by an array of cortical brain areas in the frontal, temporal, and parietal lobes. The cerebellum and the cingulate gyrus are suspected to play some role in language processing as well. Importantly, current research suggests that the computations involved in speech and language processing are not limited to areas in the left hemisphere; a number of right hemisphere areas are also regularly implicated (Wise et al., 1991; McLaughlin et al., 1992; Zatorre et al., 1992; Binder et al., 1994). The results indicate that in normal language processing the non-dominant hemisphere contributes significantly to the process, at least when assessed by the functional neuroimaging methods.

A second important issue is that although functional neuroimaging studies often activate the classical speech and language areas of Broca and Wernicke, it remains mysterious exactly which cognitive, computational aspects of a given experimental task activate these areas. For example, recent experiments that have generated frontal opercular activation are rather divergent and include a number of complex subcomponents. Cognitive computations that have been attributed to the left frontal operculum on the basis of PET or fMRI measurements include the following: verbal working memory (Awh et al., 1994, 1995); phonological encoding (Bartlett et al., 1987); a "possible role in phonological processing" (Demonet et al., 1992, 1994); a specialization for rapid temporal processing, including speech (Fiez et al., in press); (subvocal) word generation (Hinke et al., 1993); lexical processing (Mazoyer et al., 1993); subvocal rehearsal (Paulesu et al., 1993); phonological processing (Shaywitz et al., 1995); syntactic processing (Stromswold et al., in press); and accessing articulatory representations for speech perception (Zatorre et al., 1992). The task demands that have led to these results included conditions such as passively listening to connected speech, making rhyming judgments on pronounceable non-words, making rhyme judgments on visually presented letters, or making plausibility judgments on sentences.

Intuitively, there are two possibilities to reconcile these divergent reports. One possibility is that there is a single underlying component in all these experimental tasks that selectively engages or activates the left frontal operculum. Although there is no obvious common computational subroutine that suggests itself, a more detailed analytical decomposition of the studies reveals that they all include a substantive verbal working memory component, in the sense that to execute the task demands effectively subjects must (at least once per trial) update their mental representation of the verbal material in question, regardless of whether the task demanded phonetic, lexical, or syntactic judgments. The one candidate computation one might investigate further on the basis of these imaging findings, then, is that the frontal operculum is crucially required for the updating of representations in the verbal domain. Whether this process is an aspect of verbal working memory, such as subvocal rehearsal, is debatable - and indeed debated (see, for example, Waters et al., 1995, who postulate different processing resources for distinct aspects of language tasks, i.e. they draw a distinction between the processing resources required for 'verbal working memory' contrasted with those required for syntactic processing). Further analysis may implicate a more fundamental process - and thus a process that actually underlies subvocal rehearsal. In that sense it would be 'biologically prior' and epistemologically prior to such components of verbal working memory as subvocal rehearsal or phonological encoding (Baddeley, 1992). In absence of extremely detailed task analyses one cannot determine whether brain areas were activated by grammatical processing, sentence interpretation, speech perceptual processing, verbal working memory, or, for example, the attentional modulation imposed by experimental task demands (Poeppel, in press; Poeppel & Johnson, in press).

There is, of course, a straightforward alternative to the notion of a single process (for example verbal working memory) sketched above: it is likely that there is rich internal structure to the frontal operculum, including Broca's area (say Brodmann's areas 44 and 45). Consequently on might hypothesize that this cortical region becomes metabolically engaged by distinct tasks because there are distinct subcomponents tuned to the processing of particular representational attributes. Empirically it has been shown, for example, that one can dissociate 'verbal working memory' from the resources required for syntactic processing (Waters & Caplan, 1995), so there already exists evidence that speaks to the question of whether the frontal operculum is a unitary functional area or not. Note that prior to the method of single-unit recording in animals and the work of Mountcastle (see 1978 for review), Hubel and Wiesel, and many other electrophysiologists the primary projection areas, including somatosensory cortex and primary visual cortex (Brodmann's area 17, or V1), were assumed to be relatively homogeneous sheets of nervous tissue without varied internal structure. The model of visual cortex developed over the last 40 years now includes numerous visual areas each of which has complex subcomponents specialized for particular processing tasks (Hubel, 1988). Analogously, current intuitions about the frontal operculum probably reflect the lack of detailed knowledge of the structure of this cortical region. Should there be a functionally complex organization to these regions (a likely scenario), the challenge is how to study that organization effectively in the human given that single-unit recording is not a viable method.

Returning to the issue of the status of the standard model, the fact that the canonical view is being revised (implicitly, if not always explicitly) is not particularly surprising. The view of language that it assumed is very coarse and disconnected from research on language and cognitive psychology. A more accurate model designed to capture the relationship between the brain and speech and language processing will have at least the following features: the refined model will be grounded in a richly articulated theory of language; it will be closely tied to the results of cognitive psychology; and it may depart from the simplistic notion that one brain area is responsible for one language function - the idea that phonology lives in a particular brain area, grammar lives in a particular brain area, and so on.

The notion that a *single* brain region constitutes the neural substrate for a subroutine of language must be critically examined. The single area-single function model contrasts with the idea that what underlies the processing of particular language representations are networks of multiple areas that are heavily interconnected and that may be located in both hemispheres. Such a conceptualization has nothing to do with the currently popular idea of *distributed representations* discussed in the cognitive sciences (Rumelhart & McClelland, 1986). Rather, language processing is probably supported by a number of cortical and subcortical areas simultaneously, and that what is characteristic of a given linguistic process is the spatio-temporal pattern of activation across the array of areas. The individual areas may preferentially process certain *attributes of linguistic representations* (akin to visual areas that preferentially process perceptual attributes of visual stimuli, such as a visual representation's color, texture, or motion). But to generate a coherent linguistic representation the integrity of an array of areas is required. One might think of the array of interconnected areas as 'macroscopic circuits' made up of several brain regions that interact in special ways.

This being said, an important research goal is to develop a more articulated functional neuroanatomy by identifying the cerebral areas that make up a circuit of this sort. Methodologically, the most plausible new approach to achieve this is the possibility of combining imaging modalities (for a recent example of this approach, see Heintze et al., 1994). One promising strategy is to combine MRI, fMRI, and MEG. MRI provides the detailed structural information about the brain, fMRI yields the macroscopic functional neuroanatomy and thus identifies regions of interest, and MEG can be used effectively to generate high-temporal resolution information about a given region of interest, and perhaps even to generate functional topographic maps of areas (cf. Chapter 4). The ability to study one area in great detail and with many complementary methods will allow for a more hypothesis-driven approach. For example, once a brain region has been reliably identified as relevant to some psycholinguistic subroutine (say with fMRI), the methodical 'unpacking' of an area can begin, for example with MEG (or any other method of the neurosciences, for that matter). The detailed study of one area and its role in specific linguistic computations will eventually lead us to clearer hypotheses about the sensitivities and representational properties of cells in 'higher, cognitive' brain areas.

One fundamental scientific challenge is to clarify basic biological issues, such as the question of how categorial information in general and linguistic categories in particular are represented and processed and how is information of any kind stored and retrieved by the brain (Gallistel, 1990). It should be possible to identify physiological indices that correlate systematically with linguistic categories and their processing. A sensible approach is to probe first for physiological responses/ indices that correspond to representations which have independent theoretical motivation and independent experimental support. On the other hand, there is no reason to believe that any of the units we are aware of now are relevant from the 'brain's-eye-view.' Although it would be surprising if the brain were computing language using completely new units of linguistic representation, it is possible that one should be looking for 'non-canonical' units of representation. For example, it is conceivable that notions like *phoneme* or *word* are not physiologically identifiable units as much as other theoretically derived entities like 'feature' or 'bundle of features' or 'feature combination.' (In fact, the issue of what constitutes 'core' features of language is discussed frequently in the theoretical literature. In phonology, for example, a number of theories do not refer to introspectively obvious entities such as words but to less accessible properties of linguistic items, such as voice-onset time, aspiration, metrical structure, and so on. For example, in contemporary linguistics the concept *phoneme* is not a viable theoretical entity.) If, by properly applying the functional neuroimaging methods, we are able to identify some computationally and cognitively interpretable unit of representation that correlates systematically with some feature of physiology that we can measure, we will have come a long way towards developing a biologically grounded theory of the linguistic cognitive system.

## A COGNITIVE NEUROSCIENCE APPROACH TO SPEECH PERCEPTION

The central goal underlying the research presented in this dissertation is to clarify where speech sounds are processed in the human brain. At stake are hypotheses about the neural basis of speech that are investigated using various sources of evidence, not limited to anatomical and physiological evidence, and including, crucially, evidence from psycholinguistics.

There is a conceptual and factual distinction made in the language sciences between speech and language. Speech is one possible 'articulatory-perceptual interface' (Chomsky, 1992), another interface being the visuo-motor modality used in sign-languages. These modalities interface with a central computational system that has been called a grammar by Chomsky (Chomsky, 1986). What is labeled as 'language' in numerous studies conflates the distinction between grammar-dependent and interface-independent computational processes on the one hand, and the sensory-motor interface processes on the other. This conflation of distinct representations (and their associated processes) leads to biological modeling that is too vague with respect to independently established linguistic concepts and consequently is in danger of also underspecifying the neurobiological analysis.

The most common account of speech perception from the perspective of neurology and the neurosciences maintains that the left temporal lobe contains regions dedicated to the extraction of those acoustic features required for speech perception (see Adams and Victor, 1989, Kandel, Schwartz, and Jessell, 1991, or Kolb and Wishaw, 1985, who summarize the standard views in neurology, neuroscience, and neuropsychology, respectively). This position derives on the one hand from an analysis of patients with lesions to the left perisylvian cortex who, subsequent to injury or disease, exhibit deficits that are sometimes more specific to speech perception, and on the other hand from experimental studies using the dichotic listening paradigm (Studdert-Kennedy & Shankweiler, 1970; Hugdahl, 1992). The position is, however, not tenable in its most straightforward form. The strongest lateralization account (speech is processed in the left hemisphere) holds only insofar as the underlying model of language processing does not distinguish between the sensory-motor interface representations (speech or sign) and the input-independent (central, grammatical) linguistic representations.

Previous accounts of speech perception and brain organization most often are based on data correlating the sites of brain damage with specific functional deficits. The deficit-lesion correlation approach has generated the state-of-the-art map of where speech and language are processed in the brain (see Caplan, 1994 , for a number of references). A second category of data that have been used to argue for the left hemisphere specialization of speech perception are the dichotic listening results. These studies have established that when two distinct speech stimuli are presented simultaneously to the two ears, there is a consistent right ear (left hemisphere) advantage for stop-consonant identification as reflected by identification latencies and error rates (Studdert-Kennedy and Shankweiler, 1970).[5] Other kinds of evidence speak to the question as well. Both neuroimaging studies as well as (invasive) intrasurgical recording studies provide important data that contribute to our understanding of the neural basis of speech perception. It turns out, however, that there is substantial divergence among the results, in particular regarding the localization of phonetic and phonological processing. The cumulative recent evidence from neuropsychology (Caplan, Makris, and Gow, 1995), positron emission tomography (Poeppel, in press), functional magnetic resonance imaging (Binder and Rao, 1994), single unit electrode recording (Ojemann, 1991), chronic grid electrode recording (Boatman et al., in press), and magnetoencephalography suggests that our current understanding of the neural basis of speech perception has to be revised.

In view of the distinction between an underlying grammatical system (the language faculty) and the performance systems it interacts with (input/output systems, memory, attentional mechanisms), an hypothesis emerges that refines the model that speech is typically (in the majority of right-handers) processed in the left hemisphere. Specifically, I explore the model that a substantive proportion of the auditory system's computations that mediate speech perception are represented bilaterally, in the left and right superior temporal gyri, posterior to primary auditory cortex. I submit that the computations from acoustic signal to speech percept (registration, transformation, and integration) are *cortical* and *bilateral*.

---

[5] The speech processing model espoused in this thesis needs to be consistent with the dichotic listening results. An interpretation arguing that these results compel the conclusion that speech perception is a left-hemispheric specialization is too strong. The dichotic listening results are behavioral results that, alone, do not say much about the neurobiological organization of the domain in question. The results provide interesting evidence that speak to the connectivity and transmission properties of the system. But results like the dichotic listening facts argue neither for not against the underlying right hemispheric contribution to the task of computing phonetic-phonological representations. A *neurobiological* analysis must account for the fact that independent evidence (e.g. imaging, lesions) suggests that the right hemisphere contributes to the computation.

This dissertation evaluates these hypotheses about the functional-neuroanatomic account of *speech perception as a subroutine of language comprehension* or the basis of three sources of evidence: (1) neuropsychological evidence from a speech-specific auditory agnosia known as pure word deafness, (2) metabolic brain imaging evidence, and (3) new experimental electrophysiological evidence from magnetoencephalographic recordings.

## Pure Word Deafness

Pure word deafness (PWD) is a relatively rare form of auditory agnosia in which the integrity of the speech comprehension system is compromised, typically by an injury to some cortical area and/or the underlying white matter. As a consequence of the lesion, the comprehension of spoken language is dramatically impaired. In contrast, the abilities to speak, read and write, and to process non-linguistic auditory information remain relatively intact. The patients can occasionally make use of extralinguistic information (possibly also some suprasegmental, prosodic information) to infer properties like the gender of the speaker, emotional state of the speaker, and so on. However, segmental and syllabic information are not available in a way that allow the patients to effectively decode spoken language, for example in accessing the auditory input (mental) lexicon. PWD is a compelling clinical syndrome that demonstrates a dissociation of non-linguistic auditory processing, speech perception, and language and that consequently speaks to the functional architecture of auditory perception and cognition. Chapter 2 provides a detailed review and reappraisal of the PWD literature. The analysis of the clinical findings in PWD shows that PWD is either a consequence of bilateral lesions around posterior superior temporal gyrus or of a left unilateral deep temporal lesion that disconnects the left and right posterior superior temporal gyri. The PWD results are consistent with the claim that speech information is processed bilaterally in the superior temporal gyri.

## Neuroimaging speech perception: metabolic imaging results

The functional imaging literature on auditory speech and language comprehension can be characterized by two kinds of experiments. One class of studies, passive presentation experiments, simply play the speech stimuli (ranging from vowels to syllables to words to connected speech) to subjects and record the neural activity. In these studies no task is required of the subjects. In contrast, some studies require subjects to attend to the auditory speech stimulation in a way that allows them to execute specific task demands. The demands vary from monitoring for a phoneme or phoneme sequence to making plausibility judgments on sentences. Chapter 3 summarizes the neuroimaging literature, organized around the distinction between passive and task-conditioned studies. The classification of the neuroimaging findings into two classes of results on the basis of the presence or absence of a task is, of course, relatively arbitrary. However, the distinction turns out to have consequences for the imaging results. In particular, in the absence of (meta-linguistic) task demands investigators typically detect significant activation of the posterior superior temporal gyrus in both hemispheres. In contrast, when subjects are required to perform a specifically linguistic, attentionally demanding task, the results are often different: there is some suggestion that in those cases the activity is more lateralized, with the left showing larger response amplitudes. In contrast, the presence of a task never affects the results such that the right temporal cortex appears to be engaged more than the left. I argue that the presence of a linguistic task can change the imaging results such that it appears as if the left hemisphere were considerably more actively engaged in decoding the speech input. I hypothesize that it is the presence of the task itself that is predominantly generating this result, obscuring that right temporal cortex is contributing to speech perception. An important methodological point to be explored on the basis of the meta-analysis of this literature is that when one uses attentionally demanding tasks one might be selectively biasing the results in a particular direction. In the presence of tasks one is in danger of actually studying the neural basis of task demands, not the neural basis of the cognitive domain under consideration.

### Magnetic source imaging experiments on acoustic phonetics

Chapter 4 presents new experimental work. The experiments are concerned with how speech sounds are represented and processed in supratemporal auditory cortex. The studies use magnetic source imaging (MSI). MSI combines the high spatial resolution of magnetic resonance imaging (MRI) with the high temporal resolution of magnetoencephalography (MEG). MEG uses extremely sensitive SQUID-based sensors to sample the magnetic field at the scalp. It is possible to reconstruct the spatial position in time of the current that generates the extracranial field. Motivated by the meta-analytic results from pure word deafness and metabolic neuroimaging, four experiments were done: a passive presentation of tones, a passive presentation of vowels, a passive presentation of CV syllables, and a discrimination study with the same CV syllables. From a methodological point of view the results show that MSI can be an effective tool for investigating the neural basis of speech perception. The experimental results suggest that auditory cortex goes considerably beyond spectral analysis: rather than passively registering the frequency bands out of which speech stimuli are constructed, auditory association cortex plays an extensive role in the construction, representation, and processing of speech sounds. In addition, the results show that the presence of task demands differentially affects activity in left and right auditory cortices: the amplitude of the evoked auditory response in auditory association cortex is *increased in the left* and *decreased in the right.*

### Lateralization predictions

The first hypothesis investigated in the MEG studies was that there exist no qualitative hemispheric differences (no lateralization) in the early physiological responses to speech stimuli. Claims to the contrary notwithstanding, the results on pure word deafness (Chapter 2) and the results of a number of functional imaging studies (Chapter 3) suggest that both superior temporal lobes participate in the analysis of speech signals. To test the hypothesis we ran a battery of experiments and analyzed the results with respect to the variable *hemisphere*. Participants were presented with pure tones and speech sounds (vowels and CV syllables). The prediction was that there would be no hemispheric effect for any of these stimuli when looking at the evoked response component of the neuromagnetic field and the localizations of the responses. Insofar as systematic and interpretable differences are detected - either in localization or in response amplitude - it will be important to address whether the difference could have consequences for perception. For example, a difference in location of activity may be a difference that is merely an accidental consequence of the morphology of a participant's planum temporale. Alternatively, it may reflect neural activity in distinct processing areas.

The second hypothesis was that attentional demands selectively change the activation in left supratemporal auditory cortex. The simplest model that the data suggest is that speech perception is mediated bilaterally and that the attentional modulation of speech perception leads to a 'gating' towards the left hemisphere. Although the left hemisphere appears to preferentially process speech and language information, we may be seriously underestimating the contribution of the right temporal cortex to speech perceptual processing.

To recapitulate, the core assumption motivating the set of experiments is that speech and central grammatical representations are processes that must be methodically distinguished with regard to their neural basis. Specifically, I hypothesize that speech perception proper is mediated by one or more cortical areas in both the left and right posterior superior temporal gyri and that attentional mechanisms interact differentially with the two hemispheres. The chapters proceed from a more macroscopic to an increasingly microscopic perspective. In Chapters 2 and 3 I summarize and analyze the pure word deafness literature and functional imaging literature, respectively, and derive the two hypotheses that are tested in the experiments presented in Chapter 4.

# CHAPTER TWO

## PURE WORD DEAFNESS

### 2.1 Pure word deafness: definitional preliminaries

One important source of evidence for the investigation of the neural basis of speech perception is the deficit-lesion correlation literature. There are a number of cerebral lesions that affect speech perception. For example, speech comprehension deficits are common in Wernicke's aphasia. Among the cerebral lesions that cause major speech perceptual disturbances, a syndrome called *pure word deafness* stands out because it is a remarkably specific deficit. Pure word deafness (PWD) is a rare form of auditory dysfunction in which the integrity of the speech comprehension system is selectively compromised. Subsequent to a cortical or subcortical lesion, typically in the region of the posterior superior temporal lobe, the comprehension of spoken language is dramatically impaired. In contrast, the abilities to speak, read and write, and (in the 'purest' form of word deafness) the ability to process (hear, recognize, identify, and label) non-linguistic auditory information, including music, remain relatively more intact. Whether there exist any truly 'pure' cases is debated, because most often these patients are also impaired along other dimensions of auditory cognition. Nevertheless, what makes word deafness a particularly compelling clinical syndrome is that its existence demonstrates a dissociation of general, non-linguistic auditory processing, speech perception, and central language processes. Pure word deafness therefore speaks to the functional architecture of auditory speech perception and cognition, and the careful correlation between lesion site(s) and clinical syndrome in PWD can contribute valuable data to advance our understanding of the neural basis of speech perception and auditory cognition.

There have been two major reviews of word deafness (auditory verbal agnosia) in the literature (Goldstein, 1974, and Buchman et al., 1986), and a major case report and reassessment more recently (Praamstra et al., 1991). Because a number of new studies have appeared, and because current thinking about the organization of speech recognition and language processing has evolved considerably in the last 10 years, a thorough review, update, and reevaluation of pure word deafness seems appropriate. This chapter is a modest attempt in that direction.

In the sense that we think about it now, PWD was first explicitly recognized and described by Kussmaul (1877), who also coined the term ('reine Worttaubheit'). While there were some academic precedence disagreements between Bastian (1897) and Wernicke (1874) about who recognized the disorder, Kussmaul was the first to carefully separate pure word deafness from alexia and other aphasic syndromes. (Together, PWD and alexia constitute Wernicke's sensory aphasia.) Other early contributors to this area of research include Lichtheim, Freud, and many other well-known neurologists (see Table 2.2).

Because selective auditory disorders are relatively rare, somewhat confusing, and almost always overlapping with the presence of other syndromes, some terminological clarification is in order. In addition to pure word deafness, there is a deficit-lesion syndrome known as *auditory (sound) agnosia* (one subtype being *amusia*), and a separate, more general category called *cortical deafness*. Table 2.1 lists the major criteria for inclusion in a particular class. Note that what sets cortical (central) deafness apart from PWD and auditory agnosia is the impaired hearing sensitivity as shown by audiometric testing.

**Table 2.1** *Auditory disorders following cortical and/or subcortical lesions*

| | *Pure word deafness* | *Auditory agnosia* | *Cortical deafness* |
|---|---|---|---|
| Speech comprehension | impaired | impaired or (rarely) OK | impaired |
| Speech repetition | impaired | impaired or (rarely) OK | impaired |
| Non-speech auditory comprehension | OK | impaired | impaired |
| Hearing sensitivity (audiometry) | OK | OK | impaired |
| Spontaneous speech | OK | OK | OK |
| Reading comprehension | OK | OK | OK |
| Writing | OK | OK | OK |

Adapted from Buchman et al., 1986.

Even more rare than PWD is the case of auditory agnosia exclusively for non-verbal sounds, i.e. the selective sparing of speech perception despite profound impairments in the perceptual processing of non-speech auditory sounds. Nevertheless, such cases have been reported, for example by Nielsen & Sult (1939), Wortis & Pfeffer (1948), Spreen et al. (1965), Albert et al. (1972), Motomura et al. (1986), Fujii et al. (1990), and Peretz et al. (1994); for discussion, see also Vignolo (1982). The existence of such auditory disorders is important for comparison with and control for PWD in that such deficits show that speech and non-speech sounds can 'doubly dissociate' in the standard neuropsychological sense. Ideally one can identify distinct lesion sites that correspond to the two types of auditory agnosias and thereby identify areas implicated in speech or nonspeech auditory processing. More frequently, patients present with 'generalized auditory agnosia,' i.e. the combination of word deafness and auditory agnosia in spite of unimpaired hearing sensitivity (Buchman et al., 1986).

This chapter reevaluates the pure word deafness literature and addresses how PWD and the associated lesions speak to the neural basis of speech sound processing. In brief, the evidence both from cortical and subcortical lesions is most consistent with the view that speech perceptual processes are mediated bilaterally in superior temporal cortices, not just in the so-called 'dominant' hemisphere. In addition, I suggest (following Auerbach et al., 1982) that there are PWD subtypes. The subtypes, dependent on lesion site (subcortical or cortical), will only be appreciated as a consequence of careful and subtle psycholinguistic testing.

## 2.2 The subjective phenomenological experience of pure word deafness

The descriptions of the subjective experience of PWD are intriguing, and there is remarkable consistency among the reports. For example, both the patient studied by Klein & Harper (1956) and the patient studied by Ziegler (1952) express similar experiences: "[Spoken language] is like a great noise all the time ... you think you can catch it and it fades away .. like foreign folks speaking in the distance. ... When people speak loudly or quickly the words just run together." "[Language] sounds like a wind in the trees, a murmuring, a foreign language" (cited in Albert & Bear, 1975). Similarly, the patient examined by Jerger et al. (1972) and by Kanshepolsky et al. (1973) said to his testers "I can hear you talking, but I can't translate it" (Jerger et al., p. 524). One patient, studied by Mendez and Rosenberg (1991), "experience[d] ... spoken words as meaningless noise, garbled sound, or a foreign language" (Mendez and Rosenberg, p.210). Another patient could "discriminate a human voice from other sounds, but he stated that it sounded like a foreign language, except for his own speech which sounded normal. ... He was however capable of identifying the number of syllables in a word and the number of words in a sentence." (Takahashi et al., 1992, p. 296).

The patient studied by Saffran et al. (1976) pointed out to the experimenters that "it's as if there were a bypass somewhere, and my ears were not connected to my voice" (p.211). These investigators report that he was "unable to describe his auditory experience beyond the frequent complaint that speech sounds did not 'register,' or that they would not 'come up'" (Saffran et al., p. 211). The patient tested by Auerbach et al. (1982) reported that both speech and music sounded like noise, although he was able to recognize other sounds. "In describing his symptoms, the patient stated that the speech of others often sounded like a 'foreign language.' Music was incomprehensible and he was unable to pick out familiar music or recognize instruments as he formerly did. Noise was bothersome and background music in a restaurant required earplugs. ... [H]e responded that people's voices seemed faster, that it helped him when people spoke slower, and that words came too quickly ..., and would run together. He did not acknowledge changes in the quality of other people's voices, but he felt that his own voice sounded slightly different. He could still recognize an individual by voice even in a crowd, although he was unable to determine what language was being used even when it was a familiar language such as French or English" (Auerbach et al., p. 273/4).

Evidently the experience for a number of patients is that speech sounds to them like a foreign language. This phenomenological feature is particularly interesting in that it suggests that they still hear speech as speech, not as noise. In future testing it would be helpful to distinguish between patients who perceive speech as speech and those who perceive speech as something fundamentally distinct from speech. One prediction one might derive from the analysis presented below is that patients in whom cortex is spared might have residual projections to posterior STG that mediate some degree of speech perception. Only total destruction of bilateral STG would cause a complete word deafness.

It is commonly reported that patients' expressive language is unaltered, fluent, unhalting, and so on. However, only rarely is production formally tested - or even informally assessed beyond mentioning that speech is fluent. There is, however, good reason to test production more systematically. The issue at stake is what role auditory feedback plays in speech production in these patients. Intuitively, PWD patients' output must sound at least slightly altered since they presumably receive no speech feedback. It would be surprising if these patients did not sound even slightly different and would suggest that the production/articulation information that one has proprioceptively is extremely rich - so rich as to override any perceptual deficits. What needs to be clarified is the extent of a closed-loop feedback system in production. That the effect is non-nil is demonstrated by speaking under delayed auditory feedback conditions, which, at least temporarily, induces anomalous production. A further possibility is that the dissociation between the perception of one's own voice and the voice of other speakers becomes particularly acute given the type of lesion that causes pure word deafness. In general, there is insufficient information on this question. Only few investigators mention any overt speech problems. Auerbach et al. (1982), for example, report that their patient slighly slurred his speech and stuttered slightly as well; however, his connected speech was otherwise fluent, gramatically correct, and reflected normal phrase length.

### 2.3 Word deafness: the data

Table 2.2 lists reports on pure word deafness, ranging from the earliest analyses (1874) to recent case studies. This table and the references below provide a relatively exhaustive database of references on pure word deafness. A number of reports involving word deafness were not included in the analysis, either because they did not provide sufficient lesion information, because they only tangentially touched on word deafness, or because they were not available. These include LeGros Clark et al., 1938; Hemphill & Stengel, 1940; Reinhold, 1950; Jones & Dinolt, 1952; Kleist, 1962; Okada et al., 1963; Tanaka, 1964; Albert et al., 1972; Horenstein & Bealmear, 1973; Nagafuchi et al., 1973; Albert & Bear 1975; Denes & Semenza, 1975; Chocolle et al., 1975; Gordon & Naeser, 1977; Shoumaker et al., 1977; Oppenheimer & Newcombe, 1978; Kirshner & Webb, 1981; Doyle & Holland, 1982; Ikemura, 1983; Rapin, 1985; Notoya et al., 1991; Seliger et al. 1991; DiGiovanni et al., 1992; Nishioka et al., 1993.

TABLE 2.2 HERE

## 2.4 Lesion site analysis

The 60 tabulated reports provide information about 62 patients. Of the 62 patients, sufficient data are available for 59 of them to assess whether the lesion was bilateral or not. The data distribute in the following way: 42 cases have bilateral temporal lobe lesions (cortical and subcortical), 17 cases have unilateral lesions (cortical and subcortical). Of the 17 unilateral lesion cases, 16 are in the left temporal lobe and one is in the right temporal lobe (Roberts et al., 1987). To evaluate whether this patient might be right dominant, the patient with the right hemisphere lesion was tested for handedness and was strongly right-handed. Moreover, Roberts et al. found no evidence of any left hemisphere abnormality, so this case remains the single outlier in the otherwise clear pattern. That pattern is that lesions leading to PWD are (1) bilateral in the superior temporal cortex or bilateral affecting primary auditory cortical fibers projecting to posterior superior temporal gyrus or (2) left unilateral and compromise the auditory radiations, or the transcallosal projections from the right temporal cortex to left superior temporal cortex. As pointed out by Takahashi et al. (1992), in cases of PWD consequent to unilateral lesions, the lesion site includes left posterior temporal subcortical areas and, crucially, the lesions interrupt the auditory projections in cortex as well as transcallosal projections from the other hemisphere, thereby essentially deafferenting posterior STG from inputs from both hemispheres. Thus, in the unilateral cases the deep temporal lesion appears to affect both STG insofar as the transcallosal fibers from the right are also compromised.

In five of 59 cases the authors argue for exclusively subcortical lesions (Henneberg, 1926; Hamanaka, 1980; Brick et al., 1985; Kazui et al., 1990; and Takahashi et al., 1992). Of these five subcortical-only cases, 2 had bilateral lesions and three had unilateral lesions affecting left STG. In no clinical case was the lesion explicitly argued to be exclusively cortical, because the researchers leave the possibility of a subcortical lesion open.

In 13 of 62 cases the authors maintained that the syndrome was 'pure,' i.e. the patients had completely preserved non-speech auditory perception and cognition (see last column in Table 2.2, where these cases are marked with a minus-sign). 2 of these were among the three cases in which one couldn't ascertain whether the lesion was bilateral or unilateral. The remaining 11 split up into 6 cases with bilateral lesions and 5 cases with unilateral lesions, so one can derive no obvious generalization between lesion site and extent and 'pureness' of the disorder.

Many of the *bilateral* cases share one crucial feature: after the first (unilateral) lesion, auditory perception, including speech perception, was intact after an initial period of disturbance. Only after the second lesion - on the other side - did permanent cortical hearing problems, in particular with respect to speech stimuli, remain (cf. animal studies discussed below, in which similar findings were obtained after unilateral and bilateral superior temporal cortex ablations). In virtually all cases with bilateral lesions there were two independent injuries affecting approximately the same area in both hemispheres. The fact that most of these patients only manifested PWD subsequent to the second lesion rather supports the hypothesis that both sides of STG mediate speech perception. Ulrich (1978) analyzed the temporal order of left and right lesions for a number of cases to ascertain whether there was any systematic pattern. He found no obvious pattern, including the possibility that PWD only becomes manifest after a second lesion on the left - in which case one might argue that it is really only left pSTG which is implicated. Rather, both left-before-right lesions and right-before-left lesions were frequently attested.

Central to one of the two core hypotheses addressed in this thesis is the finding that in the majority of cases pure word deafness is caused by *bilateral lesions of the superior temporal cortex*, including the posterior superior temporal gyri. While this is important evidence in its own right, taken together with the neuroimaging evidence (Chapter 3) these data support the hypothesis that speech sounds are neurally mediated bilaterally and probably in the posterior superior temporal gyri.

Nevertheless, almost one third of PWD cases have unilateral lesions in left temporal cortex. I propose that the left unilateral cases fall into two groups. Characteristic of one group is that the lesion is deep subcortical and affects *both* left and right STG by compromising the left ipsilateral projections to pSTG as well as the transcallosal projections from and to left and right STG (cf. Takahashi et al., 1992). Although the lesion is left unilateral, it implicates left and right STG. This pattern may be an instance of associative agnosia (failure to connect speech sounds to other relevant representations). I submit that the second group of left unilateral PWD cases is associated with lesions that include extensive portions of cortex. In those cases aspects of language comprehension other than speech perception proper are also compromised, including effectively accessing the auditory input lexicon. In any case, I maintain that the unilateral cases are consistent with the hypothesis that both STG mediate speech pereptual processes.

## 2.5 Phonetic discrimination in pure word deafness

Having assessed the lesions presumed to underlie pure word deafness, this section turns to those reports that provide more detailed information on the psycholinguistic outcomes. There are three leading questions one might bear in mind when thinking about PWD. First one needs to establish what, precisely, is the nature of the deficit from a psycholinguistic perspective (i.e. what is the status of phoneme discrimination, lexical access, and so on). Second, one needs the data on how PWD patients perform with respect to 'para-linguistic' features of sounds such as their rhythmic structure, pitch contour, whether or not their sensitivity to suprasegmental or nonlinguistic features is altered. Also, is`the deficit merely an expression of a difficulty in making within-category judgments for complex stimuli or is it truly a deficit of speech perception? Finally, it would be interesting to be able to ascertain what the status of internal speech and auditory imagery (speech imagery) is in patients with PWD. Although these are important questions, in no case have they been answered in a satisfactory manner (or at all); this level of detailed investigation will have to remain a future project. Nevertheless, there are a number of papers that test patients' speech perception abilities with some rigor. In particular, six more recent studies did some level of psycholinguistic evaluation. The findings are briefly summarized below.

Saffran et al. (1976) ran a number of psycholinguistic tests on one patient, incuding a dichotic listening protocol, a test of contextual influence on speech recognition, and a test of the effect of category cues on repetition (i.e. does repetition improve if the subject knows the topic of discourse). Importantly, they also tested the patient's phoneme discrimination and categorical phonetic boundaries using consonant-vowel syllables differing in place of articulation and voice onset time (VOT). Based on a series of forced-choice task paradigms they found that the patient classified and discriminated these stimuli abnormally. (1) Consonant perception was dependent on the following vowel (performance between 20% and 80% correct), unlike for unimpaired listeners. The investigators interpreted this as reflecting increased sensitivity to acoustic features. (2) Unvoiced stop consonants (/p t k/) were identified with better accuracy than voiced stop consonants (/b d g/). (3) The subject's VOT boundary distinguishing voiced and unvoiced consonants was apparently normal. (4) Overall, errors in voicing were more pervasive than errors in place of articulation. They interpreted their findings as showing a *pre-phonetic* impairment which was considered to be a consequence of impaired temporal resolution in the auditory domain.

Auerbach et al. (1982) examined one patient and found that (1) his vowel identification performance was 100%. (2) In a test of consonant-vowels (CV) syllable identification, his performance declined with the number of alternatives presented. In a two-alternative forced-choice test, the patient was correct 74% of the time when one feature differed between the stimuli (i.e. voicing or place) and 93% correct when both features differed. In a three-way discrimination task, performance dropped to 69% (voiced stops) and 53% (unvoiced stops). When asked to select the correct CV from 10 alternatives, the patient made numerous errors on place and voicing. (3) The VOT category boundaries were judged to be normal. These authors also concluded that the deficit was a pre-phonemic auditory disorder (possibly in auditory temporal acuity).

Miceli (1982) tested a woman with bilateral temporal lobe lesions and significant word deafness and sound agnosia. (1) Vowel identification was 100%. (2) In a CCVC identification test more than 95% of the total errors were due to errors of place. Voicing errors contributed only 3.8% of all errors. (2) In a CV identification test, the total percentage of errors was somewhat lower (29.2%), although the pattern remained the same: a small number or errors for voicing or voicing+place and a large number of errors for place distinctions alone. (3) In a VC identification task, the pattern replicated the one found for CV stimuli. (4) The patient's category boundary along a VOT continuum was normal. In further tests probing for place of articulation identification and discrimination using synthesized stimuli, performance was dramatically attenuated compared to normals. This study, then, demonstrated a sharp dissociation between voicing and place, with place of articulation being considerably more affected than voicing perception. Miceli attributed the deficit to a 'psychoacoustical deficit' that lowers the efficacy of spectral analysis. Spectral analysis is, according to this argument, crucial for stop consonant perception and, apparently, less crucial for vowel perception. One must wonder, on that interpretation, how vowels could be identified at all given a compromised spectral analysis system, since the defining feature of vowels are formants at different bands of the frequency spectrum.

Tanaka et al. (1987) performed some psycholinguistic testing, notably vowel identifications and CV discriminations. (1) In the vowel test, the patient was verbally presented with five 1-second vowels (/a, e, i, o, u/). Of a total of 150 presentations of all vowels, the patient was correct an average of 47%. (Lip reading was not allowed.) (2) In the CV condition, the patient had to perform binary decisions: she was presented with a CV and had to point to whether it was one of two visually shown syllables. The patient was tested on voicing and place of articulation distinctions (/b d g/ versus /p t k/). The test was repeated for one versus two distinctive features (i.e. /b d/ versus /b t/). The patient's performance ranged from 44-54% correct (chance). The performance for place and voicing distinctions was equal (around 50%, which was chance peformance). Thus, although according to the authors this patient was able to identify vowels above chance, vowel identification was significantly impaired. CV discrimination performance indicated a profound deficit in stop consonant decoding.

Yaqub et al. (1988) were particularly thorough in evaluating an Arabic speaking patient presenting with PWD. A passage from their report illustrates the disorder very clearly:

> "Spoken speech was fluent, grammatically correct, with normal phrase length and no paraphasic errors. He listened to a tape with Arabic songs. He could recognize and appreciate music and musical instruments, but not the words of the song, although he guessed correctly that they were in the Arabic language. When he listened to a radio broadcasting a chapter from the Holy Koran, he recognized that it was from the Koran, but did not know the chapter, and could not understand the words. He was blindfolded, and a series of oral tests showed the following correct scores: 4/10 for yes-no questions; 4/10 for oral commands; 4/10 for repetition of single words, 2/10 for sentences, 6/10 for verbs, 0/10 for prepositions and 7/10 for the comprehension of proverbs. He was able to recognize all nonverbal sounds such as key jingling, telephone ringing, rustling paper, knocking on the table, hand clapping and animal sounds.
> When he did not understand a specific verbal stimulus he could not tell whether it was a real word or a nonsense syllable complex. He was able to discriminate between a male and a female voice and between Arabic and non-Arabic languages. He was able to tell when the same sentence was delivered as a question, imperative, or exclamation. He could also recognize when the speaker's voice was emotionally neutral, happy, or angry. He could discriminate between laughter and crying" (Yaqub et al., p. 458).

In a test of (Arabic) vowel identification, this patient scored 97%. In two-way and three-way tests of phoneme discrimination (CV syllables differing in voicing, place of articulation, or both) the patient's performance ranged from 56% to 95%. He was always worse in the three-way classification, as expected on probabilistic grounds. In a ten-way classification of consonants and CV syllables (100 trials), the patient made 10 voicing errors and 40 place-of-articulation errors (cf. Auerbach et al., 1982).

Praamstra et al. (1991) performed a battery of psycholinguistic tests too large to summarize here. In brief, the findings that compare to the other studies showed that voicing discrimination was worse than place discrimination; the category boundaries defined by VOT were abnormal; and vowel identification ranged from 43% to 92%, depending on the task. Their findings are summarized in Table 2.3, together with the other 5 studies.

Notice that the results show no consistent pattern; unsurprisingly, therefore, there is no consensus about the phonetic disturbance in PWD. While it is disappointing that one cannot detect systematic findings across studies, it is impressive that these patients manifest such serious speech perceptual disorders despite relatively full competence in speaking, reading, and writing. Note, also, that there is some agreement across cases: while phoneme identification and discrimination appear to be severely impaired, phonological information such as metrical structure appears to cause considerably fewer problems for PWD patients. In almost all cases the patients are able to satisfactorily execute such tasks as detecting the number of syllables in a word, and manipulate other critical suprasegmental features.

**Table 2.3** *PWD patients' performance on phonetic identifications*

| | Consonant ID | | Vowel ID | VOT boundaries |
|---|---|---|---|---|
| | voicing | place | | |
| Saffran et al. | <<< | < | OK | normal |
| Auerbach et al. | < | <<< | 100% | normal |
| Miceli | < | <<<<< | 100% | normal |
| Tanaka et al. | <<< | <<< | 47% | - |
| Yaqub et al. | < | <<<<< | 97% | - |
| Praamstra et al. | <<<<< | <<< | 43-97% | abnormal |

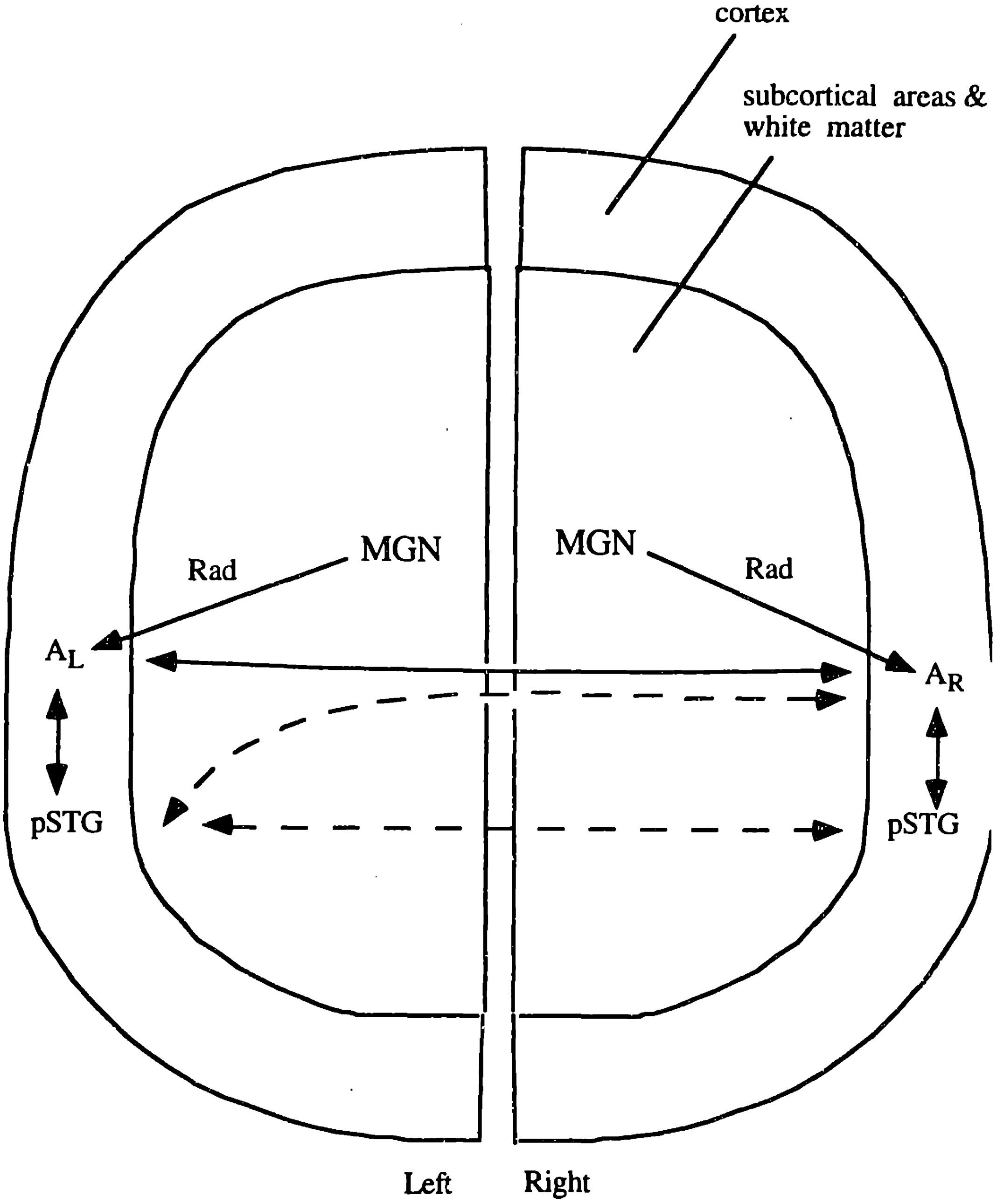
Key: < minor impairment; <<< medium imp.; <<<<< severe imp. ID - identification.

One point to bear in mind with respect to the interpretation of these findings is what the concept of 'identification,' used in virtually all the tests, can mean in the assessment of speech perception in word deafness. Because executing an identification draws on a lot of mental and neural resources, the task may already be too complex to assess early speech processing. If speech perception is really impaired at the level of constructing early representations (apperceptive agnosia), even simpler tests of recognition (e.g. same/different judgments) may be more appropriate in evaluating the degree to which patients can process speech stimuli.

## 2.6 Patterns of lesions causing word deafness

Two basic classes of lesion patterns have been proposed to best capture the phenomena of word deafness, cortical deafness, and non-speech auditory agnosia and account for the manifestations of the syndrome of word deafness and associated cortical auditory disorders: bilateral cortical lesions and unilateral subcortical lesions. This section summarizes the lesion facts. Figure 2.1 schematizes the relevant connectivity of the higher auditory system, and Table 2.4 summarizes the lesion patterns leading to PWD with reference to the areas and projections graphically shown in Figure 2.1.

**Figure 2.1 Schematic of higher auditory system connections**



MGN - medial geniculate nucleus
Rad - auditory radiations
$A_{L/R}$ - left/right primary auditory cortex
pSTG - posterior superior temporal gyrus
arrows - hypothesized (dashed lines) and established projections

**Table 2.4** *Possible lesion patterns causing pure word deafness and cortical deafness* *(Adapted and extended from Earnest et al., 1977, and Takahashi et al., 1992)*

*Abbreviations refer to Figure 2.1.*

| | | | |
|---|---|---|---|
| **A. Cortical lesions** | | | |
| bilateral | 1. | A + pSTG | => cortical deafness[1] |
| | 2. | pSTG | => PWD[1] |
| unilateral | 3. | left pSTG | => Wernicke's aphasia |
| **B. Subcortical lesions** | | | |
| bilateral | 4. | auditory radiations | => cortical deafness[1] |
| unilateral | 5. | deep left temporal lobe, probably affecting Rad and transcallosal fibers to A& pSTG | => PWD[1,2] |
| **C. Cortical and subcortical lesions** | | | |
| bilateral | 6. | right Rad + left A to pSTG projections | => PWD[2] |
| unilateral | 7. | left A to STG projections + transcallosal fibers from right A and pSTG to left | => PWD[2] |

1 - attested in Earnest et al. (1977)
2 - attested in Takahashi et al. (1992)

As Figure 2.1 and Table 2.4 show, primary auditory cortex is not always implicated in the aetiology of PWD. A possible model for PWD, then, is that while primary auditory cortex can still be largely intact (but see Coslett et al., 1984), one or more cortical fields in the posterior aspect of the superior temporal gyrus are dysfunctional, for whatever reason (i.e. direct lesion, underlying white matter lesion). Curiously, contemporary analyses were preempted by several 19th century researchers, including one particularly well-known 19th century neurologist and psychologist - confirming the suspicion that most of contemporary cognitive neuroscience constitutes a refinement of ideas introduced by 19th century cognitive neuroscientists. In this case, an excellent intuition about PWD and its lesion site was expressed by Freud in his 1891 Zur Auffassung der Aphasien. There he speculated that the lesion leading to word deafness is "not a simple disruption of tracts but incomplete bilateral lesions of auditory cortex" (Freud, 1891, as cited by Goldstein, 1974).

One final point about the lesions of auditory agnosia is worth emphasizing: in the cases of more or less 'pure' *nonverbal* auditory agnosia (see references above), the lesions have been unilateral and on the right (Fujii, 1990). Furthermore, in two recent cases of amusia without aphasia or word deafness, Peretz et al. (1994) report that the lesions are bilateral in superior temporal gyrus, but in more anterior aspects of the STG. Posterior STG was spared in those cases. Both of these findings are also consistent with the model that speech perception is mediated by bilateral posterior superior temporal gyri.

### 2.7 Are there subtypes of pure word deafness?

Given that there are different types of lesions leading to PWD and that there are distinct behavioral outcomes, one needs to consider the possibility that there are PWD subtypes. The concept of pure word deafness subtypes has been articulated before, notably by Auerbach et al. (1982). Auerbach and colleagues (see also Yaqub et al., 1988) suggested that word deafness exist in two forms. "The first is a disorder of prephonemic temporal auditory acuity and has been associated with bilateral temporal lobe lesions. The second is a disorder in phonemic discrimination that cannot be attributed to a difficulty in temporal auditory acuity and has been associated with left unilateral lesions. The former represents an apperceptive disorder secondary to pathology involving the auditory cortices whereas the latter may be a fragment of Wernicke's aphasia, especially of the word-deaf type" (Auerbach et al., 272).

There is some agreement in the literature that this view represents a sensible fractionation of PWD. For example, Albert & Bear (1975) argued that the PWD they attested in their patient was a consequence of a disorder in temporal processing. They found that their patient had a rate-dependent speech perception deficit: when spoken to at an abnormally slow rate, this patient's understanding improved significantly. Moreover, in a neurobiologically sophisticated review and analysis, Phillips & Farmer (1990) present the case for PWD as an auditory temporal processing disorder. They argue that speech perception requires intact time-resolution mechanisms in the time domain of milliseconds. Not only does speech perception require temporal sequencing of phonemes (or syllables) in time, but because critical stimulus information about the identity of speech stimuli is represented on the order of tens of milliseconds (notably formant transitions), a compromised cerebral time resolution mechanism will lead to severely impaired speech recognition. Because the type of temporal resolution they have in mind may be effectively done by neurons in primary auditory cortex (Heschl's gyrus), they suggest that (primary) auditory cortex lesions underlie PWD. In contrast to PWD that results from a temporal processing impairment, Auerbach et al. (1982) argue that in some cases temporal processing appears to be intact and the PWD has to be explained as an independent deficit in speech sound processing.

Based on the one hand on the architecture and connectivity of the higher auditory system (see Figure 2.1) and on the other on some recent functional neuroimaging results (see chapter 3), I propose a modification of this model (see also Table 2.4). The factors that come into play are (1) whether the lesion is bilateral versus unilateral, (2) whether the lesion is cortical versus subcortical, and (3) whether the lesion includes primary auditory cortex or is restricted to posterior STG.

In maintaining the central hypothesis that speech sound processing is mediated by the posterior superior temporal gyri in both hemispheres, the phenomenological predictions are as follows:

> 1. If the lesion includes primary auditory cortex and pSTG bilaterally, the disorder will be cortical deafness or sound agnosia including word deafness. This type of disorder has been called 'pre-phonemic,' and is consistent with the notion of impaired high temporal resolution mechanisms. To use an older distinction made in

the neuropsychology literature between *apperceptive* versus *associative* agnosia, this lesion would lead to apperceptive agnosia (Lissauer, 1890).

2. If the lesion is restricted to pSTG bilaterally, the patient presents with *pure* word deafness. Speech perception is fully compromised and speech sounds are predicted to be perceived as noise. Call this *profound word deafness.*

3. If the lesions are subcortical with cortical sparing (due to a unilateral or bilateral lesions; cf. Table 2.4), the syndrome is the one commonly designated as 'disconnection syndrome' by Wernicke and Geschwind (1965). Speech sounds are perceived as belonging to a foreign language. The recognition of speech as speech may be mediated by residual projections (thalamocortical or cortico-cortical) that are still intact, but the access to the areas is severely compromised (deafferenting of pSTG). This would be an instance of associative agnosia.

There does not yet exist sufficient information to distinguish systematically among the possible subtypes and to correlate lesion type with the deficits, therefore any conclusions about the sybtypes must remain speculative. It is obvious, however, that any predicted subtype will only be identified following careful and subtle psycholinguistic testing. The major clinical distinction is predicted to be a consequence of whether posterior superior temporal cortex is destroyed bilaterally or not. If pSTG is destroyed bilaterally, speech sound processing is predicted to be profoundly absent. The distinctions made here are uncontroversial. The only refinement is that an explicit distinction is made due to the role that pSTG is hypothesized to play.

## 2.8 Monkey lesion evidence: effects of bilateral auditory cortex lesions

There is some important evidence from the literature on auditory processing in monkeys (macaques, squirrel monkeys, and cebus) following unilateral and bilateral ablation of temporal cortex that speaks to the question of unilateral versus bilateral mediation of auditory processes. The disadvantage of studying monkeys when investigating speech perception is obvious. The advantage, however, is that one can make very precise lesions and test the animals at various stages of neurological impairment. It is acknowledged that drawing from comparative anatomy can be quite problematic in this domain. Animals are typically found to be less lateralized. In monkeys, for example, nobody has demonstrated consistent anatomical asymmetries like those observed in the human brain. Nevertheless, some of the major findings, summarized below, are quite suggestive.

In a review article of the superior temporal cortex lesion literature, Elliot & Trahiotis (1972) report that, in cat and macaque, *unilateral* superior temporal cortex lesions have little or no effect on discrimination of frequency, intensity, duration, complex sprectral differences, and other tasks. In contrast, *bilateral* lesions of the same areas have considerable effects on the execution of the same tasks. Importantly, the size of the lesion determined the extent of the deficit. That is, on average larger lesions led to worse discrimination, not complete elimination of discrimination performance. Iversen & Mishkin (1973) bilaterally lesioned either primary auditory cortex (area TC), the surrounding auditory association cortex (area TA), or both in macaques. When assessing the animals' frequency discrimination performance, they found that the TC or TA lesions caused mild deficits, but the combined lesion caused severe frequency discrimination deficits as well as affecting the frequency thresholds.

Some experiments speak more directly to questions of the neural basis of the auditory processing of complex sounds. Hupfer, Jürgens, and Ploog (1977) tested squirrel monkeys' ability to discriminate between conspecific vocalizations and non-conspecific but acoustically equally complex sounds (e.g. the vocalizations of other species). They tested the animals after small, medium, and large bilateral ablations of superior temporal cortex (presumed STG homologue). Only medium and large bilateral lesions led to severe difficulties in executing the previously learned discrimination tasks. Importantly, testing after the first (unilateral) lesion, whether it was right or left, revealed at most minor deficits. Only subsequent to the bilateral cortical lesions were the deficits profound.

A particularly important series of experiments was done by Heffner & Heffner (1984, 1986ab, 1989). They tested the ability of Japanese macaques to discriminate between different types of species-specific coo vocalizations after unilateral and bilateral lesions of the superior temporal cortex. They tested the macaques after the (right or left) unilateral ablation and after the bilateral ablation. Unilateral left lesions caused an initial discrimination impairment that resolved. In contrast, unilateral right lesions caused no detectable deficit in coo discrimination. Subsequent to the bilateral lesions, the animals became permanently impaired on the discrimination of species-specific vocalizations. Although they were able to discriminate between tones and coos, they could no longer discriminate within coos. Moreover, Heffner & Heffner (1989) were able to ascertain that the inability to discriminate species-specific coos was a consequence not of a generalized hearing loss but constituted a separate and independent auditory disorder (cf. PWD).

Finally, Colombo et al. (1990) tested auditory short term memory in cebus monkeys using a delayed matching-to-sample task before and after lesions of the superior temporal cortex. These investigators lesioned area TA ('secondary' auditory cortex), ostensibly sparing areas TC (primary auditory cortex) and TB (a further auditory association area). They reported that testing after the first (unilateral) lesion revealed at most mild deficits in task performance. However, after the bilateral lesion performance in the auditory short-term memory task was severely compromised. Colombo and colleagues argue that bilateral superior temporal cortex lesions lead to serious deficits in auditory short-term memory performance as assessed by their experiments but that it is not clear whether the deficits are due to the processing of the auditory information or its storage and/or recall.

In conclusion, the monkey lesion data also constitute evidence in favor of a model that posits bilateral processing of complex auditory stimuli such as speech sounds. These data, then, converge with the word deafness findings from humans. Comparable to the word deafness studies, every animal study shows that only bilateral superior temporal cortex lesions lead to the permanent inability to discriminate among and process complex auditory stimuli. Taken together the data are one critical piece of evidence lending support to the hypothesis, investigated in the imaging experiments discussed in Chapter 4, that speech sound processing is carried out by auditory cortex in both hemispheres, not just the presumed 'dominant' hemisphere. A valuable source of evidence that could speak to the bilaterality hypothesis is the split-brain literature. Demonstrating (or failing to demonstrate) recognition in the non-dominant hemisphere in a callosotomy case would constitute important evidence for the role that the right STG plays in speech perception.

### 2.9 Nagging questions, open issues

On the face of it, the existence of PWD suggests that there is a dedicated speech processor because it is possible to selectively affect speech perception. However, the evidence to date is not sufficiently rich to warrant that conclusion. There are two main shortcomings. First, although we have seen that word deafness is reported reasonably often, word deafness is rarely *pure* (see Table 2.2, last column; see also Buchman et al., 1986). Rather, there is always some impairment of non-speech sound processing associated with the lesion. To be persuaded that PWD unambiguously implicates a dedicated speech processor, one would want to see more cases of *pure* word deafness. One reason that pure cases may be rare is that lesions are typically large and presumably extend into distinct functional areas. If there is a speech recognition area (say, for example, posterior STG in both hemispheres), it is likely that only extremely circumscribed lesions that do not affect neighboring tissue will lead to PWD.

Second, the possibility has not been excluded that speech perception is relatively more impaired because it is an example of the processing of complex sound patterns (recall that PWD patients often do well with 'simple' auditory tasks like frequency, intensity, and duration discrimination with tones and non-speech sounds). For example, the lesion(s) could affect the signal-to-noise ratio for acoustic information and all speech perceptual deficits could be attributed to the problem of detecting patterns in noisy data. Indeed, some researchers maintain that speech recognition is an instance of pattern recognition of highly overlearned stimuli, and to effectively recognize the patterns, the integrity of the entire neural system is essential. To address such a position, one would want to test PWD subjects on other very familiar and overlearned stimuli (such as musical stimuli for musicians, say). If speech recognition is general pattern recognition overlearned for a special case, then both music and speech should be affected in comparable ways. A similar issue is being debated in the face recognition literature: some investigators conclude that there is a dedicated face recognition module while others maintain that face recognition is an example of processing extremely complex stimuli effectively because of the amount of expertise accrued over the years of exposure.

One question generated by these findings, and motivated more generally by the proposal of this thesis is this: if both hemispheres, specifically both posterior superior temporal gyri, play critical roles in the computation of speech signals, do both hemispheres support the same representations and processes or do they compute different aspects of the acoustic speech signal? This problem comes up again and again and is, in some sense, the next experimental question to pursue. Insofar as it is possible to ascertain that speech perception really is mediated bilaterally, one must begin looking into the computations executed in each area. One very general hypothesis to pursue in this direction is that both STG participate in early speech perception, but not to an equal degree. Although the number of architectonic areas in left and right superior temporal cortex does not differ, the sizes do differ, and size may matter.

# CHAPTER THREE

## NEUROIMAGING OF AUDITORY VERBAL STIMULI

This chapter summarizes the state of the art of metabolic functional imaging of auditorily presented verbal material. Both the early imaging studies and the results of the more recent positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) studies are discussed. Unlike previous reviews of the literature, however, I assess the results with respect to two particular factors in mind: first, was the verbal information presented passively or were subjects required to execute the demands of some experimental task? Second, was the activation reported bilateral in the superior temporal gyri or not. I derive two generalizations. (1) Speech stimuli appear to be processed in the superior temporal gyrus (STG) in both hemispheres. This observation is true for results across studies and distinct neuroimaging methods. (2) There is a suggestion of a systematic interaction between the presence or absence of experimental task demands and the physiological results: in the case of the overt execution of metalinguistic (meta-phonological) demands, the physiological results often reflect amplitude increases in the left hemisphere. In contrast, in the absence of explicit tasks ('passive presentation'), speech stimuli activate the posterior STG in both hemispheres and there is less hemispheric asymmetry. These two hypotheses are experimentally investigated in the MEG studies described in Chapter 4.

The analysis presented in this chapter speaks on the one hand to the functional organization of the speech recognition system, and on the other to an important methodological issue, namely that comparisons between task and no-task experimental conditions can be misleading in a very fundamental way by masking the role cerebral areas might play in processing particular information because they are selectively upregulated or downregulated by the attentional requirements of the experimental tasks.

## 3.1 Passive versus task-conditioned speech perception results

The neurobiological foundations of speech and language are 'frontiers' in the sense that for the first time one can systematically explore the functional anatomy of these cognitive faculties in humans by using different functional brain imaging methods. One way to carve up the studies in thinking about this literature is along the lines the language system is intuitively organized: there are studies on phonological processing, lexical processing, semantic processing, syntactic processing, and so on. A number of reviews discussing these studies and assessing the advantages and problems of the functional imaging methods have been published, including debates on the failure of a number of studies to converge (Demonet et al., 1993; Petersen & Fiez, 1993; Poeppel, in press; Stowe et al., 1994).

In this chapter, I focus on those functional imaging experiments in which auditory verbal stimuli were presented. The studies draw from a variety of stimulus modes: some studies presented non-word phoneme sequences, some presented single words, some presented connected speech in a familiar or unfamiliar language (Demonet et al., 1992, 1994; Knopman et al., 1980; Mazoyer et al, 1993; Kushner et al., 1987). There is, however, a unifying (albeit obviously arbitrary) way to classify all the studies that involve auditory speech stimuli, namely whether or not the subjects were asked to perform an explicit linguistic task on the stimuli; and whether or not the investigators detected bilateral superior temporal cortex activation. Categorizing the studies in this manner yields four possible patterns of results: no task/bilateral STG activation, no task/unilateral STG activation, task/bilateral STG activation, and task/unilateral STG activation. Upon reviewing the literature I argue that of the possible patterns two are attested more systematically: the no task/bilateral and task/left unilateral patterns. I suggest that these two patterns of results are due to the interaction of the speech perceptual systems and some aspect of the attentional mechanisms that interact with linguistic representations. Consistent with the hypothesis that speech perceptual computations are executed bilaterally in the STG and that the interaction with linguistic task demands appears to 'gate' the activation towards the left hemisphere, the task/right unilateral pattern is not attested, and the task/bilateral pattern is rarely attested.

## 3.2 The pioneering activation studies: FDG-PET and Xe-133 SPECT

Approximately since the 1970s it has been possible to study human cerebral activity in vivo. The first experimentally viable methods used fluoro-deoxyglucose PET (FDG PET) or single photon emission computed tomography (SPECT) with the 133 xenon inhalation technique to measure cerebral metabolism. (See Roland, 1993, for an extensive introduction to the techniques and review with an emphasis on the earlier literature.)

Before turning to the studies that measured the physiological responses to speech stimuli, what information is available about more basic hearing mechanisms? Unfortunately, not much neuroimaging work has been done which elucidates the psychophysical and psychoacoustic hearing mechanisms relevant to speech *and* non-speech processing. As a consequence of the sparseness of data available about fundamental psychophysiological responses of this nature, it is difficult to gauge the speech perception results described in the (by now large) literature against more basic information and build on knowledge of mechanisms presumably prior to speech perception. Examples of some relevant work include studies on tone discrimination (Roland et al., 1981; Cohen et al., 1992), tonotopic organization (Lauter et al., 1985), and the effect of rate of presentation on temporal cortex activation (Binder et al., 1994). The fMRI study by Binder et al. bears on the core issues of this thesis and will be diuscussed in more detail below. The tone discrimination studies (requiring subjects to detect same or different rhythms of 2 tone-sequence pairs or to detect the lowest in a sequence of tones) generated relatively complicated data on judgment and decisions about auditory stimuli. Moreover, the studies did not yield straightforwardly convergent results: Roland et al. found the most significant activation in the right superior temporal lobe including an extensive region at the right temporoparietal junction while Cohen et al. found mostly activity in the frontal lobes. Of the other psychoacoustic studies, the most curious is the experiment testing tonotopy in the human auditory system (Lauter et al., 1985). The authors presented two tones (500Hz and 2000Hz) to subjects while measuring activity with PET. They found that the two tones generated foci of increased activity in different parts of primary auditory cortex, and the activation pattern was consistent with the macaque physiology literature: more medial primary auditory cortex (koniocortex) responded better to higher frequencies. Since two sample points are not sufficient to generate a tonotopic map, and given that there is debate in the field about the extent of tonotopy, conclusions like the one argued for by Later et al. must remain speculative. The literature on functional imaging of hearing is reviewed in a recent article by Elliott (1994) and will not be discussed further.

Anticipating the results of this survey, there is one central finding relevant to the concerns of this dissertation: the pioneering functional imaging studies already provide solid evidence for the notion that speech sounds are processed bilaterally in the superior temporal cortex, not just in the presumed dominant hemisphere. In addition, the data turn out to provide hints of the interaction between speech perception and the presence of explicit task demands (cf. Table 3.1). Some of the major findings are briefly summarized below.

One of the seminal studies in this field was done by **Larsen, Skinhøj, and Lassen (1978)**. They had subjects recite the days of the week and months of the year while monitoring cerebral activity. In addition to some frontal foci of activity increases that included the supplementary motor area and the mouth motor area (bilaterally), the major blood flow increases were located in the posterior superior temporal areas bilaterally. The finding is particularly interesting because it is one of the few studies in the field in which subjects' overt (albeit rote) speech is monitored. Compared to the spatial and temporal resolution one can achieve with current techniques (both PET and fMRI) these older results seem extremely general. Nonetheless, these studies suggested that the superior temporal cortices appear to be differentially engaged when speech stimuli are heard by subjects.

A careful study by **Knopman et al. (1980)** used the xenon-133 inhalation technique to measure blood flow patterns in their subjects in three auditory conditions: baseline rest (quiet), noise burst intensity discrimination (500ms noise burst pairs), and detection of a target category ("something to eat") in a list of nouns. They subtracted the baseline responses from the two activation condition responses. In the noun-list/baseline comparison, the data from 11 (right-handed, male and female) subjects showed significant activation increase over baseline in the probe position over the posterior STG and a much more attenuated response in the homologous brain area in the right hemisphere. The tone-baseline comparison showed approximately the same pattern of results although the net rCBF increases were smaller, and there were no significant differences between the two conditions. Two points are worth noting: first, the left pSTG focus was the only consistent focus across subjects, and second the significantly larger left temporal cortex activation correlated with the presence of task demands (category membership detection).

In a motor control experiment in which subjects had to move their fingers through mazes, **Roland et al. (1980)** gave participants monosyllabic commands where to move their fingers. Listening to the commands increased rCBF bilaterally in the middle section of the superior temporal gyrus. The authors also reported Broca's area activation, in addition to numerous other, more specifically motor-control-related, sites. This experiment is relevant because, although the experiments included task demands, the demands were orthogonal to the speech input; i.e. there was no set of operations that had to be performed on the speech input itself. While this is not 'passive' presentation of speech stimuli, the study meets the criteria for inclusion because there was auditory verbal input and because they found bilateral STG activation. Clearly subjects had to pay attention to the verbal input in order to properly execute the task demands. The refinement of the hypothesis that these data are consistent with is that the attentional task might have to be a specifically linguistic task.

In a very prominent set of experiments, **Mazziotta et al. (1982)** investigated verbal and nonverbal responses using the FDG-PET (rCMRgl) methodology (see also Mazziotta and Metter, 1988). In the connected-speech condition, subjects had to listen to a Sherlock Holmes story. Although they were asked to listen attentively [FN: Subjects were paid according to how many details they remembered. This instruction may have lead to subvocal rehearsal in order to maximize the details remembered - which in the framework espoused in this dissertation would correlate with an activity increase in the left frontal operculum.] there were no other specific task demands. Major activity increases were seen bilaterally in the transverse temporal gyri and the superior temporal gyri, as well as the left frontal operculum, among other areas.

**Nishizawa et al. (1982)** presented their subjects with onomatopoetic utterances. Subjects were instructed to listen to these speech stimuli but no other task demands were made. The study ($^{133}$Xe method) generated increased rCBF bilaterally in the superior temporal gyri, angular gyri, frontal opercula, the frontal eye fields, and superolateral prefrontal cortex. Interestingly, with the exception of the superolateral frontal activation, all activation results had a larger amplitudes on the left, contrary to the hypothesis. Whether this amplitude-hemisphere interaction is accidental/artifactual or systematic and functionally consequential is unclear. It is interesting to bear in mind that this kind of asymmetry is an asymmetry one finds rather often in the imaging literature, regardless of whether one is considering metabolic physiological indices (as measured by PET orr fMRI) or electromagnetic recordings (as in EEG/ERP and MEG). In any case, these authors, too, find consistent bilateral STG activation in response to speech stimulation when no task is required of the subjects.

**Kushner et al. (1987)** measured participants' regional cerebral glucose metabolism (rCMRgl) while they listened to connected speech in a language they did not understand (Hungarian). Subjects were asked to monitor the speech stream for the occurence of a particular phonetic sequence (/honat/) and press a response key each time this non-word was detected. Note that Kushner et al.'s study was a 40 minute measurement of the phosphorylation rate of $^{18}$FDG. They authors report increased rCMRgl both in the left and in the right superior temporal cortex (as well as in left temporo-occipital cortex and right inferior temporal cortex).

**Bartlett et al. (1987)** measured the differential regional glucose metabolism (glucose phosphorylation rate, rCMRgl) during rest and while subjects listened to monosyllabic words. These investigators had participants monitor the speech stream for words beginning with /s/ or /b/. If a target was detected, participants had to press a response pedal. This is an example of a study in which there is simple auditory speech stimulation and a task pertaining to the stimuli - and the results show, importantly, that rather than relatively symmetrical bilateral activation showing up in STG, the activity increases are largely in the left hemisphere, including left Broca's area and Wernicke's area (defined as pSTG).

**Friberg and Lassen (1991)**, working with the intracarotid $^{133}$Xe technique to measure cerebral blood flow (CBF) and averaging the signal over approximately one minute (comparable to current PET activation studies with $^{15}$O), performed the following study. Subjects listened to connected speech in their native language (Danish). In one experimental condition the speech signal was heard normally (forwards) and in another the backwards. Subjects had to perform no specific task other than listening and understanding the speech signal. Subjects reported recognizing the temporally reversed speech signal as Danish but could not interpret the signal when played backwards. Interestingly, both presentations raised the CBF in the same cortical areas, including, crucially, the superior temporal gyrus (extending to its posterior part), the middle temporal gyrus, and the angular gyrus. The authors did not report whether the extent of activation was greater and more posterior for the speech versus the non-speech (backwards speech) stimuli.

**McLaughlin et al. (1992)**, using Xe-133 inhalation SPECT, monitored rCBF during three experimental conditions: binaurally presented connected speech (Danish), dichotically presented words, and white noise. When listening to connected speech, subjects were asked to listen attentively in order to remember the content, but otherwise there were no specific task demands. In the noise-connected speech comparison, rCBF increases were seen in left and right superior temporal gyri as well as left Broca's area.

Roland (1993) has summarized the results of many of these earlier studies in a very useful fashion by converting the results of a number of studies into Talairach coordinates for comparison. Roland's comparative illustrates that there are in essence three cortical regions that are consistently activated when subjects are presented with speech stimuli ranging from nonsense 'speech' stimuli (e.g. backward Danish; Friberg and Lassen, 1991) to nonwords (e.g. Nishizawa et al., 1982) to connected speech (Mazziotta et al., 1982): left superior temporal cortex (STG), right superior temporal cortex (STG), and left posterior inferior frontal gyrus (Broca's area). While left and right STG activations are predicted, especially in the framework of this thesis, the consistent left inferior frontal activation is less straightforward to interpret. Because the left posterior inferior frontal gyrus (Broca'a area) has been implicated in a number of distinct tasks that don't share any obvious components, there is disagreement about the role Broca's area plays in speech and language processing (cf. discussion in chapter 1).

Due to the resolution of these two imaging methods, the interpretation of the data must be cautious. In particular, the very low temporal resolution of FDG PET (activity is integrated over approximately 40 minutes) makes any generalizations about the neural basis of cognitive functions highly problematic. Nevertheless it is worth summarizing how the data pattern with respect to the generalization we are considering. Table 3.1 summarizes the facts with respect to the variables of interest.

**TABLE 3.1**

*Xe-133 and FDG imaging studies: bilateral activation and the presence of tasks*

| | stimuli and task | bilat.STG | uni. STG | area | method |
|---|---|---|---|---|---|
| Larsen | overt recital of rote forms, *task* | √ | | pSTG | Xe 133 |
| Knopman | monitor for category membership, *task* | | √ | left pSTG | Xe 133 |
| Roland | motor study with verbal commands, *no ling. task* | √ | | m/pSTG | Xe 133 |
| Mazziotta | connected speech, *no ling. task* | √ | | STG | FDG |
| Nishizawa | onomatopoetic utterances, *no ling. task* | √ | | STG | Xe 133 |
| Kushner | foreign language speech, detect targets, *task* | √ | | sup.temp. ctx. | FDG |
| Bartlett | detect phonemic targets, *task* | | √ | left pSTG | FDG |
| Friberg | connected speech, *no task* | √ | | m/pSTG | Xe 133 |
| McLaughlin | connected speech, *no ling.task* | √ | | pSTG | Xe 133 |

Key: STG - superior temporal gyrus; p - posterior; m - middle

Notice that the pattern of results falls in a particular direction, with only one exception to generalization: of the 9 studies, 5 had no explicit metalinguistic tasks and all six of them report significant bilateral activation in STG, among other areas, although there is some evidence for a left/right asymmetry without a task as well (Nishizawa et al). Of the four studies in which there were overt linguistic tasks, two pattern in the hypothesized direction (Bartlett, Knopmann) and two show bilateral activation (Larsen, Kushner). The study by Kushner et al. was done using FDG PET and required continuous auditory task performance, so bilateral activation is not all that surprising an outcome. In summary, these experiments help in constructing and clarifying the hypotheses that speech stimuli are processed in both superior temporal gyri and they are mostly consistent with the hypothesis that task demands alter the outcome such that one detects more activity in the left superior temporal cortex.

## 3.3 The recent activation studies: PET and fMRI

Most of the recent PET and fMRI activation studies that used auditory verbal presentation are extremely elaborate experiments, and it is beyond the scope (and intent) of this chapter to discuss each one in detail. Instead, I summarize the findings from the perspective of whether there were task demands or not and whether there was bilateral STG activation or not. Table 3.2 shows the combined results.

TABLE 3.2 HERE

**Petersen et al. (1988, 1989)** presented words binaurally at an approximate rate of 60 words per minute (variation due to word length). In addition to the STG activation, the (passive words - visual fixation) comparison yielded activation in the left anterior STG, left temporoparietal cortex, right lateral temporal cortex, and left anterior inferior cingulate cortex. (For a more detailed discussion of this study, see Poeppel, in press.)

**Wise et al. (1991)** studied six male subjects to whom they binaurally presented nonwords, noun-noun pairs, and verb-noun pairs at rates ranging from 40 to 60 words per minute. The experiment also included a rest condition and a verb generation condition. The rest condition yielded the control and subtraction data used for comparison with the different activation tasks. The non-word condition was a passive listening condition; in the word-pair activation conditions, subjects had to make particular categorizations. In addition to the STG activity, Wise et al. report foci of increased activation in left and right Heschl's gyrus, left Broca's area, left posterior middle frontal gyrus, and midline supplementary motor area. The results of this experiment are problematic for the hypothesis being developed in that the activation observed was the same in all three conditions, i.e. bilateral STG. The working hypothesis suggested that the non-word (passive listening) condition should have generated the data that are in fact attested, whereas the task-conditioned word-pair conditions are predicted to show primarily left-lateralized activation. In constrast to this predicted pattern, the task-dependent activations also led to bilateral STG activation. The amplitude changes in the left and right STG were also equivalent, ruling out the possibility that there was a significant amplitude modulation of the reponse as a condition of task presence or absence.

**Frith et al. (1991)** contrasted an 'extrinsic activation' task (lexical decision) with 'intrinsic activation' tasks (counting, enumerating words beginning with /a/). In the auditory lexical decision task, the activation was bilateral in the STG (Wise et al. do not provide the Tailarach coordinates), with the activation in the left slightly exceeding that in the right, but not so much that the authors considered the results to reflect significant hemispheric asymmetry (given the resolution of the method with respect to response amplitude measurements). Other foci included area 41 (bilaterally), left frontal pole (area 9), and right ventral prefrontal area 47. The authors report an interesting relationship between activation and the degree to which a task is internally or externally driven: when the task demanded verbal fluency and intrinsic generation (correlated with left dorsolateral prefrontal cortical activation), the superior temporal activation was inhibited. In contrast, lexical decision correlated with a (left) increase in STG activation and and the inhibition of left dorsolateral prefrontal cortex.

In **Howard et al.'s (1992)** very careful study on the neural representation of the mental lexicon and lexical access, involving both visually and auditorily presented words and controls, the authors found bilateral STG activation for the acoustic control task (hear a stimulus/say the same word each time) but unilateral left STG activation for a word repetition condition (hear a word/repeat word). In the hear/say control task, subjects listened to backwards speech (unrecognizable) and had to say the same word after each stimulus presentation.

**Price et al. (1992)**, in an important psychophysical experiment, varied word presentation rate from 10 words/minute to 90 words/minute. The stimuli were nouns. These authors found bilateral activation in the primary auditory cortices and in the middle and posterior STG at all presentation rates. The parametric aspect of the study generated the following finding of note: while the primary auditory cortices and the middle STG showed rate-dependent activity increases (higher stimulus rate correlated with higher rCBF), left posterior STG did not behave in the same way. Left posterior STG increased its activity uniformly if the stimuli were verbal. The authors argue that these results demonstrate the transformation of time dependent sensory input to time-invariant representations in left posterior STG (Wernicke's area).

A particularly provocative set of results was published by **Zatorre et al. (1992)**. They presented subjects with speech syllables in three conditions: passive listening (see Table 3.2), phoneme ('rhyme') judgment, and pitch judgment. In the passive syllables condition (compared with noise bursts), they found several bilateral sites in the STG. In addition, they reported foci in the left middle temporal gyrus, the right anterior STG, and the left inerior frontal gyrus. In the two conditions with overt task demands, the maximal activation foci they reported were dramatically different. Specifically, in the phonetic judgment condition, the control for which was the passive listening condition, the major focus was left Broca's area. In the pitch judgment condition, the control for which was also the passive listening condition, the major focus was in the right inferior frontal gyrus. They interpreted these results as suggesting that Broca's area plays a crucial role in the perception of phonetic contrasts. Noteworthy is the finding that the passive condition yielded bilateral STG activations whereas the task-dependent phonetic condition showed largely left unilateral (although not temporal cortex) foci and did not include any prominent activation of the left STG. (For extensive discussion of this study, see Poeppel, in press). The attent. Suppose that passive speech perception generates very large (ceiling level) activity in both STG. Then the attention-asymmetry hypothesis would predict apparent negative activation in right STG for the phonetic judgment condition.

**Mazoyer et al.'s (1993)** experiments included a number of conditions designed to fractionate language processing along some major seams, e.g. separating out acoustic, phonological, lexical, prosodic, syntactic, and conceptual processing systems. The experimental conditions included listening to a story in Tamil (not spoken by subjects), listening to a story in French (native language), listening to word lists, and listening to anomalous sentences. Importantly, Mazoyer et al. found that in all their auditory presentation conditions, the superior temporal gyri bilaterally showed significant activation increases, although the left increases were somewhat larger when the language was understood (i.e. French versus Tamil).

**Demonet et al. (1992, 1994a, 1994b)** have done a series of studies on phonetic processing. With respect to the questions at hand, their findings are the following: the comparison of a speech condition (in which subjects had to detect a phoneme in nonwords) with a tone control (in which subjects monitored tone-triplets for rising pitch), yielded bilateral posterior STG activation in two tasks. In addition, they found activity in Broca's area, the left fusiform gyrus, left inferior temporal gyrus, left motor cortex, and the posterior cingulate. These results lend support to the model of bilateral STG speech processing but does clearly not lend support to the notion that task demands modulate the processing such that the left temporal activity appears to be increased. The conditions were task-dependent and linguistic, and the activation was clearly bilateral.

The two most recent studies included in this analysis, by **Fiez et al. (in press)** and **O'Leary et al. (in press)**, are very relevant to the hypothesis. Both studies were concerned with the dependence of the rCBF results on the presence of tasks. Fiez et al. (in press) investigated temporal processing aspects of speech perception, guided by the hypothesis (developed by Tallal and colleagues) that the left temporal cortex and possibly the left frontal operculum play a special role in processing the rapid temporal transitions characteristic of phonetic stimuli (formant transitions). Fiez et al. used task conditions which are relevant to how non-speech and speech stimuli differ when there are tasks or not. Specifically, they presented speech stimuli (vowels, CV syllables, and CVC words) in passive conditions and in conditions where a feature had to be overtly detected. The critical finding - from our perspective - was the more lateralized response in the attentionally conditioned task. In a comparison of a target detection condition with a passive listening condition, the left STG was significantly more active in the detection condition than the passive condition; a similar effect for the right STG did not hold. In short, while there was significant bilateral STG activation in both speech stimulus conditions, the presence of the task skewed the results to reflect increased left STG activation, as predicted by the hypothesis. Interestingly, in a direct comparison of linguistic versus non-linguistic stimuli, the *right* STG showed a significantly higher response for linguistic stimuli than the left posterior STG, contrary to common intuitions and other findings in the field.

O'Leary et al. (in press) measured rCBF in 10 subjects while they listened to environmental sounds, phonologically possible nonsense words (pseudowords), or words that were presented either binaurally or dichotically. The baseline measurements were established by monitoring rCBF to infrequent (1 every 10 seconds) tone bursts. In the binaural conditions subjects were merely instructed to listen attentively. In the dichotic conditions, the subjects were required to attend either to the left ear or right ear input. The passive binaural condition generated symmetrical rCBF increases in posterior STG for all stimuli. The dichotically presented stimuli resulted in asymmetric activation of the STG: rCBF increased contralateral to the attended ear and decreased in the other temporal lobe. The authors interpret their findings as suggesting that sounds, meaningless words, and words yield bilateral STG activation. Dichotic stimulation generates results that alter the activation pattern in the temporal lobes. The authors argue that the differential activation in the dichotic listening conditions is a consequence of attentional mechanisms that are differentially allocating resources to the hemispheres based on the fact that the auditory areas of the two hemispheres are suited to certain kinds of processing. Attention is spatially directed, not conditioned by the *type* of processing. The elaboration of this hypothesis with respect to dichotic listening was first made explicit by Kinsbourne (1973), who saw attentional gating is an alternative to the explanation of the dichotic results (suppression of ipsilateral projections) presented by Kimura (1961). Very recent PET results by Hurtig, O'Leary and colleagues replicated the earlier findings and confirmed that the results were likely due to the interaction of attentional mechanisms with the speech stimuli.

Two recent fMRI studies speak to the issue of the involvement of the bilateral STG in speech sound processing. In a major functional MRI study of auditory speech perception, **Binder et al. (1994)** investigated the response of temporal cortex to white noise, pseudowords (monosyllabic nonsense strings), words (monosyllabic concrete English nouns), and connected speech (story by Tolstoy). Stimuli were presented at the rate of about one word per second. The authors observed signal increase after the onset of stimuli in the STG in all conditions. Word or pseudoword presentation activated a significantly larger area than noise. With the noise stimulus the activity was typically more confined to areas on or immediately around Heschl's gyrus. In response to speech stimuli the authors observed larger areas of activation including large parts of dorsal STG and extending to the superior temporal sulcus in all subjects. The authors did not find any consistent response differences among the speech conditions. Importantly, there was no noticable asymmetry in left and right STG activations. Note, also, that there was no task. **Maddock et al. (1994)** also recently reported on some fMRI results on the processing of auditorily presented words and pseudowords. In three different word conditions the common finding was that bilateral STG were activated, again supporting the notion that *at least* speech sound processing (and perhaps other aspects of language processing and representation) are mediated bilaterally.

To conclude the discussion of the functional imaging results, the generalizations stated at the beginning are problematic but not completely unreasonable. There are two redeeming features: the hypothesis was not explicitly addressed (and destroyed) by any of these studies, and second the worst possible pattern is never attested, namely task-driven measurements in which the right activations exceed those in the left. Bilateral activation of STG is a relatively standard pattern, and often the presence of experimental task demands alters the physiological outcome in the hypothesized direction. Beyond the extent to which the findings summarized in Table 3.2 support or weaken the hypothesis, the data illustrate some important points, two of which deserve emphasis. First, no study has presented auditory verbal stimuli and <u>not</u> found STG activation, to my knowledge. Second, despite vastly different experimental requirements, measurement techniques, and data-analytic strategies, the findings are reasonably consistent from an anatomical point of view. Consider the Tailarach coordinates used to standardize the results. Given that this sort of standardization introduces considerable error, the range of values is reasonably small. In the left temporal cortex, the maximal distance between values reported for peak STG activations is 1.2 cm (medio-lateral), 3 cm (superior-inferior), and 1.8 cm (anterior-posterior). In the right temporal cortex, the maximal distance between values reported is 1.8 cm (medio-lateral), 2.8 cm (superior-inferior), and 1.6 cm (anterior-posterior). While this level of variability would be at best disappointing within a single study's dataset, the values are quite encouraging given that they were acquired with extremely varying methods. In a nutshell, it has been possible to localize the rCBF increases associated with speech perception to the superior temporal cortex with a precision on the order of 2 cm. While 2 cm of cortex presumably extends over more than one functionally homogeneous cortical area, the exciting possibility has now been created that one can investigate these cortical regions parametrically in order to understand what kinds of features of auditory verbal stimuli differentially engage these areas.

## 3.4 Discussion

The hypothesis that speech sounds are processed bilaterally in the posterior STG holds up reasonably well when considering the metabolic brain imaging studies. The hypothesis that the presence of overt, perhaps meta-linguistic (meta-phonological) tasks makes the rCBF results appear to support a stronger lateralization model is, for the moment, less defensible. Regardless of the specifics of that hypothesis, however, it is relevant to evaluating the literature that having a task may alter results in a systematic way. In particular, the performance of some attentionally demanding or unusual task during the measurements can affect the results in a way that makes it extremely difficult to evaluate whether one is measuring the task or actually tapping into the neural basis of the cognitive system supposedly under investigation. Insofar as the question is about the neural basis of the speech and language system, it is worthwile to work with no-task conditions and present speech and language stimuli passively. The automatic processing associated with passive presentation approximates the natural computational processes much better in the sense of ecological validity. Furthermore, in terms of interpreting the results, presumably the dependent variable of interest is the physiological response. It may be beneficial to have additional behavioral data to confirm that subjects are in fact executing particular task demands. However, one should be cognizant that one may be providing more evidence about particular *task demands* or the *interaction* of linguistic-computational systems with attentional systems than about the neurobiology of the speech and language system. If the goal of cognitive neurocience research in the domain of speech and language is to understand the neural basis of speech and language processing, one needs to focus on understanding how the brain mediates the linguistic representations that form the basis for the execution of task demands, rather than being misled to focus on the tasks themselves. Presumably it is representations that are represented and computed in the brain, not task demands.

An final methodological conclusion arises from this analysis: insofar as it is true that task and no-task experiments yield systematically different results, one must be extremely cautious in the comparison across studies. For example, the direct comparison of passive conditions with task-driven conditions is not appropriate. If the analytic strategy was based on paired-image subtraction (Fox et al., 1988) it is especially problematic to interpret results derived by such comparisons unless one is aware that one generated fundamentally distinct data to begin with. That is not to say that the approach of studying language using these techniques cannot be profitably employed. Insofar as one is explicit that one is assessing language processing *and its interaction with task-demand execution*, one can presumably gain insight into the neurobiology of language through the window of language's interaction with processing systems of attention and memory.

Practically every study presenting auditory stimuli has found increased activity bilaterally in primary auditory regions (Brodmann's area 41) as well as auditory association areas including STG (areas 41 and 22). Is it possible that pSTG does 'generalized' auditory analysis and therefore always lights up but plays no special role in speech perception? That possibility is unlikely, on both empirical and conceptual grounds. Empirically, it is established that activation in and immediately around Heschl's gyrus, including the portion of STG lateral to Heschl's gyrus, increases for a wide range of auditory stimuli, including, for example, clicks, tones, musical stimuli, backwards speech, non-words, and connected speech. However, posterior parts of STG in both hemispheres seems to depend on more than auditory presentation; typically, posterior STG and temporoparietal cortex shows activation increase in response to speech stimuli . In contrast, tones, clicks, and other non-speech stimuli do not appear to drive pSTG as effectively, if at all (Petersen & Fiez, 1993).

The discussion of pure word deafness in Chapter 2 ended with the question of whether the cortical regions in posterior STG perform the same (kind of) computations or different computations from each other. Both pure word deafness and neuroimaging implicate the posterior STG in the processing of verbal information, especially when it is presented auditorily, but in neither literature are there any serious hints as to what the nature of the computations might be. The simplest hypothesis is that both STG execute the same analysis of auditory speech stimuli. This simplest hypothesis cannot be exactly right given that STG interacts differentially with attentional mechanisms. If the identical computations were executed, and assuming that auditory attention can be distributed equally across the hemispheres (a weak assumption in the visual domain), one might expect to observe the same kinds of responses in the STG in both sides rather than a modulation of left and right activations dependent on the presence of task demands. The only conclusion that one can safely draw at this juncture is that pSTG areas typically activate bilaterally to passively presented speech stimuli.

# CHAPTER FOUR

## MEG STUDIES OF SPEECH PERCEPTION

### Introduction

This chapter presents new empirical research using the brain imaging method *magnetoencephalography* (MEG). The deficit-lesion correlation approach, exemplified by the pure word deafness results summarized in Chapter 2, provides one with information about a cognitive system's steady state subsequent to brain injury. The metabolic neuroimaging studies discussed in Chapter 3 have a temporal resolution on the order of seconds and provide a measure of activity sampled over a large region. In contrast, MEG measures neuronal activity with (sub)millisecond resolution. The spatial information obtainable with MEG data alone is not as good as that obtained with, for example, fMRI. However, the combination of MEG with structural MR images from the same subjects (magnetic source imaging, or MSI) provides remarkably high-resolution information about the neural basis of sensory, motor, and cognitive processes (Näätänen et al., 1994).

The experiments investigate the contribution of left and right supratemporal auditory cortex to the processing of speech signals. The review of the results from pure word deafness (Chapter 2) and from metabolic imaging (Chapter 3) converged on the finding that the posterior superior temporal gyri (STG) in both hemispheres are implicated in speech-perceptual processes, although the precise nature of the computations is, of course, not understood. In addition there was some data that hinted that the outcome of the imaging results can depend on the presence or absence of an overt linguistic task. The presence of overt task demands often appeared to correlate with a change in the results such that the cognitive function under consideration appeared to be mediated by the left temporal cortex. Given that the superior temporal gyrus is 'downstream' from primary auditory cortex both anatomically (receiving its main inputs from koniocortex and parakoniocortex) and functionally (by assumption), one simple hypothesis to investigate is that auditory cortex generates more or less equivalent kinds of responses on the left and right sides in unattended conditions. By hypothesis, then, the bilateral hemispheric symmetry of responses seen in the superior temporal gyri will also be reflected in the MEG recordings from the auditory cortices.

Note that a 'bottom-up' perspective also motivates the (simpler) hypothesis that (at least primary) auditory cortex processes information the same way in both hemispheres. The most direct anatomical pathway from the periphery to the auditory cortex (cochlea -> cochlear nucleus -> inferior colliculus -> medial geniculate body -> auditory cortex, ignoring the parallel projections involving the superior olivary complex and the nuclei of the lateral lemniscus) is not associated with any obvious asymmetries. The majority of projections from both cochlear nuclei cross over and initiate a series of parallel paths. There is no particular feature of the ascending system that would suggest asymmetries, at least at the morphological level (Geschwind & Galaburda, 1987).

On the basis of the above considerations, the following questions are investigated: (1) is the processing of speech sounds in the auditory cortices qualitatively and/or quantitatively the same or not? (2) Do task demands (attentional demands) modulate the neural activity in left supratemporal auditory cortex?

An essential feature of using relatively simple speech stimuli is that one can take advantage of the known physical basis of the signals. In particular, one can take advantage of the well-understood difference between the acoustic/physical properties of stimuli and properties of the stimuli that are of linguistic consequence. The stimuli are digitally synthesized and single variables can be manipulated parametrically. The effects of the variation of certain parameters on the acoustic signal and the percept have been systematically assessed, and these variations in the acoustic signal can now be correlated with the behavioral and physiological measures that are acquired. In effect, the approach advocates a *psychophysics of speech*, and the subsequent *correlation of psychophysics with physiology*. This aspect places the research program into the Helmholtzian psychophysiological tradition that has yielded a wealth of information in other psychological domains. The hope is that using these well-understood stimuli one can learn in a paradigmatic way how categorial representations in the verbal domain are represented and processed. In other words, if we understand how phonetic categories are processed and effectively bridge the gap between anatomy and phonetics/phonology using a physiological approach, perhaps our understanding will extend to other linguistic categories.

A second reason for using acoustic-phonetic stimuli is that they are easily constructible and (in a non-technical sense) simple: using phonetic stimuli minimizes interference from so-called 'complex' levels, i.e. levels with which we intuitively associate meaning, notably lexical items and phrases. A given response to an experimental task will be minimally contaminated by other representations, access to which is not required for task performance. In the ideal experimental case one could study phonetic and phonological processing without any reference to the level of morphemes, listed lexical items, or larger phrases. This experimental desideratum, however, is not realistic. It is also not necessarily theoretically coherent. Although there are independent levels of description and explanation, these levels are interdefined and rely in crucial ways on each others' representations. While phonetic stimuli possibly *minimize* the interference due to higher levels, it is an empirical issue whether one can effectively *exclude* interference from other linguistic levels. For example, it is possible that every phonetic stimulus initiates an automatic search for a lexical item beginning with or including that sound (Forster & Bednall, 1976). In fact, there is ample experimental evidence that linguistic representations (sometimes called 'codes') not required for the execution of a task are nonetheless activated. Following Fodor (1983) one could argue that the computational system is such that all components of the system are automatically engaged when a stimulus is recognized as linguistic in the broadest sense (cf. Mattingly and Liberman, 1988). Exposure to a linguistic stimulus simply triggers the system, which then mandatorily computes all the relevant representations. Importantly, this view implies certain fundamental processing similarities between language and other systems, notably vision. That is, barring gross pathology, a stimulus of a given modality will set off the system's computational machinery automatically. When confronted with a visual stimulus, for example, the visual system necessarily computes representations of color, form, motion, depth, texture, etc., and any available cue is used to construct perceptions that the organism's visual system can interpret. Note that it is not possible to view a scene and *explicitly not* compute some perceptual attribute like color or texture or stereoscopic depth. It is proposed that in a similar fashion the linguistic cognitive system takes advantage of all available cues to construct interpretable utterances. The speech perception subroutine of auditory language comprehension will automatically generate linguistic output based on acoustic information. In other words, the speech perception module acts much like other sensory modalities. Whether or not components of this cognitive system compute the input representations in an encapsulated fashion is not at stake. Relevant is the hypothesized property that the entire speech and language system is triggered into action automatically and opportunistically (using any available cue) constructs representations that are interpreted.

An important consequence of using phonetic stimuli is that one can actually compare the physiological results with existent physiological results from human studies and studies investigating other species. For example, several investigators have recorded in different parts of the monkey auditory cortex using speech stimuli of precisely the sort that are used in the experiments discussed in this chapter (notably Steinschneider et al., 1994, 1995). It is important to connect with the animal results because it may help make the human results interpretable by clarifying the underlying mechanisms generating evoked fields.[1] The substantial literature on speech sound processing at earlier stages of the auditory pathway in macaques and other mammals, including the auditory nerve does not speak directly to the problem of how speech sounds are coded in the human auditory cortex and will therefore not be addressed further (Popper & Fay, 1992; Edelman et al., 1988).

## Magnetic Source Imaging (MEG and MRI)

In the MEG experiments discussed below the region of interest is defined a priori (in this case supratemporal auditory cortex in both hemispheres), and the experiments are designed to investigate questions about the activity of that region of interest in response to particular auditory stimuli. Previous electrophysiological research (using MEG, EEG, subdural grids, and deep penetrating electrodes) has demonstrated that electrophysiological responses are sensitive to linguistically meaningful distinctions (MEG: Eulitz et al., 1995; Kuriki & Murase, 1989, Kaukoranta et al., 1987; Hari, 1991; Sams et al, 1991, Hayashi et al., 1992, and others. EEG/ERP: Neville et al., 1989, and many others. Subdural grid electrodes: Boatman et al, in press. Penetrating electrodes: Steinschneider et al., 1994, 1995; Ojemann, 1991; Nobre et al, 1994; Creutzfeldt et al., 1989). MEG source localization algorithms are well suited to localization and outperform EEG algorithms. Commercially available dipole source localization algorithms perform extremely well, particularly in studies of the auditory system, localizing single equivalent current dipoles to areas of the auditory cortex with confidence volumes of $0.01 cm^3$ and better. One important reason for this difference is that EEG measures electric fields generated by intracranial

---

[1] Insofar as the electrophysiological results from the MEG recordings converge in clear ways with some of the animal results from single and multiple unit recordings, one can argue with more confidence that the instrument is making measurements from auditory cortex that have a reasonable interpretation. Moreover, there are postulated similarities between the auditory systems of macaques and humans, so that the evidence from macaques and other species serves as an important constraint.

volume currents, and electric fields are attenuated by the conductivity properties of tissue and bone. In contrast, MEG measures magnetic fields, which are not grossly distorted. For physiological, physical, and geometric reasons one can be very confident from which cortical area one is recording with MEG. Consequently it is easier, using appropriate constraints, to reconstruct the current distribution underlying the generation of evoked neuromagnetic fields.[2]

### The neurophysiological foundations of magnetoencephalography

Recall the fundamental observation from physics that all currents (moving charges) are associated with magnetic fields (typically denoted as **B**). The strength of the magnetic field **B** associated with a given current is given by the Biot-Savart law, or, generalized in the framework of Maxwell's equations, by Ampere's law (see Halliday and Resnick, 1988, for an introduction to electromagnetism and derivations of these equations). In the context of neuromagnetic recordings, two properties of magnetic fields as described by the Biot-Savart law or Ampere's law are noteworthy. First, the magnetic field associated with a given current is always in a particular spatial-geometric relationship to the current. Specifically, the field is orthogonal to the current and has an orientation described by the 'right-hand rule': if one points one's right-hand thumb along the direction of current flow, the associated magnetic field lines lie in the direction of one's curled fingers, i.e. in concentric rings around the current. Second, the strength of the magnetic field (if generated by a current dipole) falls off with the square of the distance from the current. (If the field is

---

[2] Because the sensors used in MEG are 'blind' to sources that are oriented in a particular direction in the brain (radial sources), one can exclude certain areas as contributing to the extracranially measured field. Of course, this advantage is a serious disadvantage if one is attempting to study parts of cortex that have predominantly radial orientation of dipole sources (gyral surfaces, for instance). One ultimately wants to combine the sulcal recordings from MEG with the gyral fields from EEG to obtain more extensive data. Nevertheless, insofar as one is studying auditory cortex, MEG is a good choice because large portions of auditory cortex lie on the superior surface in the Sylvian fissure. That means that the majority of cells (specifically, the majority of dendrites of pyramidal cells in A1) are oriented tangential to the cranial surface and can therefore be effectively recorded with MEG.

generated by a current better modeled as a quadrupole, the field falls off with the cube of the distance from that current source.) In what way do these fundamentals of physics apply to the measurement of *neuromagnetic fields* and the reconstruction of their origin in time and space?

The natural connection, of course, is that much of neuronal activity *is* electromagnetic activity. In particular, there are different classes of current flow in nervous tissue. Current is associated with action potentials, transmembrane potentials (ionic currents and displacement currents), intracellular post-synaptic dendritic currents, various extracellular volume currents, and others. As given by the Biot-Savart law, all fields associated with neuronal currents superpose to yield the net magnetic field. In principle, the net magnetic field reflects the superposition of the magnetic fields generated by all current elements. However, only the magnetic fields generated by certain current elements are measurable extracranially. Swinney and Wikswo (1980), for example, have shown that transmembrane currents (ionic and displacement currents) contribute very little to the measured magnetic field because transmembrane currents are typically radially symmetrical around the neuronal processes and therefore cancel each other.

Two types of neuronal events can generate extracranially measurable magnetic fields: currents that are associated with action potentials and currents that are associated with postsynaptic potentials (currents in dendritic shafts). Both of these currents are intracellular events. The field one can measure at the scalp, generated by axonal current or dendritic shaft current, is conditioned by (at least) three factors: (1) the spatial and temporal characteristics of the underlying current distribution, (2) the geometry of the head (including the morphology of the brain), and (3) the electrical and magnetic conductivity and permeability properties of the head (i.e. the conductivities of the scalp, the skull, the cerebrospinal fluid, and neural tissue).

To develop an intuitive grasp for the relationship among currents, magnetic fields, and neuromagnetic measurements and to get an understanding of how the three factors mentioned above constrain the measurements, the central concept to keep in mind is the *current dipole*. A dipole is defined as two equal but opposite charges between which primary current flows in a straight line and between which return currents flow in a complex pattern of field lines, symmetrical around the primary current. I first work through the example of the magnetic field generated by postsynaptic current dipoles in dendritic shafts (see Lewine and Orrison, 1995, for a very detailed explanation).

The precise configuration of postsynaptic current flow in the apical dendrites depends on two postsynaptic properties: first, is the synapse excitatory or inhibitory, and, second, where along the dendritic shaft is the synapse. If the synapse is somewhere along the middle of the dendritic shaft, the consequence is a quadrupolar current configuration, and therefore not readily visible to MEG measurements. There remain, then, four possible current configurations that yield measurable fields. The synapse can be excitatory or inhibitory and distal or proximal to the cell body. The four configurations generate two net current flow outcomes, towards or away from the soma. A dipole *towards* the soma comes from excitatory, distal synapses and inhibitory, proximal synapses. A dipole *away from* the soma comes from excitatory, proximal synapses and inhibitory, distal synapses. This illustrates that there is some ambiguity associated with the origin of the fields even at the level of the single cell. Nevertheless, these considerations show that one can strongly constrain the interpretation of where and how neuromagnetic fields are generated.

A single postsynaptic event does not create a neuromagnetic field that is measurable at the surface of the scalp. Lewine and Orrison (1995) estimate that, given the small size of the current associated with a single postsynaptic event, approximately one million synapses have to be synchronously active to yield the type of evoked field that one typically registers at the surface. Since there are over 100,000 pyramidal cells per $mm^2$ of neocortex, and since each pyramidal cell is associated with 1,000 or more synapses, coherent activity of 1,000,000 synapses is not an unrealistically large number. Lewine and Orrison speculate that "as few as one synapse per thousand is probably adequate to generate a detectable extracranial signal" (p. 376).

Consider, in comparison, how action potentials might contribute to the extracranial field recordings. An axonal impulse moving from the soma towards a synapse is defined by a leading edge of depolarization and a trailing edge of repolarization. Because of the two polarization 'directions' (the action potential can be characterized by a current dipole away from and a current dipole towards the soma), the axonal impulse is a current quadrupole. Importantly, the magnetic field associated with a current quadrupole falls off with the *cube* of the distance, unlike a current dipole (which falls off with the square of the distance). Given the field decrease with the cube of the distance, in order to generate a field strong enough to be recorded at the scalp, an enormous number of neurons have to fire action potentials simultaneously. Lewine and Orrison (1995) suggest that it would require the synchronous firing of hundreds of thousands of pyramidal cells to generate a measurable signal. They argue that "the short duration and restricted spatial extent of the intracellular component of action currents make it extremely unlikely that sufficient temporal and spatial synchrony of events is generally obtained" (p. 376). Consequently, axonal impulses have a considerably more restricted role in the generation of extracranially measured magnetic fields.

The overwhelming source of neuromagnetic fields measured at the scalp, then, is believed to be the current flow in apical dendrites of pyramidal cells. The cell bodies of these pyramidal neurons are typically in lamina 3 (Cajal, 1911). Because of their number, their appropriate morphology (open field configuration, unlike, say, cortical stellate cells), and their geometry relative to the head (apical dendrites oriented parallel to skull surface), the concurrent activation of pyramidal neuron apical dendrites in response to an external or internal stimulus is the most likely source of the scalp-recorded evoked neuromagnetic fields.

## How is the neuromagnetic field measured?

The neuromagnetic signal associated with the activity of neurons (even thousands of neurons) is extremely small. For example, the earth's steady magnetic field is 8-9 orders of magnitude larger than the average auditory evoked response we are measuring, and even the field associated with cardiac activity is 2-3 orders of magnitude larger than a cortical evoked response. Some typical values for familiar magnetic fields are shown in Table 4.1.

**TABLE 4.1**

*Approximate values for some illustrative magnetic fields ($1T=10^{15}fT$)*

| | | |
|---|---|---|
| Field applied in magnetic resonance imaging | $10^{15}$ | fT |
| Near a small bar magnet | $10^{13}$ | fT |
| At the surface of the earth (earth steady field) | $10^{11}$ | fT |
| Urban noise | $10^{9}$ | fT |
| Abdominal currents | $10^{5}$ | fT |
| Cardiac activity, skeletal muscle | $10^{4}$ | fT |
| Epileptic spikes | $10^{3}$ | fT |
| Cortical evoked field | $10^{2}$ | fT |
| SQUID noise | 10 | fT |

Given the exceedingly small signal and the large amount of background noise, the challenge is to extract biomagnetic signals reliably. Four factors contribute to the successful measurement of evoked neuromagnetic fields. (1) Extremely sensitive superconducting sensors (Superconducting Quantum Interference Devices or SQUIDS) register the neuromagnetic fields. (2) The measurements using the SQUIDS are performed in magnetically shielded recording chambers to exclude extraneous magnetic fields as effectively as possible. (3) Large sensor arrays (instead of single SQUIDS) permit the simultaneous acquisition of neuromagnetic signals from large regions of the head. (4) Repeated stimulus presentation and signal averaging help to generate satisfactory signal-to-noise ratios. Each of these factors is described briefly below.

(1) The technological innovation at the core of magnetic field measurements is the exquisitely sensitive detector, which consists of one or more detection coils inductively coupled to a superconducting quantum interference device, or SQUID. A magnetic field near the coils forces current through the wire-loops of the detector. The coupling of the coils to the SQUID allows the induced current to be measured and translated into voltage (electronics). The current forced through the SQUID detector coils is proportional to the strength of the magnetic field through the coils. In detail, the detection process is extremely complex, involving superconducting and cryogenic technology, because the detectors must be maintained at superconducting temperatures in order for the quantum effect (Josephson junction) to consistently convert field measurements to electronics. The cryogenic aspect of the task is commonly accomplished by suspending the detectors in liquid helium. There are several types of detectors, distinguished by how many coils there are and how the coils are positioned and wound: magnetometers, planar gradiometers, first-order gradiometers, and second-order gradiometers. The type of SQUID-based sensor used in the experiments discussed in this chapter is an axial first-order gradiometer. In a first-order gradiometer, the wire loops of the detector are separated along the long axis and wound in opposite directions and the SQUID measures the difference in magnetic field between the loops. The advantage of such a sensor design is that far-away fields do not generate significant magnetic flux density differences between the two axially separated coils but near sources do (field falls off with square of distance). These detectors, then, are effective at registering fields generated by nearby sources, such as the brain, and are equally effective at 'ignoring' far-away sources. Figure 4.1 schematizes the logic of how a SQUID-based

sensor (first-order axial gradiometer) registers the magnetic field generated by an intracranial current source.

FIGURE 4.1 HERE

(2) The magnetometer recordings are usually done in a magnetically shielded room (MSR). MSRs are made of materials that have high magnetic permeability. When an extraneous magnetic field encounters the high permeability material, the field is diverted from its path (for example through the recording chamber) along a path defined by the high permeability material. "The shielding capability of the room varies as a function of the frequency characteristics of the noise. For a typical room, magnetic noise of less than 0.1Hz is shielded by only 20 to 30dB, whereas higher-frequency signals (e.g. 100Hz) are attenuated by as much as 50 to 60dB" (Lewine and Orrison, 1995, p.384). Because lower-frequency signals can influence the field as measured at the detector despite the shielding effects of the MSR, a sensor like an axial gradiometer further helps to limit the measurements to the sources of interest.

(3) Because single SQUIDS sample only from a very restricted surface area over the scalp and because one needs to sample the evoked field as broadly as possible to reconstruct its origin effectively, contemporary experiments are typically done using a magnetometer with a large sensor array. Characteristic time-varying evoked magnetic fields recorded from several channels simultaneously are shown in Figure 4.2. In this case, the stimulus was a 300ms auditory stimulus (vowel /a/). The measurement period extended from 100ms before stimulus onset to 200ms after stimulus offset (total epoch duration 600ms). On the surface, the recordings look very much like EEG recordings of evoked responses. Note that even in a single epoch one can detect the large deflection around 100ms. This response component is the M100 (or N1m), comparable to the N1 component in ERP recordings. (The standard auditory evoked response assessed by MEG has three major components, at

50ms, at 100ms (the M100 or N1m component), and at 180-200ms. In addition, Pantev et al. (1994) and Eulitz et al. (1995) have argued that acoustic stimuli generate a sustained field (SF) that remains visible for the duration of a stimulus and ceases with stimulus offset.)

FIGURE 4.2 HERE

To illustrate multi-sensor array recordings, Figure 4.3a shows the layout of 37 sensors, which are placed in a hexagonal grid arrangement. The recording surface (the 'electrode-tip,' as it were) is 14.4cm in diameter. The data shown represent the identical epoch depicted in Figure 4.2.

FIGURE 4.3 HERE

(4) To optimize the signal-to-noise ratio, it is necessary to do signal averaging. Figure 4.3b shows the average for 100 trials of the vowel /a/. Figure 4.3c shows the data of Figure 4.3b low-pass filtered (1-20Hz). In the low-pass filtered data one can discern the distinct response components very clearly. Importantly, the dipole localization algorithms employed work effectively on data smoothed by low-pass filtering. Figure 4.3d represents a contour map of the magnetic field at a particular slice in time (108ms, the M100 peak). The magnetic field generated by a dipole underneath the measurement surface ( a so-called 'half-space') is defined by regions of magnetic flux emerging from the skull on one side and reentering into the skull the other. The contour map separates field lines coming out of the skull and reentering the skull, in other words it is a representation of the field direction (into or out of the skull) and amplitude at various points. The lines are iso-field lines, i.e. the field is of equal strength and points into the same direction at any point along a particular isocontour line. The source underlying the field lies underneath the contour, most likely somewhere along the line between the polarity reversal of the field (thick line).

## Reconstructing the current source parameters: dipole modeling

With the data averaged and smoothed, one can now estimate the parameters of the current sources that underlie the magnetic field. The process of estimating the current source parameters is known as dipole fitting or modeling. The problem of fitting single or multiple dipoles to extracranially measured fields is an instance of an inverse problem, which is computationally ill-posed: there are no unique solutions to inverse problems. In absence of constraints, one cannot find unambiguous inverse solutions for magnetic measurements. Because inverse problems are very common in the sciences and particularly challenging (as a classical instance of the problem of sparse data), the issue has stimulated much new computational research. For MEG and EEG measurements, relevant work has been done by Scherg et al., 1989, Dale and Sereno (1993), and Cuffin (1990, 1993, 1995).

In modeling the current source a number of parameters can be estimated? Assuming a single-dipole solution, one minimally determines the position (x, y, z coordinates), orientation (Qx, Qy, Qz), and strength (||Q||) of the dipole, for each millisecond. The factors Qx, Qy, and Qz correspond to the orientation of the dipole vector. Dipole strength ||Q|| is canonically expressed in units of Ampere (nanoAmperes). In the dipole modeling software used for these experiments, the program calculates, at each latency: dipole location (x, y, z), dipole orientation (Qx, Qy, Qz), dipole strength (||Q||), radius of the dipole location from the head center (r), root mean square value (indicates signal strength value across all sensors), correlation between the predicted and observed values, goodness of fit, and the number of iterations it took the algorithm to find the fit.

The dipole source localization algorithms are typically iterative and use least-squares minimization procedures. The iterative process begins with an initial guess about the position, orientation and strength of the dipole. The (well-established) forward equations are then used to compute the field generated by the first guess dipole. The computed field is compared to the observed field, an error term is calculated, and the iteration proceeds by changing the parameters until the error term is minimized. Several assumptions go into this type of modeling. The first assumption is, obviously, that the source is a dipole (a point source). Two important additional assumptions are (a) that the source is embedded in a sphere and (b) that the head is a spherical volume of uniform conductivity. The computational assumption that one can ignore field inhomogeneities is reasonable. Although neural tissue, cerebrospinal fluid, bone, and skin have different conductivities, they are well modeled by concentric spheres of distinct conductivities. However, if each sphere is distinct but of homogeneous conductivity, the magnetic field is insensitive to those differences. Consequently one can assume a single spherical volume conductor as a headshape. Independently of the conductivity assumption, the *geometric* assumption of sphericity is somewhat more problematic. At the very least, this assumption is a source of error for the localizations one must bear in mind. Nevertheless, insofar as one's main concern is localization and one is committed to using a single dipole source localization method, one can get reliable localizations within a resolution error of about 3mm (i.e. one can successfully resolve 2 dipoles that are 3mm apart in Euclidean space). Calculations by Cuffin (1993) suggest that this level of spatial resolution is maintained even under the assumption of a spherical head model.

It is commonly accepted that a single dipole is only an extremely crude approximation of the current source. Ideally one would like to model current *distributions* rather than idealized single point sources. Even multiple dipole models are, of course, no more than approximations to the real neuronal activity generating extracranial fields. A complex cognitive process (any cognitive process) is presumably not mediated by a group of neurons all contained within one cortical column and therefore conveniently captured by a single underlying dipole. It turns out, however, that for studies of the auditory system single dipoles are often extremely accurate descriptions of the data. For example, in the presentation of pure tones (Experiment 1 below), the single dipole model for the M100 response component (which is dipolar; see Figures 4.2 and 4.3) is extremely good. The correlation between the observed field and the field generated by a single dipole estimated through inverse-modeling is over 0.99; the goodness-of-fit is 0.99 and better; and the confidence volumes around the computed dipole are on the order of 0.01cm$^3$. One can now overlay the computed dipoles on MR images obtained from the same subjects. The overlay of MR images with dipole data from pure tone presentation is illustrated in Figure 4.4 for one subject.

FIGURE 4.4 HERE

When using single dipole localization (calibrated by pure tone stimuli) we are confident that we are recording *predominantly* from auditory cortex (on the superior temporal plane) for three reasons. First, the predominant source of extracranially measured evoked neuromagnetic fields is activity in apical dendrites of pyramidal cells. The apical dendrites of auditory cortical pyramidal cells are appropriately positioned within cortex to be registered (Orrison & Lewine, 1995). Second, it is established that pure tones drive primary auditory cortex effectively (Romani et al., 1982). Moreover, it has recently been shown by single-unit recording studies that bandpass filtered noise and complex stimuli such as species-specific vocalizations drive auditory association areas on the STG surface more effectively than pure tones (Rauschecker et al., 1995). Consequently, when dipole modeling attributes the source of a pure tone to supratemporal auditory cortex, the result is physiologically and functionally plausible. Finally, a substantial amount of modeling work, based on data obtained with different recording equipment and using different inverse algorithms, has converged on the finding that primary auditory activity underlies the M100 and the N1 component recorded in MEG and EEG/ERP (Elberling et al, 1982; Romani et al, 1982; Pantev et al., 1990; Yamamoto et al., 1988). For example, Scherg et al. (1989) modeled the sources (multiple equivalent current dipoles corresponding to N1, P2, the mismatch negativity MMN, and the sustained positivity SP) of the midlatency auditory evoked responses with EEG and found that the N1 is most reliably analyzed as having an underlying dipole source in koniocortex. Näätänen and Picton (1987) review the N1 and N1m/M100 component and its locus of generation and conclude on the basis of the cumulative literature that there are several overlapping sources underlying the N1 but that the two initial sources of the M100, including the component we are concerned with, are likely generated in primary auditory cortex. The precise locus of generation of the M100 remains unspecified. Recordings discussed in the next section reliably localized the M100 dipoles for tones, vowels, and syllables to non-primary auditory areas using MRI data.

## MEG RESULTS ON SPEECH SOUND PROCESSING

In recent years a number of investigators have recorded auditory evoked neuromagnetic fields elicited by speech sounds. The speech stimuli have included vowels (Kuriki & Murase, 1989; Hayashi et al., 1992; Eulitz et al., 1995), syllables (Kaukoranta et al., 1987; Kuriki & Murase, 1989; Sams et al., 1991; Hayashi et al., 1993), and words (Hari et al., 1989). Other studies have used pure tone processing to elucidate the functional organization of the auditory cortex (Elberling et al., 1982; Romani et al., 1982; Hari et al., 1984; Hari et al., 1987; Mäkelä et al., 1987; Yamamoto et al., 1988; Pantev et al., 1988, 1994; Kuriki & Murase, 1989; Cansino et al., 1994; Eulitz et al., 1995; Nakasato et al., 1995). The latter studies have been able to characterize in detail the components of the auditory evoked response and localize these components to regions of supratemporal auditory cortex. A number of distinct auditory evoked response components have been described. Following a suprathreshold auditory stimulus, one typically detects a response 40-50ms after stimulus onset, a prominent response around 100ms (the M100, or N1m), and a response of opposite polarity to the M100 at 150-200ms. In addition, some studies have documented a sustained field (SF) response that remains visible throughout the duration of a stimulus (Pantev et al., 1994; Eulitz et al., 1995) and a response to stimulus offset comparable to the M100 (Hari et al., 1987).

There is consensus that for auditory stimuli the M100 evoked response is particularly important (Näätänen & Picton, 1987). The M100 is virtually always elicited by auditory stimuli and is sensitive to acoustic stimulus features (Hari, 1991). The M100 exhibits specificity, as well: its amplitude and latency vary systematically with certain physical and temporal aspects of stimuli (Näätänen & Picton, 1987). Importantly, a number of investigators have been able to take advantage of the dipolar nature of the M100 to compute the single equivalent current dipole that underlies the evoked neuromagnetic field at that latency. Although acknowledged to be an idealization, single dipole modeling has generated a wealth of data which suggest that auditory cortex has at least one and perhaps multiple tonotopically (cochleotopically) organized fields (Elberling et al., 1982; Romani et al., 1982, Pantev et al., 1988; Yamamoto et al., 1988; Cansino et al., 1994).

Attempts to correlate response components with specific acoustic or linguistic-phonetic features of stimuli have generated more complex results. Kaukoranta et al. (1987), for example, detected a response component that was unique to certain of their syllabic stimuli. When presenting subjects with a fricative consonant-vowel (CV) syllable such as /hei/ or /ssseee/, they recorded a response that looked much like a second M100 (same polarity, same latency) that appeared to be triggered by the onset of the vowel. While the (standard) first M100 was presumably generated by stimulus onset, they argued that the second response, which they termed N200m or N1m', was generated by the onset of the vowel part of the syllable. The experiments revealed that the second response occurred with any vowel, did not occur when syllables were presented in the other direction (/eeesss/ versus /ssseee/)), and originated somewhat more anteriorly than the M100. Kaukoranta et al. concluded that the N1m' was specific to acoustic parameters, albeit ones that may reflect some important property of speech processing. Mäkelä et al. (1988) extended these findings by demonstrating that they could elicit the second peak N1m' with noise-square wave non-speech stimuli and manipulate its characteristics by temporal and amplitude changes of the stimuli.

Kuriki and Murase (1989) recorded the evoked magnetic field generated in response to pure tones (170Hz and 1kHz) and speech sounds (/a/ and /ka/) from both temporal lobes. Single dipole modeling for both tone and speech stimuli revealed that the sources in the right hemisphere (for the same stimuli, i.e. comparing vowel with vowel, tone with tone) were typically 1-2 cm more anterior than the comparable dipoles in the left hemisphere. This overall asymmetry for dipole location has been reported in other studies (Eulitz et al., 1995), and is presumably associated with the gross morphological asymmetry of the superior temporal plane in many human subjects (Geschwind & Levitsky, 1968). With respect to the speech stimuli, they found no left-right asymmetries when looking at the waveforms generated by the vowel or the syllable. However, dipole modeling showed that, in the left hemisphere only, the dipoles chosen as representative for syllables were significantly more posterior than the dipoles for vowels. These authors, too, argue that their results reflect the sensitivity of the M100 to acoustic features.

Eulitz et al. (1995) recently reported a systematic difference between tone and vowel stimuli, but not as early as the M100. They recorded the evoked neuromagnetic field generated by pure tones and vowels from both hemispheres and showed that there was a clear (localizable) effect of the sustained field (SF) that distinguished tone and speech stimuli. Although these investigators found no main effects of hemispheric asymmetries between tones and vowels at the M100, they demonstrated an amplitude increase in the SF to vowels as compared to the SF to tones. The SF increase was particularly strong in the left hemisphere. Moreover, the SF had a dipolar distribution which they were able to localize to sites slightly more anterior than the M100. They attributed the increased left temporal SF response to speech-specific processing.

Generally, in analyses of the latencies and amplitudes of the major evoked response components most studies have failed to detect any significant asymmetries between the left and right temporal cortices before the SF response approximately 200ms after stimulus onset, at least with repect to vowels (Eulitz et al., 1995) and words (Hari et al., 1989). Moreover, in a review of the EEG literature on asymmetries in speech and nonspeech processing, Rugg (1983) maintains that the results are contradictory, and some investigators have failed to detect any asymmetric responses with speech stimuli such as vowels when explicitly looking for them (Molfese & Irwin, 1981).

The apparent symmetry of earlier (<100ms) evoked responses to speech is surprising if one subscribes to a model in which the analysis of speech sounds is performed exclusively by left temporal cortical mechanisms. In particular, because speech sound recognition is an extraordinarily fast process (*cortical* responses elicited by auditory stimuli can be measured as early as 15ms after stimulus onset in direct recordings; Celesia, 1976), one might reason that responses as late as 100ms should reflect higher-order linguistic processing, and because there is ample evidence for left-hemisphere specialization of many language processes one might expect that type of stimulus to preferentially engage the left temporal cortex. On the other hand, these data are compatible with a model in which speech sounds are processed in both superior temporal cortices.

In addition to the MEG and EEG studies that have failed to show systematic left-right asymmetries attributable to speech processing, other electrophysiological research with human subjects has yielded similar findings of symmetrical processing. For example, Creutzfeldt et al. (1989a, 1989b) and Ojemann (1991) report on intracranial recordings performed during surgery during which they presented patients with a variety of speech stimuli and recorded from the left and right superior and middle temporal gyri. Both single- and multi-unit recordings in superior temporal gyrus did not show any obvious left-right differences, although the responses appeared to be differentially modulated by task requirements. In a similar vein, Boatman & Gordon (p.c.) have monitored patients' speech discrimination performance while recording from the temporal cortex using subdural grid electrodes. These investigators have found the same type of effects when recording and stimulating in the left and right posterior superior temporal gyri (STG). Recordings revealed a significant activation increase in posterior STG during speech perception while stimulation of a posterior STG electrode pair dramatically impaired phoneme discrimination across subjects. Finally, from a 'bottom-up' perspective, electrophysiological recordings in the auditory nerve and the brainstem nuclei in macaques and other species have not revealed any obvious morphological or functional asymmetries that one could link in an direct way to a left temporal cortex preference for speech sound processing (Clarey et al., 1992).

Data from human brain research methods with lower temporal resolution speak to this issue as well. First, the evidence from pure word deafness, presented in Chapter 3, is consistent with the claim that both the left and right superior temporal gyri play an essential role in speech sound processing. Second, a survey of the original (FDG PET and 133-Xe) and more recent (PET and fMRI) functional neuroimaging literature reveals that speech perception more often than not activates both superior temporal gyri, regardless of whether the stimuli were single phonemes, syllables, words, or connected speech (Chapter 3). In summary, although there exists ample evidence that links many language processing subroutines to a number of cortical areas in the left hemisphere, including left posterior STG, the neurobiological evidence that speaks to the substrate for speech processing (construed as a subroutine of language comprehension) more strongly supports the hypothesis that speech perception is subserved by the posterior superior temporal gyrus in both hemispheres.

If the decoding and interpretation of speech sounds is in fact mediated by both supratemporal cortices, a number of new issues are raised. First, one now owes an explanation addressing at what processing level the extensively documented asymmetries for language processing arise. (This issue invites speculations such as central modality-independent representations are mediated by the left hemisphere. But see Baynes, 1990, and Gazzaniga, 1983, for reviews that argue for richer language representational capacities of the right hemisphere based largely on split-brain evidence). A second issue, and one that is more directly experimentally tractable, is whether or not the left and right auditory cortices compute speech input in a similar manner.

To shed light on this latter issue, we recorded neuromagnetic data from six subjects from the left and right superior temporal lobes independently while the subjects listened to tone (Experiment 1), vowel (Experiment 2), or syllable (Experiments 3a and 3b) stimuli. Because of the importance of the M100 for the assessment of auditory stimuli, we restricted the present analysis to M100 latency and amplitude and to estimates of the single dipole that contributes to the generation of the speech-elicited magnetic field around the M100.

In addition to the questions of lateralization and localization of function in the verbal domain, independent observations coming from the neuroimaging literature motivated the exploration of the interaction of speech stimuli with attentional requirements, such as those one will engage by having to execute particular linguistic task demands (Experiments 3a and 3b). There is a large literature devoted to the study of attention both with EEG/ERP and, more recently, with MEG. The studies investigating attentional effects have been quite distinct from the studies presented in this chapter. Two types of experimental protocols are common, tone oddball paradigms, in which one measures the mismatch response to an oddball stimulus in a sequence of standards (Rif et al., 1991), and 'alternating ear' or dichotic listening paradigms, in which one is asked to selectively attend to the input delivered to the left or right ears (Rif et al., 1991; Woldorff et al., 1993). A central finding of these studies has been that attending to the input of one or the other ear modulates the response amplitude of the M100. For example, Woldorff et al. (1993) studied the dichotic presentation of pure tones and reported that the same tone elicited larger M100 amplitudes when attended than unattended.

The study presented in this dissertation (Experiments 3a and 3b) addresses a simpler problem. Subjects hear the same stimulus set (monaurally presented to the contralateral ear) in two conditions. In one condition (Experiment 3a), they 'passively' listen to syllables. In the attended condition (Experiment 3b), they listen to the same stimulus set but they are asked to make an overt discrimination on the stimuli. I hypothesize that selective attention to the stimuli will modulate the response properties. Specifically, based on some findings in the neuroimaging literature (Chapter 3) I hypothesize that the attentionally conditioned response modulation will be most evident in amplitude increases in the left hemisphere recordings.

A final concern is what kind of functional topographies might exist in the auditory system. Two possible outcomes are anticipated. Under one model, speech sounds are 'assembled' in tonotopically organized auditory cortex. These tonotopic areas do not care about whether a stimulus is linguistic or not but are wired such that a phonetic stimulus generates a representation which can interface with linguistic systems in an effective manner. An alternative possibility is that there exist (one or more) topographic maps specifically organized for the representation of speech stimuli. Insofar as there is topological systematicity in the representation of speech sounds, a detectable functional map of phonemes, or 'phonotopy', we obviously want to understand how it functions. The question is explored in a preliminary fashion in the analysis of the vowel data (Experiment 2).

Recapitulating, the genesis of the hypotheses, we reasoned as follows. There is a tension between much previous research on asymmetries in language processing and many contemporary findings coming from MEG, PET, and fMRI. By explicitly separating speech perceptual processes from language processing and by investigating speech as a subroutine of language comprehension we aim to elucidate the neural basis of speech perception. Because we have a good understanding of auditory evoked responses, particularly of the M100, the experiments investigate the contribution of left and right auditory cortices to to the processing of speech stimuli and the interaction with attentional mechanisms as reflected by properties of the M100 auditory evoked neuromagnetic response.

Two major results derived from these data are discussed. The analysis of the M100 evoked field suggests that both superior temporal cortices contribute to the construction of speech sounds. In particular, the data suggest that left and right superior temporal cortices are not just performing spectral (acoustic) analysis on the stimuli but generate more articulated responses. Second, the comparison of passive and task-dependent syllable perception shows that the superior temporal cortices respond differentially to task demands. Passively listening to syllables showed complex responses in left and right superior temporal cortices. In contrast, the execution of a simple categorization task on the same stimuli revealed that the same region in left temporal cortex was up-regulated (large M100 amplitude increases) while the the same region in right temporal cortex was down-regulated (large M100 amplitude decreases).

## GENERAL METHODS

### Subjects

Six normal-hearing healthy volunteers (three women and three men, ranging from 23 to 36 years, mean age = 29.6 years) participated in the experiments. Each of the six participants ran in each of the four experimental conditions. The volunteers gave informed consent, separately for the MEG and the MRI components of the study. The experimental protocol was approved by the University of California at San Francisco Medical School Committee on Human Research and the Massachusetts Institute of Technology Committee on Human

Subjects. Information about the subjects is summarized in Table 4.2. None of the subjects reported any hearing loss or other audiological abnormalities. Note that subject IS is a 31-year-old male whose native first language (L1) is German, not English, like the other five subjects. He is also a left-hander, with no first-degree left-handed relatives. Because the syllables used in these studies contrast phonetically in German and are also nonwords in German, this subject's data were considered comparable to the data generated by the other subjects. Another subject (JM) is a strong left-hander, by self-report. She also reports having no first-degree left-handed relatives. One subject (RR) has a blind spot in the fovea of the right eye, due to retinal (not cortical) damage. Because handedness interacts in complex ways with hemispheric asymmetry as assessed by different measures for language and other cognitive domains, the data for the two left-handed subjects are explicitly marked in the graphs and the statistical analyses are done both including and excluding those two subjects.

**TABLE 4.2** *Subjects*

| Subject | Date of birth | Sex | Handedness | L1 |
|---|---|---|---|---|
| Subj. 1: JM | 3/12/68 | female | left | English |
| Subj. 2: MM | 9/9/70 | female | right | English |
| Subj. 3: RR | 12/5/71 | female | right | English |
| Subj. 4: AM | 1/31/59 | male | right | English |
| Subj. 5: JH | 8/22/62 | male | right | English |
| Subj. 6: IS | 10/10/63 | male | left | German |

### Magnetic field measurements

The recordings were done in a magnetically shielded room using a 37-channel magnetometer with SQUID-based first-order gradiometer sensors (Magnes, Biomagnetic Technologies, Inc., San Diego, CA). The SQUID sensors, arranged in a hexagonal grid of 14.4 cm in diameter, were supported inside a liquid-helium containing dewar. The dewar was placed over the left or right temporal lobe and its position was adjusted to record optimally from auditory cortex (see below). The position of a participant's head with repect to the sensors was recorded by a transceiver-based system which localized fiducial landmarks (see below), thereby also setting up a spatial reference frame for the MEG data.

The stimulus durations were 400ms (tones) and 300ms (vowels and syllables). We therefore acquired 600ms epochs which included a 100ms pre-stimulus interval and a 100ms or 200ms post-stimulus interval. The start of acquisition of each epoch was triggered by a TTL-level trigger sent from the Macintosh Quadra to a dedicated data acquisition processor controlling recording of the neuromagnetic fields. We recorded neuromagnetic data using a 1.0Hz high-pass cutoff. In the tone and syllable conditions the sampling rate was 1041.7Hz with a bandwidth of 400Hz. In the vowel condition, we acquired epochs with a sampling rate of 2083.3Hz (a bandwidth of 800Hz). We acquired data at such a high sampling rate in the vowel condition because we were interested in any responses in higher frequency bands that might have systematic relationships to the formants of vowels. All epochs were digitized for later processing.

### Magnetic resonance images

For each subject, structural MRIs in the three standard planes (axial, coronal, and sagittal) were acquired with a 1.5T magnetic resonance scanner at the UCSF MRC facility. Examples are shown in Figure 4.4. For all three views the resolution (pixel size) was 256mm x 256mm. Coronal images were aquired with repetition time TR: 5000 and echo time TE: 96ms. Axial images were acquired with TR: 36 and TE: 8. Sagittal images were acquired with TR: 100 and TE: 7. The MR images were judged to be normal by a neuroradiologist.

## Procedure

For each individual subject, the experimental protocol entailed the following steps. Before entering the magnetically shielded room, each subject was demagnetized (using a standard handheld demagnetizer ) to eliminate any magnetization in a participant's body (such as metal dental fillings, magnetized blood from having been in an MRI scanner, etc.) that may affect the signal measured in the magnetometer and therefore significantly contaminate the evoked fields one is attempting to measure.

The first measurement for a given subject was a noise test. With a subject lying down on his or her back, the magnetometer was placed over the head with the concave recording surface of the dewar approximately 1-2 inches above the frontal lobe and centered over the frontal lobe. This initial noise measurement insured (a) that the individual sensors - SQUIDS - were properly functioning, and (b) that we were able to obtain adequate recordings from the subject. When the noise was large, subjects were demagnetized again and the procedure was repeated. There are subjects who are not appropriate for MEG recording because some materials in their bodies (such as metallic dental fillings) generate large artifacts that do not average out sufficiently to reach an adequate signal-to-noise ratio.

Following the noise test, subjects were given a brief pretest (20 trials) of the CV discrimination condition (BPDT discrimination condition; experiment 3b) to allow for some practice and familiarization with the test.

With the subject lying on his or her back, the headshape was digitized and fiducial points were marked. The headshape digitization was done using a digitizing stylus with which we outlined the entire headshape, thereby sampling several thousand points from the skull (typically 3000 points). The sphere with respect to which the localization algorithm computes dipole positions is generated based on these digitized headshape data. That is, from the headshape digitization data a sphere is generated that best fits the sampled cranium. When the dipole localization algorithm is run (for any data acquired from that subject), the sphere that enters into the localization computation is the sphere most appropriate to a given subject's head (see below).

Five fiducial points were also marked: the two preauricular points, nasion, and the inion and Cz positions (left and right forehead, respectively, unlike their 'classical' positions in the center and back of the head along the midline). The fiducial points were used (a) as reference points with respect to which the sensors were lined up, in order to locate the head with respect to the sensor positions, and (b) to line up the derived MEG localizations with independently acquired structural MRIs obtained from each subject. The preauricular points and nasion are reliably identifiable on structural MRIs. Consequently one can carefully line up the electromagnetically measured data with the structural data for each subject.

The subjects were now positioned the way they remained for the duration of the experiment. During the recordings, the subjects lay on their left or right side, depending on the side of recording. They were made as comfortable as possible - typically by positioning pillows around them - because the duration of recording lasted two to three hours for each hemisphere.

The stimuli were presented to the contralateral ear (contralateral auditory presentation typically yields larger response amplitudes; Näätänen, 1992; Nakasato et al., 1995). A small but tightly fitting earphone was inserted into the stimulated ear (right ear in the case of left hemisphere recording, left ear in the case of right hemisphere recording). Although it was extremely quiet in the magnetically shielded room, the other ear was plugged using

30dB sound-attenuating earplugs to eliminate possible distracting noise from the ventilation system or any other ambient sound.

The first study for each subject was a pure tone presentation (Experiment 1). This was done (a) to achieve an optimal positioning of the dewar over auditory cortex by computing the single equivalent current dipole for the responses to the pure tone, and (b) to acquire pure tone data that speak to the issue of the relative lateralization of the responses. That pure tones generate strong auditory cortex responses is well established (Elberling et al., 1982; Romani et al., 1982; Näätänen, 1992).

Following the pure tone run (Experiment 1), subjects went on to one of the three speech-stimulus conditions: the vowel study (Experiment 2), passive syllable (BPDT) study (Experiment 3a), or syllable discrimination (BPDT_ID) study (Experiment 3b). The order of these studies was varied between subjects, and was varied within subjects between hemispheres. Detailed descriptions of the conditions are provided below.

## Data analysis

**1. Averaging and Filtering**. For each hemisphere, the 128 epochs (in the case of tones) or 100 epochs (vowels and CVs) per stimulus were averaged for further processing. The averaged waveforms were digitally filtered from 1-20Hz, using a low-pass filter of filter-order 4 (20dB attenuation at 25Hz), and all further analyses were performed on the filtered data.

**2. M100/N1m RMS peak and latency**. The root mean square (RMS) of the field strength across the 37 channels was calculated for each datapoint. The peak RMS of the M100 (N1m) and its latency were chosen to lie in a latency range from approximately 60 to approximately 160ms. The rationale for choosing the amplitude and latency of the RMS peak as a measure to be interpreted was primarily that RMS is the most direct ('least' derived) measure insofar as it reflects the field value at the sensors. RMS is also used in other studies to summarize the M100 evoked component and is therefore appropriate for comparison across studies. RMS is computed as $RMS = [(\Sigma_{i=1} m_i^2)/N]^{1/2}$, where $m_i$ is the measured field at a given sensor and N is the number of (active) sensors. Other available measures include the peak correlation between the single-dipole predicted data and the observed measurements, the peak dipole moment, goodness of fit, and number of iterations to find a fit.

**3. Localization**. To estimate the source parameters of the M100, we used a single equivalent current dipole model. The localization program inversely (iterative least-squares minimization) finds a single dipole in a spherical volume of uniform conductivity. Parameters of interest were the position, orientation, and strength of the computed dipole and the confidence limits for that localization in the head-centered coordinate system. Representative *single* dipoles were chosen in the following way: the M100 peak was determined (latency of peak amplitude), and then all the dipoles in a 10ms window around the M100 peak were computed (peak latency +/- 5ms). Of all the dipoles in that interval the dipole with the best (smallest) confidence volume was chosen as representative for that condition.

**4. MR overlay**. The left and right preauricular points and nasion were marked during the MEG recordings, and set up a coordinate system with respect to which one could interpret the MEG data. By superimposing these fiducial landmarks on the MR image of the appropriate subject, one can visualize the position of the computed point sources with reasonable anatomical accuracy (error on the order of few millimeters). For the MRI dipole overlays, we chose larger intervals (e.g. M100 RMS peak +/- 15ms) and plotted all the dipoles in that interval.

**5. Statistical analysis**. The M100 RMS data were analyzed using standard analysis of variance (ANOVA) methods. In the report of ANOVA results, an effect size estimate, *eta*, is provided for each factor analyzed. Eta, which can take on a value between 0 and 1, is computed as $(\text{SumSquares}_{\text{effect}}/\text{SS}_{\text{effect}} + \text{SS}_{\text{error term}})^{1/2}$. Analogous to the effect size estimate r (correlation), eta assesses the relationship among factors. However, unlike r, eta is more general in that it assesses relationships that go beyond the degree to which there is a linear relation between variables. $\text{Eta}^2$, like $r^2$, provides an estimate of the coefficient of determination, an estimate of the amount of variance accounted for by a given independent variable.

## EXPERIMENT 1: TONES

### Introduction

The presentation of pure tones constitutes one of the simplest experiments to evaluate processing in auditory cortex, only clicks being acoustically simpler stimuli. Single-unit recordings in macaques (for review, see Clarey et al., 1992) as well as EEG/ERP and MEG recordings in humans have established that pure tones are effective stimuli to engage processing in auditory cortex (Näätänen & Picton, 1987; Näätänen, 1992). To what extent these stimuli more effectively drive cells in primary or association auditory cortex is an open question. Interestingly, pure tones have been argued to drive primary auditory cortex better than bandpass noise with a center frequency equivalent to the pure tone, whereas bandpass noise has been argued to drive neurons on the lateral surface of the posterior superior temporal gyrus better than pure tones (Rauschecker et al., 1995).

We presented subjects monaurally (contralateral to the recorded hemisphere) with 500 Hz pure tones. There are a number of factors we wanted to clarify in this condition. (1) Are the evoked neuromagnetic fields comparable to the data in the published literature? Specifically, are the M100 amplitude and latency approximately equal to the measurements obtained in similar studies? (2) Are the response amplitudes we are measuring from the two hemispheres in the speech sound studies selectively biased in the direction of one hemisphere? Are all left (or right) hemisphere response amplitudes larger, for whatever reasons (systematic or artifactual), or is there variability? (3) Are we localizing the neuronal activity underlying pure tone perception to the same area (auditory cortex) as previous studies - in other words, are our results comparable to the results in the literature in that the localization algorithm converges with others? Clear answers to these issues are critical for the interpretation of the speech perception studies that follow. Insofar as the values we obtain in the pure tone condition are similar to other researchers' findings, we can interpret our new results on speech perception with considerably more confidence. In that sense, the pure tone condition also serves as an important baseline control.

## Materials and Methods

The tones were generated by a Wavetek™ function generator. We used a 500Hz sinusoidal pure tone of 400ms duration (in 1 subject, S5: JH, we also presented a 1500Hz and 5000Hz tone in the right hemisphere recordings, and in another, S2: RR, we presented a 1500 Hz tone to the left instead of a 500 Hz tone). The stimuli were amplified and delivered to the subjects via plastic tubing. 128 repetitions of the tone were given in one block. The inter-stimulus interval was 1 second.

Pilot experiments had established that the SQUIDS are so sensitive that they recorded the stimulus at the amplifier when the amplifier was in the magnetically shielded room and itself not shielded. The cortical responses were thus contaminated by the 'response' of the amplifier. The stimulus amplifier was therefore outside of the recording room. This required using long plastic tubing to deliver the stimuli to the subjects. Naturally, using long tubing has its own problems for sound delivery. In particular, the higher frequencies of speech sounds (e.g. the contributions of the 2nd and 3rd formant) are attenuated. Some subjects did report some loss of clarity in the upper frequency range, but all stimuli were clearly audible and identifiable.

The amplitude of the sounds was adjusted manually so as to be comfortable for the subjects. Once a subject selected a loud and clear but comfortable amplitude, that value was maintained throughout the run of recordings collected from one hemisphere. To minimize eye-movement related artifacts, subjects' eyes remained closed during data acquisition.

The evoked field was recorded in 600ms epochs at a sampling rate of 1041.7Hz (bandwidth 400Hz), averaged, and low-pass filtered (1-20Hz low-pass filter). Contour maps of the evoked field were then computed for the strong evoked responses, notably the M100. Insofar as the sensors were positioned over the auditory cortex in a way that effectively recorded fields generated there, the contour maps showed the field exiting the cranium on one side and reentering the cranium on the other side. In other words, we aimed for a positioning in which the zero line of the contour map was in the center of the sensor array. If the contour map information or sensor layout information indicated that the dewar position was not optimal, the sensor array was repositioned over the subjects' head and the fiducial points were registered again. This procedure was repeated until the sensor array was in a position that best sampled from the auditory cortex.

It is possible that the region of cochleotopically organized auditory cortex that responds to a particular frequency is in an unusual orientation (as was the case for one subject, S2: RR; see discussion below). If there is morphological variation within auditory cortex such that some region is not in an orientation from which one can record effectively, it is necessary to use tones of different frequencies - that are therefore represented in a different place in auditory cortex. For example, if a 500Hz tone fails to produce the expected responses, a 1500Hz tone may produce them if that frequency is represented 'far enough away' from 500Hz. In the best of all possible worlds (Leibniz, 1714; Voltaire, 1759) one can then present several different frequencies to each subject in order to establish a tonotopic mapping, should it be evident in that particular subject. The experimental protocol of this study was so long that only one frequency, 500Hz, was used to evaluate pure tone processing.

## Results

### Evoked magnetic field results

The M100 RMS peak and the latency of the RMS peak for the left and right hemisphere measurements for each subject are summarized in Table 4.3. The latencies of the M100 ranged from 111.4 ms to 138.2 ms. For subject three (RR) a 1500Hz tone was used in the left-hemisphere/right-ear recording instead of a 500Hz tone. Unlike the measurements from the right temporal lobe, it was not possible to obtain adequate signals from this subject using a 500Hz tone. The subject had no difficulty perceiving the 500Hz tone, so a microlesion selective for that frequency-coding region of cortex was considered unlikely. Closer inspection of the subject's MRI revealed considerable 'roughness' within the Sylvian fissure at the presumed depth of auditory cortex. It is possible that in that subject frequencies around 500Hz are represented in one of the small gyri within the Sylvian fissure. Insofar as the pyramidal cells in these microgyri are not in an orientation visible to magnetic measurement (i.e. the apical dendrites of these pyramidal cells are not tangential to the scalp surface), one cannot effectively measure their activity. Repeating the protocol with a 1500Hz tone yielded a good localization in primary auditory cortex, suggesting that the the failure to detect reliable respones to 500Hz was in fact a consequence of that frequency being represented at a place not adequate for neuromagnetic measurement.

TABLE 4.3 HERE

Figures 4.5 and 4.6 show the peak RMS results and the latency of peak RMS results, respectively. Figure 4.5 illustrates three features of these data. (1) There was large variability in response strength to pure tones across subjects, presumably due to factors such as head size, skull thickness, relative orientation of auditory cortex, the position of the sensor array, and so on. In the group of six subjects studied, field strength values ranged from 64.857 fT and 67.232 fT (subject 3; RR) to 167.212 fT and 144.776 fT (subject 5; JH) for the left and right hemispheres, respectively. (2) There was no systematic relationship between field strength and hemisphere for this group of subjects. For three of the subjects the responses were stronger on the right, for the three other subjects the responses were stronger on the left.

FIGURES 4.5 AND 4.6 HERE

The analysis of variance summarized in Table 4.4a confirms that with respect to peak RMS there is no significant main effect of hemisphere, F(1, 4)=0.245, p=0.647, eta=0.44. (3) Note that in Figures 4.5 and 4.6, subjects S1-S3 are female and S4-S6 are male. Although it was not a variable of a priori interest in these experiments, there is a significant main effect of gender, F(1, 4)=23.30, p=0.008, eta=0.97 (see Table 4.4a), with the peak RMS values being significantly larger in the recordings from male subjects. Table 4.4ab shows that the Hemisphere x Gender interaction was not significant.

TABLES 4.4a & b HERE

Because of the two left-handed subjects and the often distinct data patterns observed to be generated by them, the data were recoded two ways and subjected to ANOVAs. One reanalysis involved recoding all subjects by presumed dominance, which in effect meant recoding subject 6 (IS). This reanalysis of the M100 RMS and latency data revealed the same patterns of results. For RMS, there was a main effect of gender, F(1, 4)=23.30, p=0.008, no main effect of Hemisphere, F(1, 4)=3.76, p=0.124, and no significant interaction, F(1, 4)=0.093, p=0.776. The latency data revealed no main effects or interaction. In a second reanalysis, the two left-handed subjects were omitted. The main effect of gender was still large, F(1, 2)=46.61, p=0.021, and there was no effect of Hemisphere, F(1, 4)=5.98, p=0.134, and no interaction. The latency data also revealed no significant effects.

Figure 4.6, displaying the latencies of the peak RMS values for each subject and each hemisphere, shows no obvious differential effects of hemisphere or gender of subject on the latency of the RMS peak. The analysis of variance, summarized in Table 4.4b, confirms that there were no significant or marginally significant effects. Figure 4.5 shows that the two left-handed subjects had the largest asymmetric response amplitudes, although in opposite directions. Subject 6 gave the only datapoint that showed a definite trend in the non-preferred direction, with his right side response amplitude being much larger than in the left. ANOVAs on the RMS data with the two left handers excluded revealed no effects different from those reported in Table 4.4a (see above).

### Dipole localization results

Table 4.5 summarizes the parameters of the single dipole that was chosen as representative for the condition. The single dipole chosen was that dipole within a 10 ms interval around the latency of the RMS peak that had the smallest confidence volume (see Methods). The goodness-of-fit and correlation values indicate how good a model for the data the single dipole was. The values for both measures range from 0.96 to 0.99 and suggest that the single dipole model is a reasonable estimate for the field generated by the presentation of a pure tone. Nevertheless, since it is not uncommon in neuromagnetic recordings elicited by pure tones to observe correlations consistently above 0.99, these values are relatively low. There is no account for why the correlations and fits were less accurate than is optimal.

The 95% confidence volumes for these dipole localizations were on the order of 1-10 cubic millimeters, which can be considered to be remarkably good localizations.

TABLE 4.5 HERE

Figure 4.7 plots, in a head-centered coordinate frame, the x (anterior-posterior), y (medio-lateral), and z (superior-inferior) coordinates of the single dipole in the left and right hemisphere for the six subjects. The grouping of the single dipoles in the right hemisphere shows little variation given that these are the dipole values for six different subjects. In the left hemisphere, there is slightly increased variation, particularly in the anterior-posterior dimension. The one dipole that stands out in the left hemisphere because it is considerably more medial is the dipole for Subject 3 (RR). As discussed above, for this subject we had to use a 1500Hz tone when evaluating the response from the left temporal cortex. Note that the more medial location of the dipole for a higher-frequency tone is consistent with reports in the literature that higher frequencies are represented more medially (Romani et al., 1982; Pantev et al., 1988).

FIGURE 4.7 HERE

For three subjects we examined the location of the single 'best' dipole to the tone with repect to their neuroanatomical MR data (see Figures 4.4 and 4.34). For all three subjects (S1: JM, S2: MM, S5: JH), the dipoles were judged by a neurologist and experienced interpreter of MR images (Dr. Albert Galaburda, Beth Israel Hospital, Boston, MA) to be outside of primary auditory cortex. For subject S5: JH, the dipole in the left and right hemispheres was estimated to be on the planum temporale in 2nd order auditory association cortex, perhaps area Tpt. For S1: JM both left and right dipoles were not in primary auditory cortex but noticably posterior to it. For subject S3: MM the dipole in the left was in Heschl's gyrus but outside of primary auditory cortex; in the right hemisphere the dipole was at the edge of the superior temporal sulcus although probably still within auditory association cortex, and not polymodal association cortex. Importantly, the these localization estimates take into account the possible localization error of the method.

## Discussion

With respect to the three leading questions we intended to clarify, we found the following. (1) The strength and latency of the M100 auditory evoked field that we recorded are comparable to other studies. (2) We found no significant or marginally significant effect of hemisphere. The absence of any lateralization effects is important, particularly if we should find such effects in other conditions using speech stimuli. Insofar as we detect lateralization effects for speech stimuli, they cannot be interpreted as effects generated by any random auditory stimulus.

(3) Our single dipole localizations successfully placed the source of the tone-elicited M100 in auditory cortex. However, unlike what is often reported, we did not localize the single dipoles to primary auditory cortex, in either the left or right hemisphere. As determined for three of our subjects by overlaying the MEG data on structural MR images, the single equivalent current dipoles were in auditory association cortex. Although we could not ascertain whether the dipoles localized to first-order or second-order auditory association cortex (possibly as distal to primary auditory cortex as the superior temporal sulcus, which contains polymodal association areas), it was clear that the dipoles did not localize to primary projection auditory cortex. Näätänen & Picton (1987) argue that the N1/N1m auditory evoked response is probably generated by up to six different cerebral processes that contribute to the extracranially recorded N1/N1m. The earliest two of these subprocesses are hypothesized to localize to distinct parts of supratemporal auditory cortex.

It appears that the component we are recording is not generated by primary auditory cortex but by an auditory association area.

Lesion evidence by Woods et al. (1987) suggests that N1/N1m is probably generated over a wider area than one might assume. Heschl's gyrus/primary auditory cortex lesions did not abolish the response; only bilateral lesions extending posteriorly to temporoparietal cortex abolished the evoked N1. In addition, considerations about the timing of responses in auditory cortex suggest that it is plausible - perhaps even more likely - that the M100 is generated outside of primary auditory cortex. Signals are transmitted to primary auditory cortex within approximately 15 ms (Celesia, 1976). Although that does not mean that the M100 cannot be generated in A1 (it could be a highly processed, top-down modulated response in A1), the data are also consistent with the assumption that one is measuring neuronal activity of a downstream projection area. In any case these data suggest that one must be rather cautious when attributing neural activity as summarized by a single dipole to an area in auditory cortex.

There was an effect of gender on response amplitude of the M100. The response strength difference between males and females is somewhat paradoxical. One might expect that (i) because one average males have larger heads and (ii) because magnetic field strength falls off with the square of the distance, male brains should generate smaller magnetic fields (assuming that an approximately equal number of neurons are active in male and female brains when a tone is presented). In light of the findings reported for the other studies, in which such sex-dependent amplitude effects have not been reported, we have no interpretation for the result, and so it must remain, for the moment, a curiosity.

# EXPERIMENT 2: VOWELS

## Introduction

The rationale for this study was to use an extremely simple experimental paradigm, comparable to passive tone presentation but with the added dimension of using speech stimuli. What sets vowels apart from tones is their complex internal structure: the physical features that give vowels their identity are the bands of energy peaks at particular frequencies, the so-called formants. Although the acoustic structure of vowels resembles a combination of tones, vowels are not perceived as tone-combinations (chords). Rather, vowels have unique perceptual characteristics defined by various articulatory factors that form the basis for their production (height, frontness, and rounding, for example).

With regard to the lateralization of responses for vowels the findings are contradictory. The classic dichotic listening studies by Shankweiler & Studdert-Kennedy (1970) and many others showed that vowels had a relatively smaller lateralized effect than consonants in dichotic listening experiments. Unlike consonants, which are defined by rapid (<30ms) formant transitions, vowels, despite being speech stimuli, were supposed to be processed more or less symmetrically in both hemispheres because they did not carry the rapid-transition temporal information thought to be charateristic of speech. However, direct recordings, e.g. by Lund et al. (1986), showed a marked left superior temporal cortex preference for vowels as well as consonants.

Several groups have investigated the neuromagnetic field evoked by vowel perception in recent years using MEG, including Kuriki & Murase (1989), Hayashi (1992) and Eulitz et al. (1995). Kuriki & Murase (1989) and Eulitz et al. (1995) reported findings that differentiated vowels from tones. Kuriki & Murase localized the M100 equivalent current dipole generated by vowels consistently more anterior to the dipole generated by tones in the left hemisphere. Eulitz et al. (1995) found that vowels generated a significantly larger sustained field (SF) responses over the left temporal cortex than tones. This sustained field was dipolar, and they were able to localize it to a cortical field more posterior than the M100 dipoles.

In order to further evaluate whether there were any hemisphere-specific effects in the responses to vowels, we presented 3 vowels (/a i u/) at two fundamental frequencies (100Hz and 200Hz) to our six subjects while recording, sequentially, the evoked neuromagnetic field from the left and right temporal cortices. Because pure tones (i) revealed no hemispheric effects with respect to M100 amplitude and latency, and (ii) localized to non-primary auditory areas, we expected to find some lateralization effects with the speech stimuli, either in latency, amplitude, or localization.

## Materials and Methods

### Stimuli

The vowel (and CV) stimuli were created using a Klatt formant synthesizer (SenSyn™, Sensimetrics Inc., Cambridge, MA) running on a Macintosh Quadra 800 microcomputer. The vowels were synthesized with the goal to be as natural-sounding as possible. Spectrograms, waveforms, and power spectra of two representative vowels are presented in Figures 4.8 ('male' /a/; $F_0$=100Hz) and 4.9 ('female' /a/; $F_0$=200Hz). The waveforms, spectra, and spectrograms were generated using the Signalyze™ software package, InfoSignal, Inc., Lausanne, Switzerland. The source parameters *amplitude of voicing* and *amplitude of aspiration* and the frequency time course of the *fundamental frequency* $F_0$ were as follows. The duration for all vowels was 300ms. The fundamental frequency went linearly from 120Hz to 90Hz over 300ms. The amplitude of voicing ramped up from 20 to 60 dB in the first 50ms, then slowly declined from 60 to 55dB till 200ms, declined more sharply from 200 to 250ms (55 to 45 dB), and went down to zero dB between 250 and 280ms. The amplitude of aspiration ramped up from 0 to 5dB in the first 30ms, remained constant till 290ms, and sharply declined to 0dB.

FIGURES 4.8 AND 4.9 HERE

Table 4.6 lists the formant values and bandwidths used for the six vowels presented in this experiment.

**TABLE 4.6** *Vowel formants: spectral composition (in Hz)*

| | F0 | F1 | bandwidth | F2 | bandwidth | F3 | bandwidth |
|---|---|---|---|---|---|---|---|
| a(m) | 100 | 710 | 90 | 1100 | 100 | 2540 | 90 |
| a(f) | 200 | 850 | 90 | 1220 | 100 | 2810 | 90 |
| i(m) | 100 | 280 | 90 | 2250 | 100 | 2890 | 90 |
| i(f) | 200 | 310 | 90 | 2790 | 100 | 3310 | 90 |
| u(m) | 100 | 310 | 90 | 870 | 100 | 2250 | 90 |
| u(f) | 200 | 370 | 90 | 950 | 100 | 2670 | 90 |

### Stimulus presentation

The stimuli (the vowels as well as the CVs used in studies 3a and 3b) were loaded into an experimental psychology software package, PsyScope, developed at Carnegie Mellon University (Cohen et al., 1993). Through the interface with a custom-designed output device (MITBox, Elron Yellin, 1994), the PsyScope software, running on a Macintosh Quadra 800 microcomputer equipped with a 16-bit sound card (Audiomedia II™), delivered the stimulus to the subject, and delivered the stimulus identification code and a TTL-level epoch-acquisition trigger to the data acquisition processor that registered the magnetometer output. Each vowel was presented 100 times to allow signal averaging. The stimuli were pseudorandomly intermixed. Thus, in the vowel condition, a subject heard six distinct stimuli, /a/, /i/, and /u/, once with a female pitch (200Hz) and once with a male pitch (100Hz), each one 100 times, in pseudorandom order. The ISI varied between 750ms and 1250ms, to avoid habituation and artifacts of periodic origin.

### Recording

In the vowel condition we acquired data with a sampling rate of 2083.3 Hz (bandpass of 800Hz) to evaluate whether there were any effects of the lower formants of the vowels on the neuromagnetic signal when analyzed for power in spectral components.

## Results

Evoked magnetic field results

The M100 RMS peaks and their latencies for each subject and each vowel are listed in Table 4.7. Figures 4.10 - 4.13 display these results graphically, and Tables 4.8 and 4.9 summarize the analysis of variance results for these data. Figure 4.10 shows the mean results for the six subjects for the amplitude of the M100 RMS peak. Although in the mean group data the response amplitude was larger in the left hemisphere for all six phonemes, suggesting an effect of hemisphere, the main effect of hemisphere was not significant, F(1, 5)=0.725, p=0.433, eta=0.35, as can be seen in Table 4.8. Table 4.8 shows that there was a significant main effect of Phoneme, F(2, 10)=8.135, p=0.008, eta=0.79, and a significant Fundamental x Phoneme interaction, F(2, 10)=4.163, p=0.048, eta=0.67. There were no other significant or marginally significant effects.

TABLES 4.7, 4.8, AND 4.9 HERE

Figure 4.11 summarizes the RMS peak data for individual subjects and phonemes. For three subjects (JM, AM, and JH), the response amplitudes measured from the left hemisphere were larger for each stimulus, and for two subjects (RR and IS) the response amplitudes from the right were always larger. In one subject (MM) there was no differential response effect of hemisphere.

FIGURES 4.10, 4.11, 4.12, AND 4.13 HERE

The mean latency of the RMS peak is shown in Figure 4.12. The mean latency of the M100 peak for all stimuli was in the interval from 100 to 140ms after stimulus onset. With the exception of the /u/ response, the left hemisphere response latencies exceeded the right hemisphere responses, although the effect did not reach statistical significance, F(1, 5)=4.007, p=0.102, eta=0.67. Note, however, that the effect size eta was rather large. As Figure 4.12 suggests, the two /a/ stimuli had an earlier peak latency than the other vowels in both hemispheres. Statistically, the main effect of Phoneme was, in fact, highly significant, F(2, 10)=26.98, p=0.000, eta=0.92 (Table 4.9).

The graphical intuition derived from Figure 4.11 is that there is a considerably stronger effect of hemisphere than revealed by the statistical analysis. Consequently the data were recoded and reanalyzed, specifically with S6: IS in mind. S6 showed the reverse pattern of responses on virtually all of our tests. Because he is a left-hander with left-handed relatives, he has a fair chance of being right dominant, which would account for his reversed pattern of responses. When the data were coded by 'presumed dominant hemisphere' (i.e. S6: IS was recoded), a reanalysis of the RMS data (see Table 4.20) revealed that the only significant effects were still a main effect of Phoneme, F(2, 10)=8.135, p=0.008, and a Fundamental x Phoneme interaction, F(2, 10)=4.163, p=0.048. The factor Hemisphere gained in strength but did not reach significance, F(1, 5)=4.215, p=0.095, possibly due to the low power of the study. When the two left-handed participants were omitted from the analysis of variance (Table 4.21), only the main effect of Phoneme remained significant, F(2, 6)=10.13, p=0.012. The same reanalysis of the latency data showed that there was a significant main effect of Phoneme (Tables 4.22 and 4.23).

TABLES 4.20, 4.21, 4.22, AND 4.23 HERE

Dipole localization results

Table 4.10 summarizes the estimated source parameters for the dipoles computed for the two types of the vowel /a/. The correlations ranged from 0.87 to 0.99 and the goodness of fit ranges from 0.86 to 0.99. On average, the correlations and fits were worse than for pure tones.

TABLE 4.10 AND FIGURES 4.14 AND 4.15 HERE

Figure 4.14 plots the single dipole data from Table 4.10 computed for the two /a/ stimuli for the first three subjects (S1 - S3). Figure 4.15 plots the same data for subjects 3-6. The localizations for vowels were completely overlapping with the localizations for tones, in particular if one takes into account the typical error in localization. In all three dimensions, the single dipoles group in the same general area of auditory (association) cortex posterior to Heschl's gyrus. Note in Figure 4.14 that both in the left and the right temporal cortex the responses are extremely close for two of the subjects (S1: JM and S3: RR). For these two subjects, the dipoles for the two /a/ stimuli with different fundamental frequencies were remarkably close to one another in both hemispheres. For the third subject, S2: MM, the responses were not as tightly grouped by phoneme. In Figure 4.15, a similar situation obtains. For two of the subjects (S4: AM and S5: JH) the dipoles generated by the two /a/ sounds are extremely close in space. For the third (left-handed, outlier) subject, S6: IS, the dipoles were not as close.

To demonstrate that the dipoles generated by another vowel group together as tightly as those for /a/, Figure 4.16 plots, in an analogous fashion to Figures 4.14 and 4.15, the dipoles for /u/. For four of the subjects (S1, S3, S4, S5), the dipoles were, again, closely spaced, typically within 0.5cm of one another. As was true in the analysis of the responses to /a/, subjects S2 and S6 (left-hander) generated a different result, with their /u/-dipoles being spaced apart more widely. The same situation obtained for comparisons of dipole positions for the vowel /i/. One noteworthy feature about the dipole positions is that there is very little variation in the superior-inferior (z) dimension. Even when dipoles to the vowels were slighly further apart in the x-y plane, the z-coordinates (for all three vowels at two F0 values) remained within a small range for the four subjects that showed consistent localizations.

FIGURE 4.16 HERE

In Figure 4.17, the dipole localizations for all three vowels at both F0 values are plotted for one subject (S3: RR) to evaluate if there is any systematic grouping of vowel locations ('phonotopy'). The cumulative plotting of these data show that all of the presented vowels localize to small areas in auditory (association) cortex (revealed by MR overlays; see below). Noteworthy is the very tight localization in the right hemisphere: all 6 dipoles fall within 0.5 cm in each dimension. In the left hemisphere, the variation is somewhat more extensive, but even in the most variable dimension (anterior-posterior) the dipoles still localize to within one centimeter of one another. There is no evidence of any systematic grouping that suggests phonotopic organization.

FIGURE 4.17 HERE

4.18 displays, for one subject (AM) the trajectories of three dipoles in one dimension, the medio-lateral (y) axis, for the recordings from the left hemisphere of three (female, 200Hz F0) vowels. The data were graphed in this way to compare our results to those of Hayashi (1992). When Hayashi plotted the dipole trajectories for 5 female-pitched, natural Japanese vowels they appeared to line up by F1 and crossed over to line up by F2. We sought to replicate this report in order to ascertain, for vowels, whether formant-structure relates in an interpretable way to evoked responses as assessed by M100 single dipole esimates. The data in Figure 4.18 indicate that the trajectories line up and cross over in a comparable manner. Whereas there is a clear medio-lateral shift for the vowel /a/, there is no noticable shift for /u/. The trajectory for /i/ is complex and does not suggest any obvious interpretation.

FIGURE 4.18 HERE

## Discussion

Other than the effect of phoneme, as a main effect and as an interaction effect with the fundamental frequency, there was no statistical evidence of any hemispheric asymmetry for vowel processing. The factor *Phoneme* (/a/ versus /i/ versus /u/) generated significant results, both when looking at the amplitude and the latency of the RMS peak. The phoneme /a/ may in fact be coded such that its unique features generate a different evoked field. The main feature differentiating /a/ is the notably the higher energy at the first formant (a defining component of vowel height). It is critical that the interaction (Phoneme x Fundamental) was significant. Insofar as an interaction in the same direction is evident in recordings from both hemispheres, the result supports the model that the two auditory cortices are contributing to the representation of speech sounds.

Despite the absence of effects as revealed by statistical analysis, the graphically derived intuition still remains that there are effects of hemisphere. The individual RMS data in Figure 4.11 show that in three subjects the left response was without exception larger than the right response. In contrast, in two subjects (RR and IS), the right amplitude was always larger than the left amplitude. There is no conclusive way to disentangle these data, and so it remains open whether there are hemispheric lateralization effects for vowels. As mentioned in the Introduction, the literature on vowel perception and laterality is contradictory, so there is no independent evidence to help us clarify these findings.

Particularly intriguing results were presented by Hayashi (1992). Hayashi presented his subjects with five Japanese vowels and reconstructed the single dipole underlying the evoked field. He computed the dipoles around the M100 for an interval from 80 to 140 milliseconds and mapped the variation of the apparently moving single dipole coordinates in that interval. In the anterior-posterior (x) dimension, the dipoles for all five vowels moved anteriorly over the interval, with all five vowels remaining along a similar trajectory. Similarly, in the superior-inferior (z) dimension, the dipoles for all five vowels moved inferiorly over the interval. In the lateral dimension (y), the dipole trajectories lined up in an interesting way. Although Hayashi did not interpret these data, at the beginning of the interval, the dipoles appeared to 'line up' from medial to lateral by increasing second formant frequency (i.e. from back vowel to front vowel). Whether this was a consistent finding across subjects was not clear from these data.

Hayashi's data were possibly the first evidence that the linguistic-specific internal structure of vowels, the formants, may provide clues to the processing of speech stimuli. To engage these questions (from the angle of left right symmetry), we analyzed the the dipole trajectory around the M100 for six subjects with three vowels (/a i u/) at two different fundamental (100Hz and 200Hz). We presumed that we might detect formant structure as well as the interaction of vowels with fundamental frequency.

The dipole localization results exemplified in Figures 4.14 and 4.15 are provocative. In 4 of the 6 subjects, the responses to the two /a/ stimuli of different fundamental frequencies were completely overlapping, suggesting that the phoneme was possibly coded as a phoneme independent of its perceived pitch. The result for the vowel /u/, depicted in Figure 4.16, was similar: for four subjects, the dipoles are very closely spaced and for the other two (MM and IS) there was larger variability, especially also in the superior-inferior aspect. The data for one subject displayed in 4.17 illustrate that all the dipoles localized to a well circumscribed area and that there was little superior-inferior variability. While Figure 4.17 provides no hint of a vowel space, or any underlying 'phonotopic organizational principle,' localizing the activity elicited this way to such a small area does suggest that we are successfully engaging a region that responds vigorously, if late, to speech stimuli. It is important to bear in mind, too, that the dipoles for the pure tone localized to the same position, so the localizations are not distinguishing between speech and non-speech.

## EXPERIMENTS 3A AND 3B: PASSIVE SYLLABLES AND SYLLABLE DISCRIMINATION

### Introduction

Several magnetoencephalography studies have explored the evoked neuromagnetic responses to CV syllables, including Kaukoranta et al. (1987), Kuriki and Murase (1989), Hari (1991), Sams et al. (1991), and Hayashi (1993). Only Kuriki & Murase (1989) recorded from both hemispheres and compared the results with respect to the lateralization of responses. They reported that, as compared to the responses to the vowel /a/, the M100 field generated by the syllable /ka/ consistently localizes to a more posterior aspect in auditory cortex, although they do not report whether their localizations place the dipoles in primary or association auditory cortex. The question whether there are differential hemispheric responses to syllables and their intrinsic properties remains open. A number of issues arise with respect to the processing of syllables, including whether the responses are similar or different to the responses for vowels and tones, whether there are systematic effects of voicing or place of articulation of the syllable-initial stop-consonant on the responses, whether the responses correlate with the amount of energy in a stimulus, and so on.

While these factors play a key role in the analysis and interpretation of the two experiments, one core hypothesis was at stake in this pair of studies, namely whether the addition of a task modulates the responses measured from auditory cortex in a selective way. There are two aspects two the issue. The first is whether the M100 is in principle influenced by task (attentional) demands or whether the M100 reflects processing that is independent of any additional (perhaps extraneous) computations. A number of investigators have been able to show that auditory evoked response components such as the M100 can be influenced by certain task conditions (for review, see Näätänen, 1992). For example,Woldorff et al. (1993) presented subjects dichotically with tones of different frequencies and showed that the M100 generated by the stimulus in the attended ear was larger than the M100 generated by the same stimulus when presented to the unattended ear. The evidence that the M100 amplitude is subject to attentional manipulation is solid, at least for recordings from the left hemisphere.

A second attention-related issue arises from the literature on neuroimaging of speech sounds with PET or fMRI (Chapter 3). That literature is consistent with the notion that speech sounds are processed bilaterally in the posterior superior temporal gyri. Moreover, there is some evidence that there is an interaction between the bilateral activation in posterior STG and the presence or absence of tasks. Several studies suggest that when subjects execute certain task demands, typically 'meta-linguistic' tasks such as phoneme categorization or semantic classification, the observed neural activity is more highly lateralized, with an asymmetric activity increase typically measured in the left temporal cortex. The two experiments presented in this chapter test the hypothesis that there is an attention-dependent gating of activity that differentially affects the two hemispheres. The the same stimulus material is presented in two different experimental conditions: passive listening (Experiment 3a) or performing a discrimination task (Experiment 3b). The prediction was an increase of activity in the left temporal cortex in the attended condition. There was no prediction about the response of the right hemisphere.

### Materials and Methods

In the two syllable conditions, passive listening (BPDT, Experiment 3a) and discrimination (BPDT_ID, Experiment 3b), subjects heard the syllables /bæ/, /pæ/, /dæ/, and /tæ/. For information on synthesis, see the description of vowel synthesis above. Figures 4.19-4.22 show spectrograms, waveforms, and power spectra for each of these four stimuli, respectively.

FIGURES 4.19, 4.20, 4.21, AND 4.22 HERE

Analogously to the vowel condition, in the passive condition subjects were presented with the four syllables in pseudorandom order at a variable ISI (1000-1500 ms).

In the discrimination condition, subjects had to execute a 2-alternative forced-choice decision, grouping stimuli by VOT, i.e. grouping /bæ/ and /dæ/ versus /pæ/ and /tæ/ in a voiced/voiceless classification. Subjects pressed one response button on a custom-built button box when perceiving a (short, 10ms VOT) /bæ/ or /dæ/ and another when perceiving a (long, 80ms VOT) /pæ/ or /tæ/. Subjects' responses and response latencies were collected by the PsyScope software.

## Results

### Results for Experiment 3a: passive listening to CVs

Evoked neuromagnetic field results

Table 4.11 (3a) lists the M100 RMS peaks and the latencies of the M100 RMS peaks for each subject and each CV. Figure 4.23 graphically displays the individual subjects' M100 RMS peak latency data and Figure 4.24 displays the individual subjects' M100 RMS peak amplitude data. Because of the complex nature of these data, it is important to work through the patterns generated by individual subjects.

FIGURES 4.23 AND 4.24 HERE

Two basic patterns of results were evident in the latency data (Figure 4.23). For three of the subjects (JM, AM, and JH) the M100 latencies were slightly longer in the left hemisphere (black bars) than in the right hemisphere (white bars). For the other three subjects there was no hemispheric preference and no other internal pattern. Analyses of variance of the latency data reveal no significant or marginally significant main effects or interactions. Importantly, there was no effect of hemisphere, even though the Figure 4.23 suggested an effect in that direction.

The amplitude data (Figure 4.24) showed complex patterns of results. In two subjects (JH, IS) the data were extremely parallel in the left and right recordings: the voiced (/bæ/, /dæ/) syllables had larger RMS amplitudes than the voiceless syllables (/pæ/, /tæ/) in both sides. These are the data patterns one would expect if the response revealed sensitivity to the energy in the stimulus. The second systematic pattern in these results was that in four of the subjects (JM, MM, RR, AM), the voiceless syllables generated larger amplitudes than the voiced syllables in the recordings from the right temporal cortex. In an ANOVA that included all six subjects (Table 4.12), the interaction Hemisphere x Voicing came out significant, F(1, 5)=9.039, p=0.030, eta=0.80.

Dipole localization results

I present the localization results for all subjects below. The results are presented graphically in the Figures below, together with the results from the syllable discrimination condition. Briefly, the dipoles generated by the passive syllables clustered in very tight regions in the left and right auditory association cortices, in th same area that was was activated by vowels and tones.

### Results for Experiment 3b: discrimination of CVs

Behavioral Results

Of the total of 4800 trials acquired in this experimental condition (6 subjects x 2 hemispheres x 4 syllables x 100 repetitions/syllable) only 12 trials had to be excluded because of equipment failure or the failure of the subjects to press the response button. In the 4788 analyzed trials, participants performed extremely accurately. 5 of the 6 subjects performed with an accuracy rate ranging from 98.25% - 100%, corresponding to an absolute number of errors ranging from 0 (subject MM) to 14 (subject RR). One subject (S6: IS) performed worse, at 94.1%, making a total of 47/799 categorization errors. This higher error rate may be due to the fact that the subject is not a native (L1) speaker of English and may therefore have found the categorizations slightly more difficult, even though the status of these CVs as non-words and their phonetic distinctiveness is the same in that subject's native language (German).

Figure 4.25 summarizes the mean reaction time data for this condition.The analysis of variance for these data showed no significant main effect of Hemisphere, F(1,5)=3.707, p=0.112, eta=0.625, or Phoneme (CV), F(3,15)=1.012, p=0.415, eta=0.410, and no significant Hemisphere x Phoneme interaction, F(3,15)=2.233, p=0.126, eta=0.555.

FIGURE 4.25 HERE

Evoked neuromagnetic field results

Table 4.11 (3b) lists the RMS peaks and the latencies of the RMS peaks for each subject for the syllable discrimination condition. Figure 4.26 shows the individual subjects' RMS peak data. In four of the subjects (JM, AM, JH, IS), the pattern that was attested in the passive condition in two subjects obtains, namely larger responses for voiced than voiceless stops, particularly in the left hemisphere. The Task x Voicing interaction was significant, F(1,5)=8.70, p=0.032, eta=0.79. The effect of a larger RMS amplitude in the right for voiceless stops is only evident in one subject (AM). Analyses of variance of these data and those of the passive syllable condition are summarized in Table 4.12. These analyses of the latency data revealed no significant effects, particularly of hemisphere.

FIGURE 4.26 HERE

To focus on the central question, the effect of having an attentionally explicit task, it is necessary to better quantify the difference of having a task. Statistically, there was no main effect of task, F(1,5)=0.349, p=0.580, eta=0.25, no Hemisphere x Task interaction, F(1,5)=3.933, p=0.104, eta=0.66, but a Task x Voicing interaction, F(1,5)=8.70, p=0.032, eta=0.79, with the response to the voiced stops being more altered by the presence of the task. As the comparison of Figures 4.24 and 4.26 suggests, however, the intuitive difference between the passive and the discrimination conditions is the RMS increase that appears to be manifest in the left. As a way of quantifying the effect, Table 4.13 computes the difference, for each subject and each syllable, of the RMS peak in the syllable discrimination and passive conditions.

**TABLE 4.13** *Difference of task and no task RMS amplitudes for six subjects*

| | | Task -passive RMS difference values (fT) | | Average change | |
|---|---|---|---|---|---|
| | | left | right | left | right |
| | stimulus | | | | |
| S1: JM | bæ | 20.69 | -7.01 | -0.8 | -23.8 |
| | pæ | 1.6 | -31.76 | | |
| | dæ | 24.67 | -28.81 | | |
| | tæ | -50.13 | -27.71 | | |
| S2: MM | bæ | 9.37 | -11.1 | 18.48 | -25.8 |
| | pæ | -2.72 | -39.52 | | |
| | dæ | 28.06 | -19.3 | | |
| | tæ | 39.2 | -33.26 | | |
| S3: RR | bæ | 15.55 | -5.82 | 33.33 | -11.49 |
| | pæ | 36.27 | -28.45 | | |
| | dæ | 35.9 | -3.1 | | |
| | tæ | 45.62 | -8.58 | | |
| S4: AM | bæ | 1.43 | -3.86 | 4.9 | -15.79 |
| | pæ | 2.81 | -25.07 | | |
| | dæ | 21.01 | -18.29 | | |
| | tæ | -5.65 | -15.93 | | |
| S5: JH | bæ | 52.25 | -20.24 | 40.07 | -6.98 |
| | pæ | 18.46 | -10.71 | | |
| | dæ | 54.79 | -7.86 | | |
| | tæ | 34.79 | 10.9 | | |
| S6: IS | bæ | 5.72 | 63.45 | -5.86 | 26.77 |
| | pæ | -8.29 | 37.33 | | |
| | dæ | -22.17 | 44.25 | | |
| | tæ | 1.3 | -37.94 | | |

This analysis reveals a striking difference with respect to how the left and right responses are modulated by the presence of the discrimination task. Figure 4.27 illustrates the effect graphically. Consider first the left hemisphere data from subjects S2-S5 (the right-handed subjects). In 14 of 16 measurements, the task increased the mean RMS value in the left, with the increase ranging from 1.43 to 54.79 fT. In the right hemisphere measurements from the same four subjects, 15 of 16 measurements show a down-regulation in the right hemisphere, with values ranging from 3.1 to 39.52 fT. Subjects S1 and S6, the left-handers, pattern in two directions. The response pattern for S1 is comparable to the patterns for S2-S5, with the exception of one single value (/tæ/) which significantly altered the overall pattern. The comparison of responses for S6 are in the opposite direction: the values in the left hemisphere, on average, decrease, whereas it is the right hemisphere values that show a large increase. It is important to note that the changes, with few exceptions, are extremely large. Typical RMS values of evoked M100 responses (cf. Table 4.11) lie between 100 and 200 fT. The task-dependent modulations are on the order of tens of percent of change.

FIGURE 4.27

All these data were analyzed in two additional ways, once by recoding by presumed dominant hemisphere and once by omitting from the analyses the two left-handed subjects. These analyses are summarized in Tables 4.24-4.31. Particulary noteworthy about these analyses is the Hemisphere x Task interaction. When the data are analyzed as a group, that effect was not significant. When the two left-handers are omitted (Table 4.25), that interaction is highly significant, F(1,3)=39.957, p=0.008, and when the one left-handed subject is recoded, the effect is also extremely large, F(1,5)=55.307, p=0.001 (Table 4.24). The latency data show a significant main effect of Hemisphere after the recoding (Table 4.26), F(1,5)=12.214, p=0.017. That effect was no longer in evidence when the two left handed participants were omitted from the analysis.

An additional analysis was done to confirm the presence or absence of hemisphere effects: because the passive and the task condition were run separately, the data were analyzed independently (see Tables 4.30-4.33). In these ANOVAs, the M100 RMS and latency data were analyzed by themselves to ascertain any laterality effects. The important contrast is provided by the comparisons between Tables 4.30 and 4.32, and 4.31 and 4.33, respectively. In the first case, the data were recoded by presumed dominant hemisphere. In the passive listening condition (Table 4.30), there was no main effect of Hemisphere, F(1,5)=0.089, p=0.777. In contrast, when the syllables were discriminated (Table 4.32), there was a main effect of Hemisphere, F(1,5)=11.282, p=0.020. Similarly, when the two left handed subjects are omitted from the analysis (Tables 4.31 and 4.33), there is no main effect of Hemisphere in the passive listening condition (Table 4.31), F(1,3)=0.460, p=0.546, but a main effect of Hemisphere approaching significance in the task-modulated condition (Table 4.33), F(1,3)=8.432, p=0.062 despite the low power (n=4). These analyses highlight that the presence of the task 'induces' the lateralization.

### Dipole localization results

For all subjects, the best single equivalent current dipole was computed for each of the four syllables in the two conditions. Tables 4.14-4.19 provide the parameters of interest for each dipole and Figures 4.28-4.33 graphically display the dipoles for the subjects.

TABLES 4.14-4.19 AND FIGURES 4.28-4.33 HERE

Note in Table 4.14, for subject AM, that the correlations for syllable presentation are between 0.96 and 0.99, rather high correlations that correspond well to those obtained with pure tone stimuli. For subject JH, whose dipole data are summarized in Table 4.15, the correlations are slightly lower, ranging from 0.91 to 0.98. Figures 4.28 and 4.29 show that the dipoles for the two conditions are extremely closely spaced. For subject AM, whose dipole data are graphed in Figure 4.28, there is a single outlier; the remaining 15 dipoles (2 hemispheres x 2 conditions x 4 syllable types) are very tightly spaced in all three dimensions. (Some dipoles in the right are on top of on another and are consequently not seen in the display.) The left and right regions in which the syllable dipoles are clustered were the same regions of auditory association cortex in which the subject had the vowel dipole cluster (cf. Figures 4.15 and 4.16). Similar results obtained for subject JH (Figure 4.29). Particularly consistent are the values in the superior-inferior (z) dimension, in both hemispheres. Importantly, there were no systematic hemispheric differences with respect to dipole localization. Although in subject AM the cluster of dipoles was more anterior in the right, this was true for that subject with vowels and tones as well. This anterior-posterior shift is reported in almost every MEG localization study and is commonly attributed to the large morphological asymmetries between the left and right supratemporal planes (Geschwind & Levitsky, 1968). The most remarkable feature, for syllables as well as for vowels and tones, is the consistency of best single dipole location within a given subject.

### Discussion

The behavioral results from the discrimination condition (3b) are important in two respects. First, as Figure 4.25 shows, there does appear to be a trend in that the right hemisphere/left ear input RTs are uniformly longer than the left hemisphere RTs even though statistical significance was not reached. The effect sizes indicate, however, that the overall effect of Hemisphere is non-negligible. One will have to increase the power of the study to test whether these potentially important trends (that speak at a more macroscopic level to the issue of lateralization) are systematic effects. One important factor that needs to be accounted for is that subjects responded with the right hand during right hemisphere recordings and with the left hand during left hemisphere reecordings, and there may be subtle hand preferences in the use of the response button. Given that there were more right handers than left handers among the subjects, the latency difference goes in a surprising direction. One would expect the right handers to respond faster with the right hand, and these results go in the opposite direction. On the other hand this trend is consistent with the dichotic listening findings. In that literature it is often argued that (Darwin et al., 1978) the left hemisphere is specialized for speech and, as part of that specialization, happens to be good at fast temporal processing. From that perspective it would be natural to expect slightly longer latencies from left ear monaural presentation.

Second, for the interpretation of the neuromagnetic data it was important to establish that there was no difference (main effect) among the syllables, i.e. subjects were not preferentially reponding to /bæ/ or /tæ/, for example. If there had been main effects in the reaction times for CV syllable (either along the voicing dimension or the place-of-articulation dimension), it would be more problematic to compare the neuromagnetic recordings of the syllables. For example, if we had found that subjects responded uniformly faster to /pæ/ than to the other three syllables (perhaps because they were synthesized more naturally/realistically), and if we had also found an earlier peak RMS for /pæ/, we would have to entertain the possibility that the neuromagnetic signal is confounded with some orthogonal factor. Although one can never exclude that possibility, the absence of large behavioral differences among the syllables gives us confidence that the syllables were 'treated' or perceived more or less equivalently. Therefore it seems legitimate to compare the neuromagnetic signals generated in response to the syllables in the discrimination task.

Although there are a number of MEG studies investigating the effect of attentional modulation on the neuromagnetic response, the paradigm used here is somewhat different. In most studies in which attention is one variable of interest, the attentional modulation is implemented by having subjects selectively attend to the input of one or the other ear. In contrast to this type of 'ear-specific' attention, which comes squarely from the tradition of dichotic listening studies, this experiment presents the identical set of speech stimuli under two distinct attentional requirements. In experiment 3a, subjects 'passively' listened to /bæ/, /pæ/, /dæ/, and /tæ/. In Experiment 3b, subjects were presented with the same stimulus set in the same fashion but were required to make an overt classification about a stimulus property, in effect a 'meta-linguistic' judgment. If one assumes that the task, a linguistic judgment, selectively engaged the dominant hemisphere, the results for subject S6: IS, the left-hander generating the different patterns are consistent with the hypothesis that he is right dominant.

## GENERAL DISCUSSION

### Summary of results

Let us first recapitulate the results and summarize how the data speak to the two hypotheses, the bilateral mediation of early speech perceptual processes and the hemispheric interaction between linguistic-attentional modulation and speech-evoked neural activity. For 500Hz pure tones, the evoked field responses showed no hemispheric effects, in amplitude or in latency. With regard to the first hypothesis, as tested by vowels, there were no statistically significant hemispheric effects, although both the RMS and the latency data hinted at a possible 'advantage' of the left hemipshere, insofar as an RMS increase and a faster latency can be interpreted as advantageous. The passive syllable data were more complex (as was true for the task-dependent syllable data), suggesting that the M100 response codes for something more complex than stimulus energy, at least in the right hemisphere, in which for four subjects the response amplitudes went in the opposite direction to that predicted by the physical parameters of the stimuli. The RMS facts for the passive condition also support the idea of bilateral early speech processing in that there were no main effects of hemisphere in three different analyses of the data. With regard to the second hypothesis, the critical finding was that the presence of the task differentially affected the left and right auditory association cortices: the left response amplitudes increased (as hypothesized) and the right response amplitudes decreased as a consequence of executing a linguistic task.

The interpretation of the results requires some caution, of course. The sample size is small (6 subjects), and there are two left-handed subjects included in the analysis, and left-handed subjects are known to pattern differently with respect to hemispheric asymmetries. Moreover, because the large-array magnetometer recording technique is so new, we are not yet aware of all the subtleties of the data obtained with extracranial recording. For example, one cannot really interpret amplitude differences obtained for the same experimental condition in the two hemispheres (because of the multiple ways of creating amplitude differences by manipulating the dewar position) but only interactions that reflect internal structure to the response patterns.

Acknowledging the caveats and limitations mentioned above, and bearing in mind that these data are preliminary in that they come from three small studies, the overall pattern of results is consistent with the model that posits that (at least) early aspects of speech sound processing are mediated bilaterally in the posterior aspect of the superior temporal gyrus (supratemporal auditory cortex). Moreover, the data support the hypothesis that there is an attention-dependent modulation of activity such that left superior temporal cortex is affected. Unexpectedly, the right temporal cortex was also affected by attentional demands, although in the opposite direction to the left superior temporal cortex.

## Lateralization

In contemplating the patterns of data that could be generated when recording from the two hemispheres and considering what kind of data would provide evidence about processing that is demonstrably the same or different between left and right auditory association cortices, one might imagine two possible outcomes. One outcome is that all responses measured from left and right supratemporal cortices are the same along the dependent measure. This outcome would be highly problematic because it would be uninterpretable. The two sides could be computing the speech stimuli identically along every dimension or not at all. Either way, such a pattern of results could not speak to the hypotheses. The other, more ideal possibility is that the left and right responses are complex (e.g. response to af = a > uf = u > i f = i) but that the complex patterns are perfectly replicated in each hemisphere. That sort of pattern would constitute compelling evidence that speech stimuli are processed in a very similar fashion in the left and right. The real data, naturally, lie somewhere between ideal and uninterpretable.

The hypothesis, motivated by the word deafness and the neuroimaging literature, was that the left and right superior temporal gyri both construct speech sound representations, although no claim about the precise nature the computations could be formulated. There is prior evidence that suggests that there may be differences between auditory cortical computation as reflected by the perceptual attribute attended to in an experiment. For example, Zatorre et al. (1992), in a positron emission tomography study using the paired-image subtraction design, auditorily presented their subjects with CVC syllable pairs and required them to make one of two judgments. In one condition subjects had to decide whether or not the final consonants in a syllable were the same or not (phonetic judgment), and in the other subjects had to judge whether or not the pitch change at the end of the syllable was the same or not (pitch judgment). Both task activations were compared to a condition in which subjects passively listened to the same stimulus set. The results showed that the phonetic judgment condition vigorously activated left hemisphere sites, notably Broca's area. In contrast, the pitch condition showed activation foci more sharply in the right hemisphere, notably right inferior frontal areas. These results are suggestive, but do not speak to the processing in auditory cortex itself. It is still unknown whether or not auditory association areas treat speech (and other auditory) input in the same fashion.

The high-temporal resolution data from supratemporal auditory cortex reported in this chapter reveal no obvious asymmetries for any of the passive conditions, which is consistent with the hypothesis that both auditory cortices do participate in speech sound processing. Interestingly, there are nevertheless aspects of the results that point to possible differences. For instance, the M100 latency (latency of RMS peak) is virtually always slightly longer in the left hemisphere. This effect is the least obvious for the tones in Experiment 1, in which for three of the subjects there is an obvious difference, with the left being slightly longer. For the vowels in Experiment 2, the latency measured on the left is longer in for five of six subjects in a clear way, although the effect was not statistically significant (despite the relatively substantial effect size of 0.67). In experiments 3a and 3b (passive CV and discriminate CV), the left latency was uniformly longer. Although none of these latency effects were significant, this may be due to the small sample size. Alternatively, we may be measuring a real effect that remains 'statistically subthreshold.' If the latency difference effect is real, it is quite interesting because it goes in the opposite direction that one would expect. A currently popular idea is that the left hemispere as both specialized for speech and language and specialized for fast temporal processing. If that is true, it is somewhat counterintuitive that the left auditory cortex uniformly generates slower latencies for the major auditory evoked response (M100) than the right.

The important result in the vowel condition (Experiment 2) that speaks to the hypothesis is that there was a main effect of Phoneme and a Phoneme x Fundamental interaction in absence of a hemisphere effect. This is useful insofar as it suggests that both sides are processing the vowels in a fashion similar enough to generate that interaction in such a small sample size. Finally, there were no effects of F0; for example, both types of /a/ had earlier M100 latencies than the other vowels in both hemispheres. These M100 data are consistent with the proposal that vowels are processed independently of their pitch.

In the passive syllable condition (Experiment 3a), there was an intricate pattern of responses for the different subjects. For two of the subjects the patterns of results appeared to be near identical in the two hemispheres. Their results showed a pattern that correlated with stimulus energy. There is more net energy in voiced consonant syllables than unvoiced consonant (cf. Figures 4.19 - 4.22). If the responses we are meauring from auditory cortex reflect the registration of spectral energy at each frequency, the predicted responses are straightforward and would have been those for these two subjects: larger M100 amplitudes for the two voiced stimuli on both sides. On the other hand, another pattern of response amplitudes was seen that violated this pattern in the right hemisphere in four of the subjects. In those four subjects, the right responses were larger for voiceless than voiced stimuli, a response that goes against the direction dictated by the physical stimulus feature energy. On the basis of such results one can argue that the right auditory cortex generates responses to speech stimuli that suggest that it processes the stimuli in a complex way that likely contributes to speech perception.

Like other studies, we find that sometimes the dipoles in the left are more posterior than those in the right (Kuriki & Murase, 1989; Eulitz et al., 1995; Nakasato et al., 1995). Like previous reports, we see that pattern for some subjects (e.g AM) but not for others. In addition, we observe it for all stimuli. However, unlike previous work, we do not find anterior posterior differences between stimulus types: tones, vowels, and syllables localized to the same, small area in supratemporal auditory cortex. Some investigators have reported that vowel-evoked M100 dipoles are anterior to tones, and that vowels are anterior to syllables (Kuriki & Murase, 1989). These differences have been attributed to the gross anatomical asymmetry between the left and right temporal lobes (Geschwind & Levitsky, 1968), but no one has argued that this anterior-posterior asymmetry has functional consequences. Similarly, although we replicate this asymmetry for some subjects, I do not want to attribute functional consequences to it, in lime with previous interpretations. In fact, although the sizes of homologous areas in left and right temporal cortex are different, there are the same number of fields in both sides, and it is plausible that the right hemisphere homologues of posterior auditory association cortex are more anterior in the right.

### The effect of having a task

The second hypothesis, motivated by some neuroimaging results discussed in Chapter 3, was that performing a language task on the stimuli selectively changes processing in the left hemisphere, thus potentially obscuring the role of the right temporal cortex in the processing of the speech sounds in question. In fact, there was a significant difference between passively listening to the CV syllables and actively discriminating among them in a forced-choice task. Having to attend actively to the material led to a significant RMS increase in left auditory cortex. In addition, there was a large decrease in RMS amplitude in the right temporal lobe. This was a surprising new result that was not anticipated by any findings in the earlier literature on MEG or other imaging methods. Because the nature of this response can only be explained in the context of a model of the interaction of attention with speech and language processing, it will have to be further elucidated in future experiments.

The right hemisphere task-dependent amplitude decrease, implies that one must be especially cautious in comparing task-dependent and passive conditions in the verbal domain when evaluating results from other neuroimaging methods as well. Insofar as cortical areas respond differentially to attentional requirements, one is in danger of systematically over- or underestimating the contribution of a given area to the processing of the domain in question. In any case, the interaction reported here predicts that in imaging studies such as subtraction PET studies one will find such an effect as well. In particular, it is predicted that one will detect response decreases in right auitory association cortex and increases in left auditory association cortex.

Of course, the effect of attention needs to be clarified in order to ascertain precisely what kind of attentional modulation is going on. The issue one needs to pursue is what kind of attentional manipulation leads to this response attenuation or enhancement. Several possibilities exist. (1) Only attending to the specific stimulus set changes the attentional gain. The effect is therefore due to subjects having to make overt linguistic discriminations on the presented material. (2) Attending to any speech or language material during the task will result in this effect. It is not crucial to attend to the stimuli proper; attention to any linguistic material, even though it is not part of the task, will result in the change. (3) Finally, it is possible that *any* attentional ask will lead to such a modulation, including, say, the execution of a visual task.

One issue that requires special care is the possible relation of these results to hemispheric dominance for language. It appears as though the attentional reequirements may shift activity to the dominant hemisphere. In the context of such an idea, it is intriguing to find that one subject clearly showed a reverse pattern than the other five in the passive-discrimination comparison. Since that subject reversed on many other patterns and is a left-hander, he was deemed at having a reasonable chance of being right-dominant. For example, in contrast to the five other subjects, his tone RMS peak in the right was much larger than in the left. If the task requirements selectively engage the dominant hemisphere, we may have incidentally tested hemispheric dominance, even though there were no obvious differeces in the passive condition. This is an intriguing possibility. Because one typically assesses language dominance using tasks, it is easy to mask the effects of the non-dominant side.

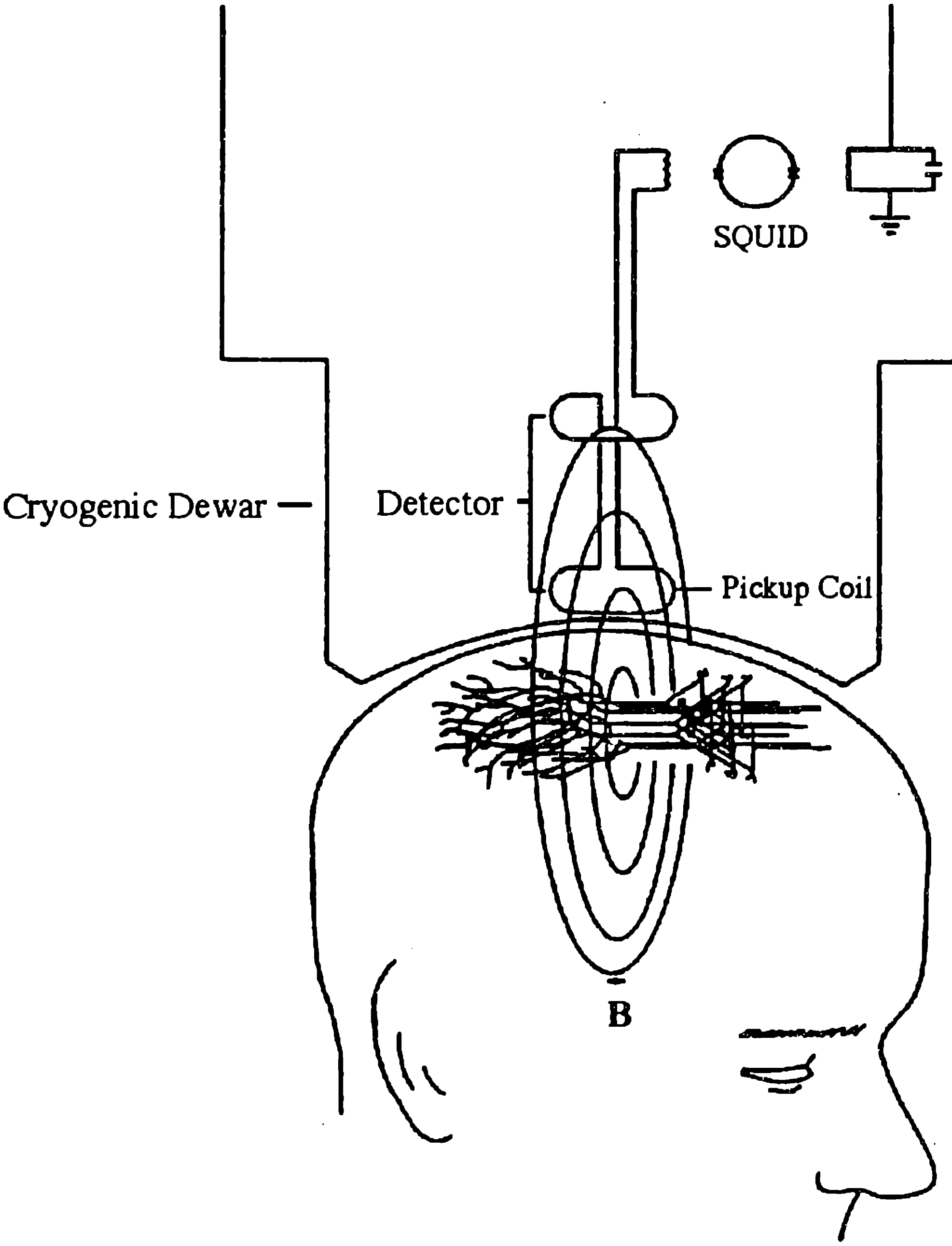
## Which area is activated and what does it do?

All dipole localizations appeared to be in approximately the same area in auditory cortex. Precisely what cortical fields in auditory association cortex are being registered is unknown. Figures 4.30, 4.31, and 4.32, and 4.33 show the dipole overlays for the four conditions (tones, vowels, passive syllables, discriminated syllables) in one subject (JH). The dipoles do not localize to primary auditory cortex. This is somewhat surpising for the tone data, and very surprising given the oft-voiced claim in the literature that the M100 source localizes to primary auditory cortex on Heschl's gyrus.

FIGURES 4.34, 4.35, 4.36, and 4.37 HERE

Both where these recordings are from and what this area does remain open questions. We do know that speech stimul reliably generate evoked responses that appear to originate in that area. It is worth beaing in mind, as well, that we are reliably localizing speech-evoked activity to a small area of cortex that is non-primary. However, given the resolution of the method, it is unlikely that one can detect an organizational principle of 'phonotopy' within such an area. Thus, there may exist a speech perception area which reflects something like phonotopic organization somewhere in auditory association cortex, but it would almost certainly be invisible to this method of recording.

FIGURE 4.1



## FIGURE 4.2

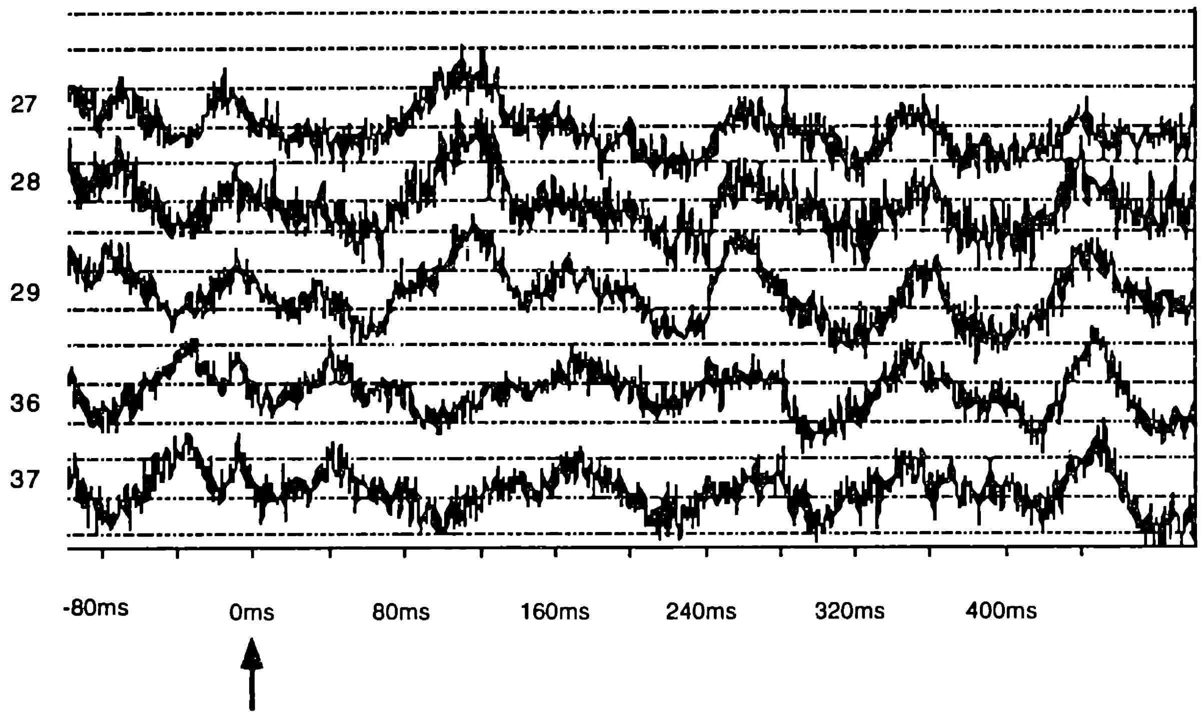
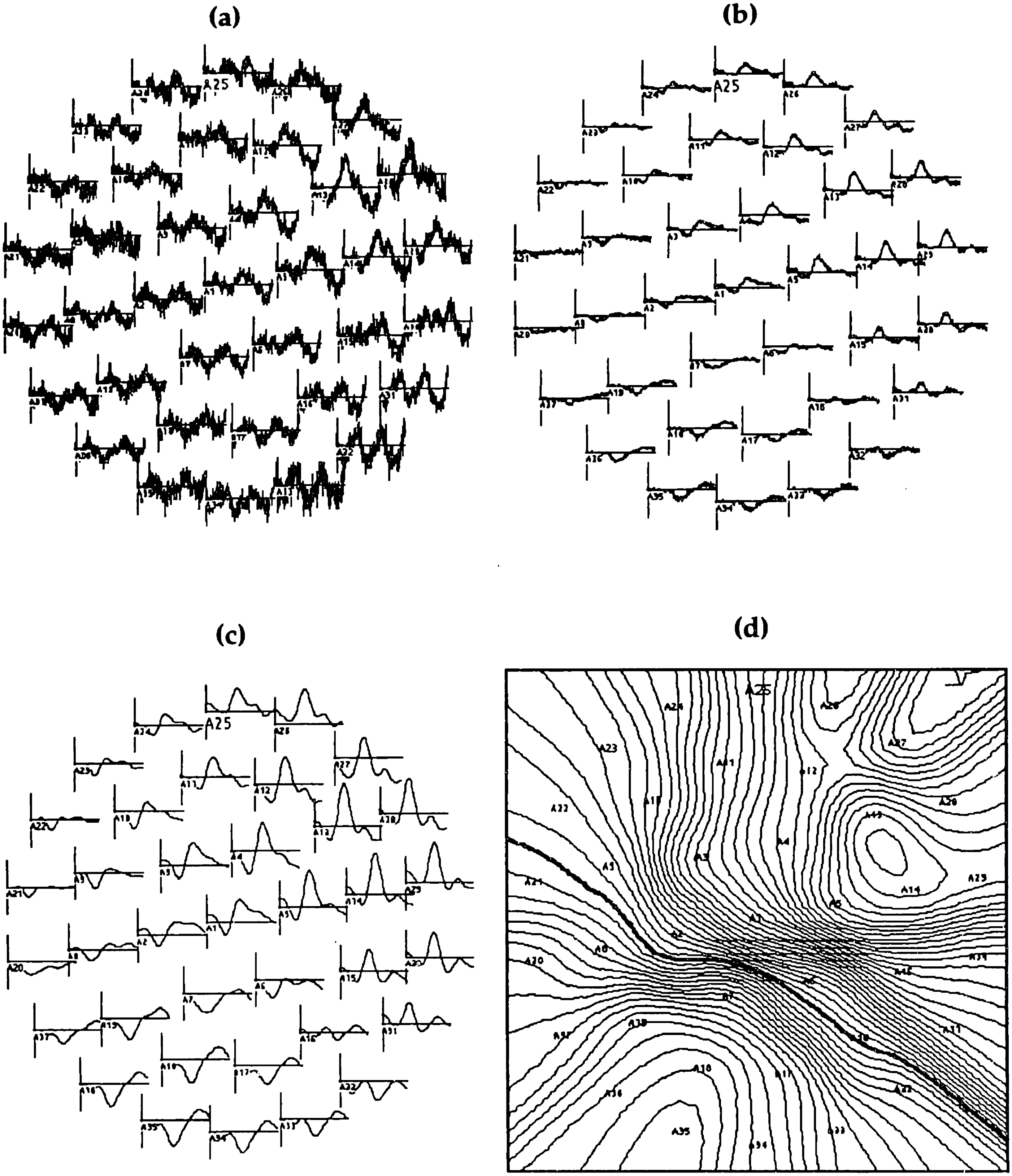


FIGURE 4.3



## MRI-MEG Single Dipole Overlay

Pure tone presentation
to left ear, MEG recording
from right hemisphere

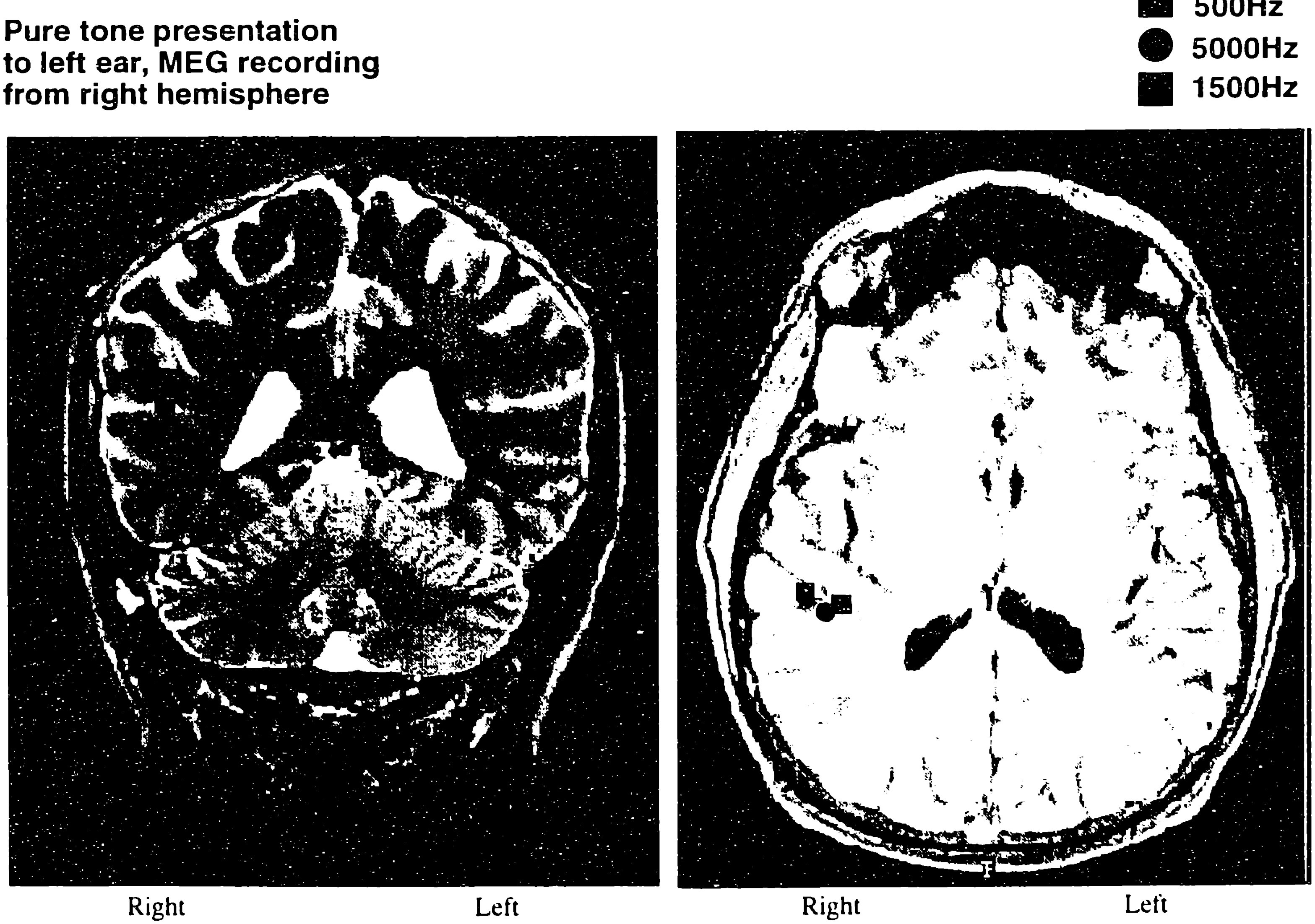
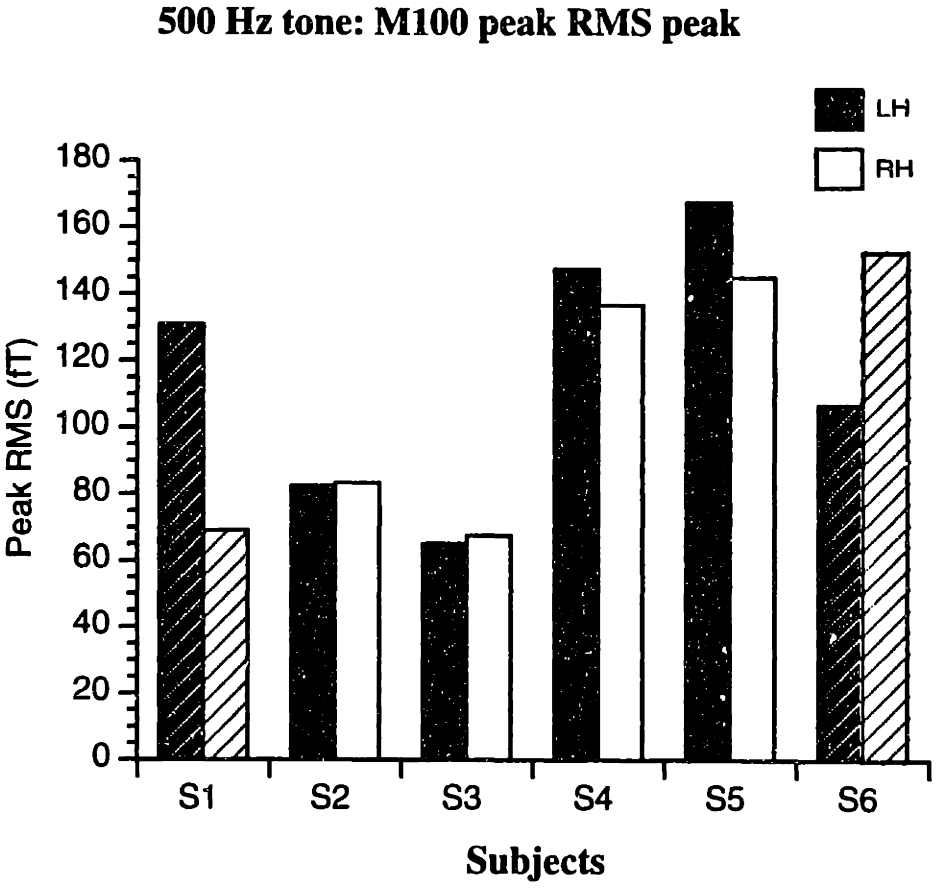


FIGURE 4.4

**FIGURE 4.5**



500 Hz tone: M100 peak RMS peak
LH
RH
Peak RMS (fT)
180
160
140
120
100
80
60
40
20
0
S1
S2
S3
S4
S5
S6
Subjects

**FIGURE 4.6**

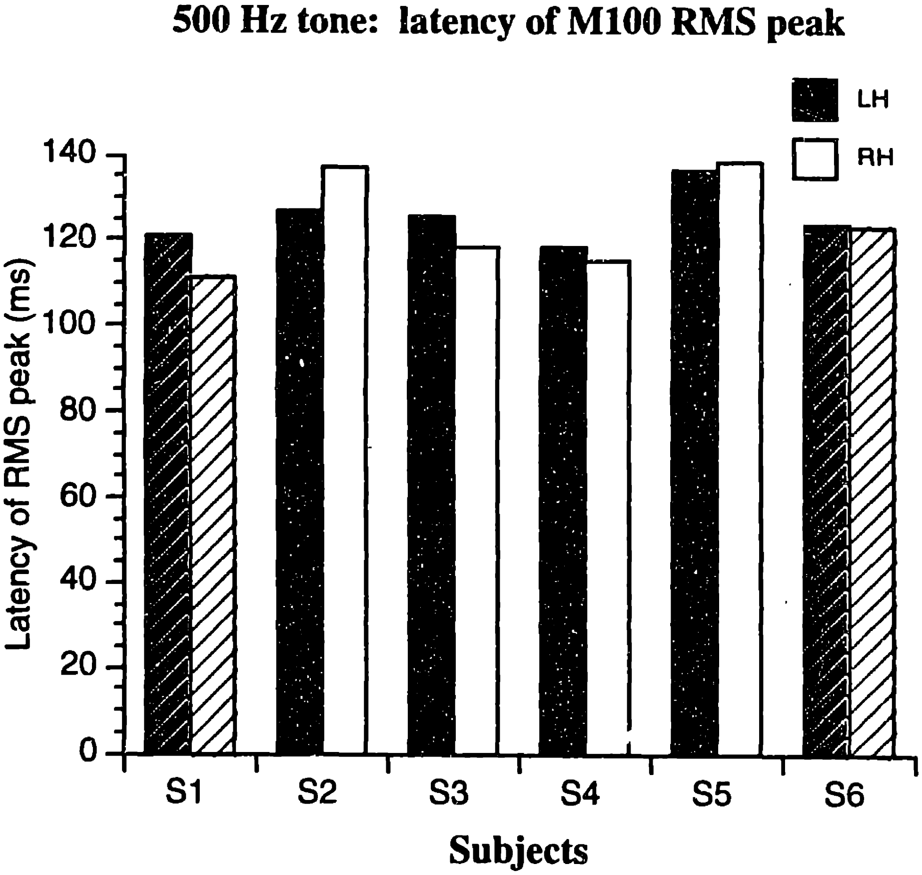

500 Hz tone: latency of M100 RMS peak
LH
RH
Latency of RMS peak (ms)
140
120
100
80
60
40
20
0
S1
S2
S3
S4
S5
S6
Subjects

**FIGURE 4.7 Single equivalent current dipole for 500Hz tone at the M100 RMS peak in both hemispheres (6 subjects)**

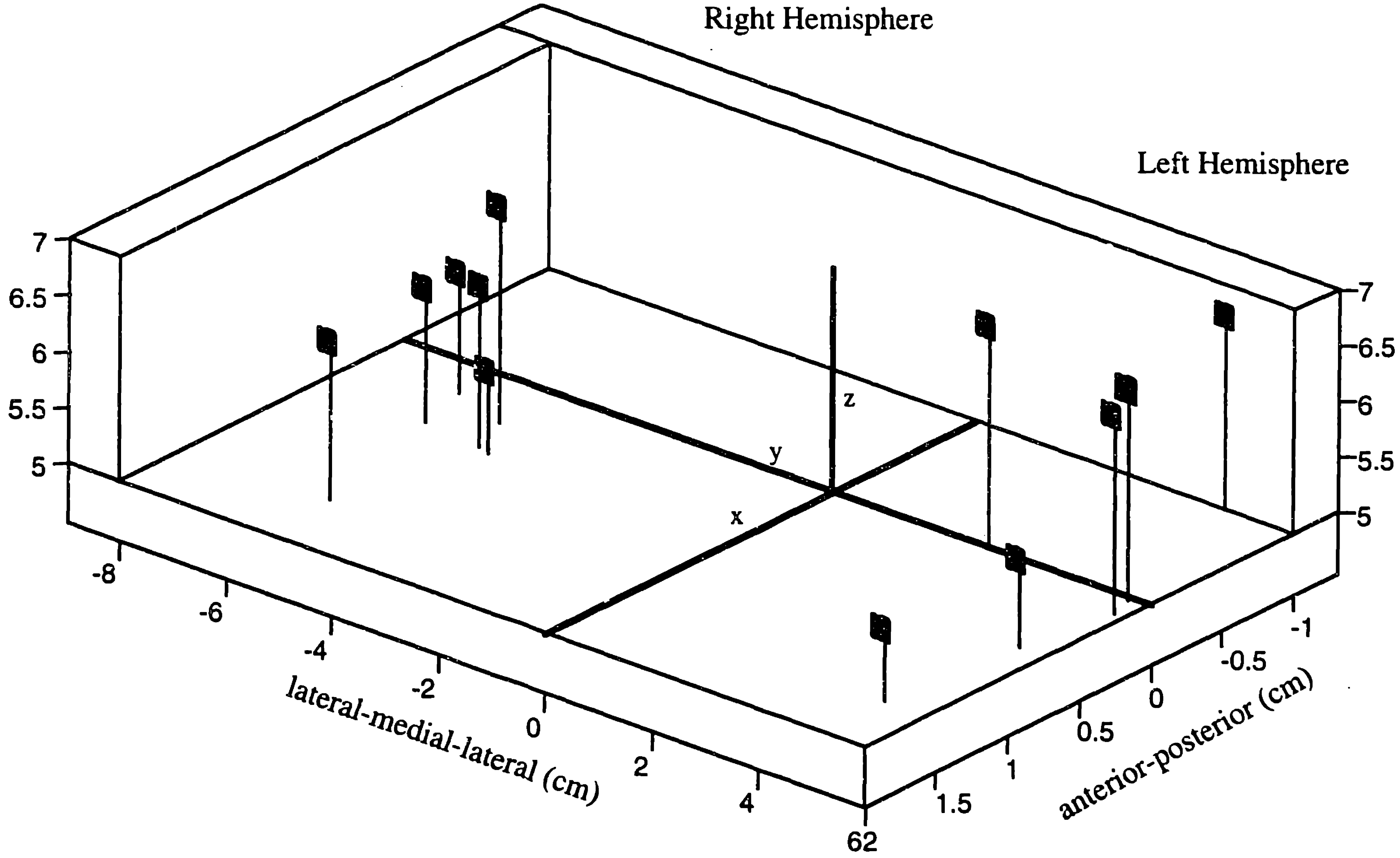
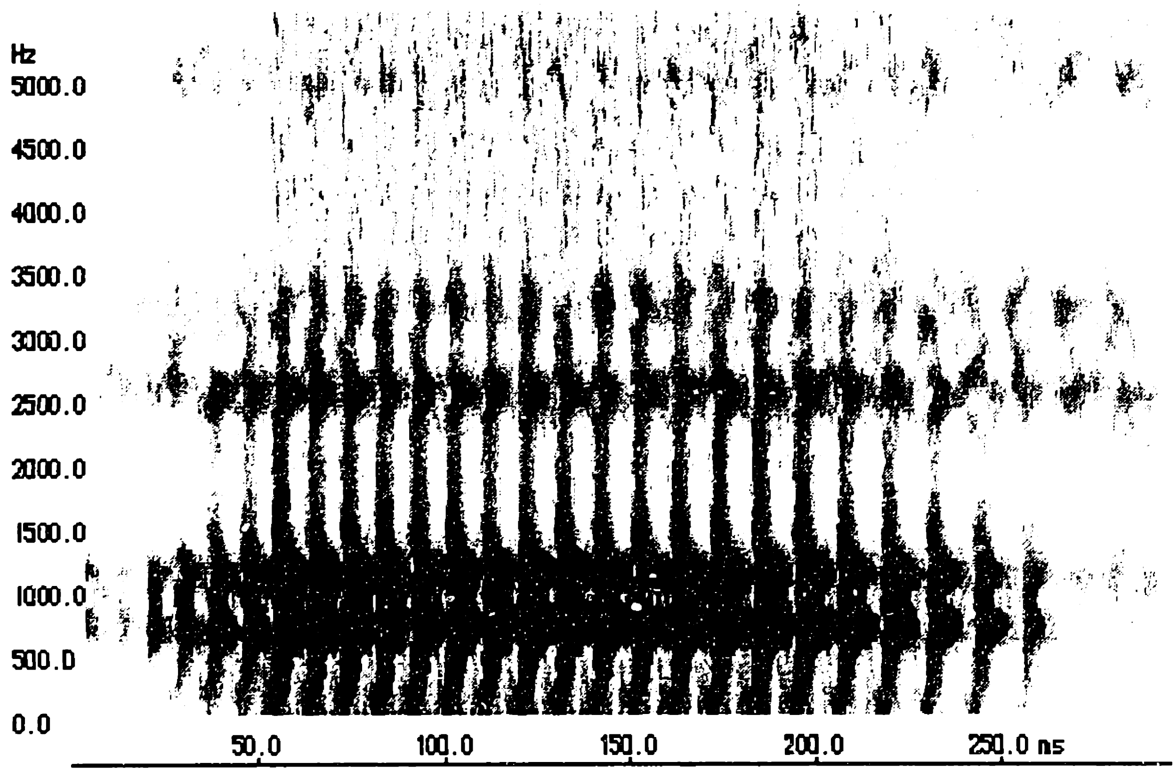


**FIGURE 4.8**

Vowel [a], F0 = 100Hz

Spectrogram



Waveform

Spectrum

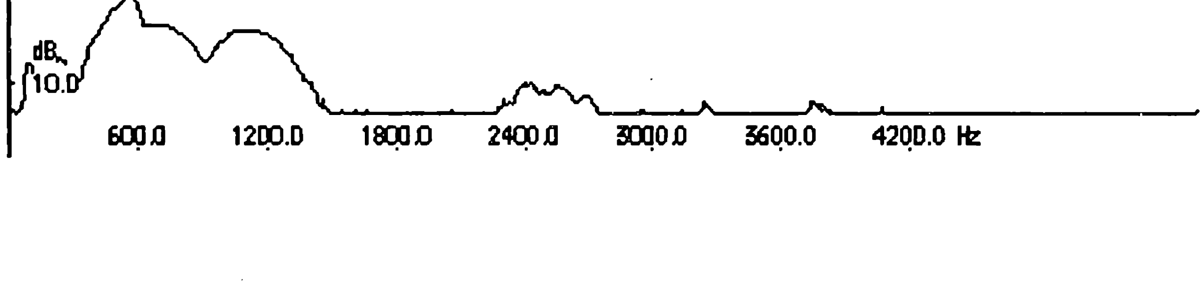
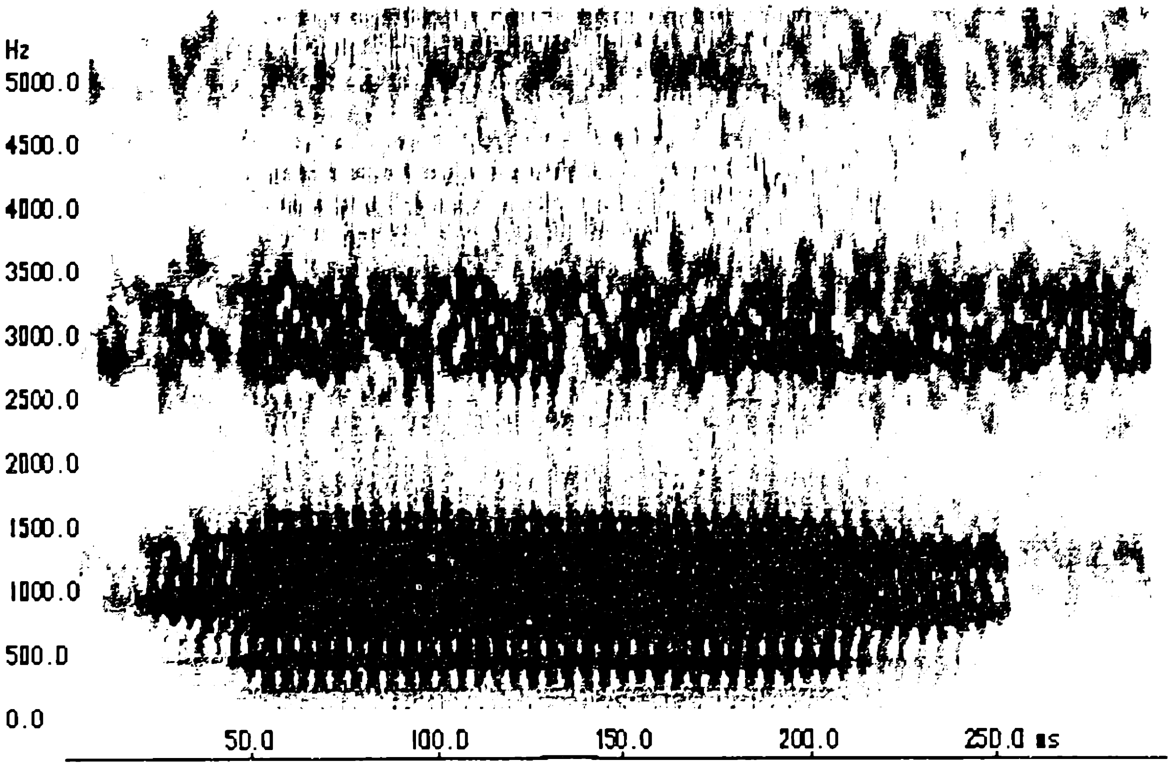


**FIGURE 4.9**

Vowel [a], F0 = 200Hz

Spectrogram



Waveform

Spectrum

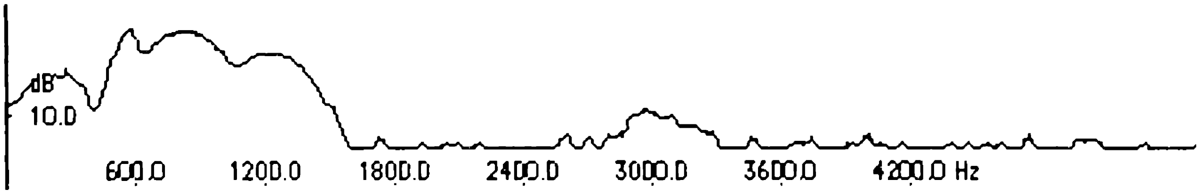


**FIGURE 4.10**

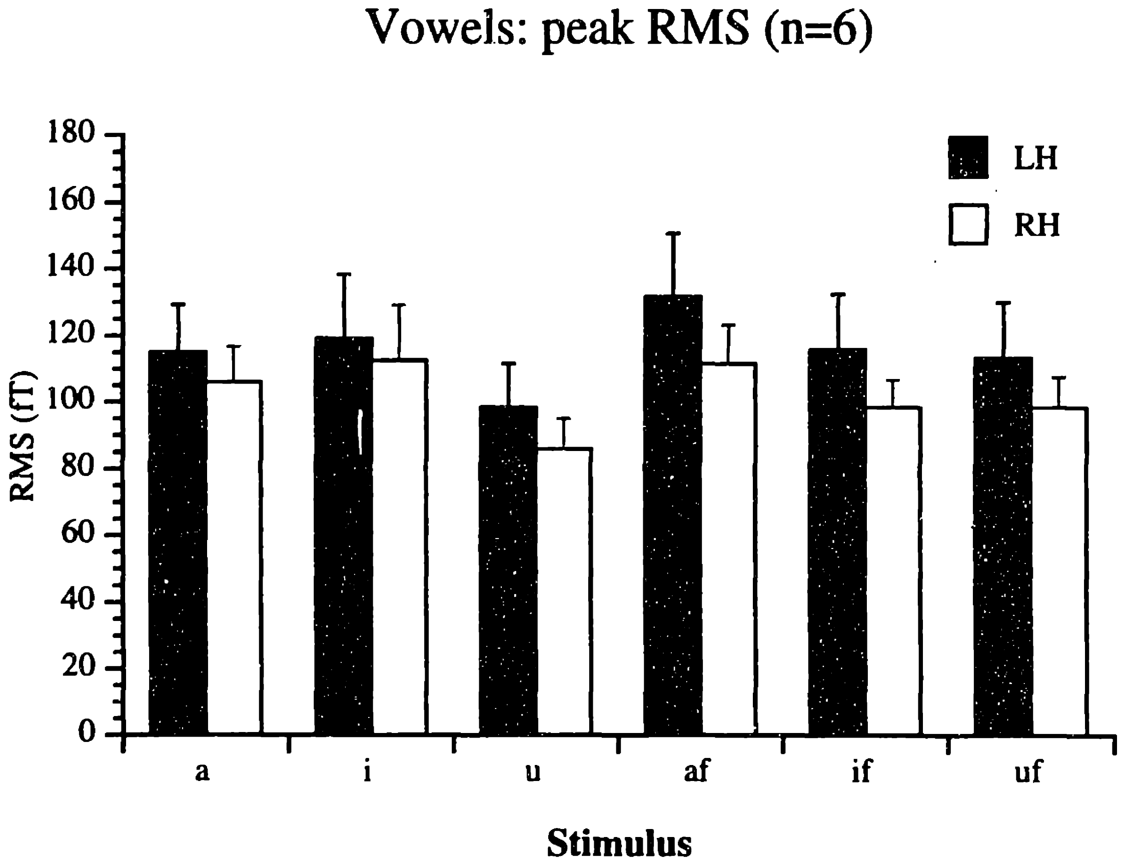

Vowels: peak RMS (n=6)
LH
RH
RMS (fT)
0
20
40
60
80
100
120
140
160
180
a
i
u
af
if
uf
Stimulus

**FIGURE 4.11**

Vowels: RMS peak for 6 subjects

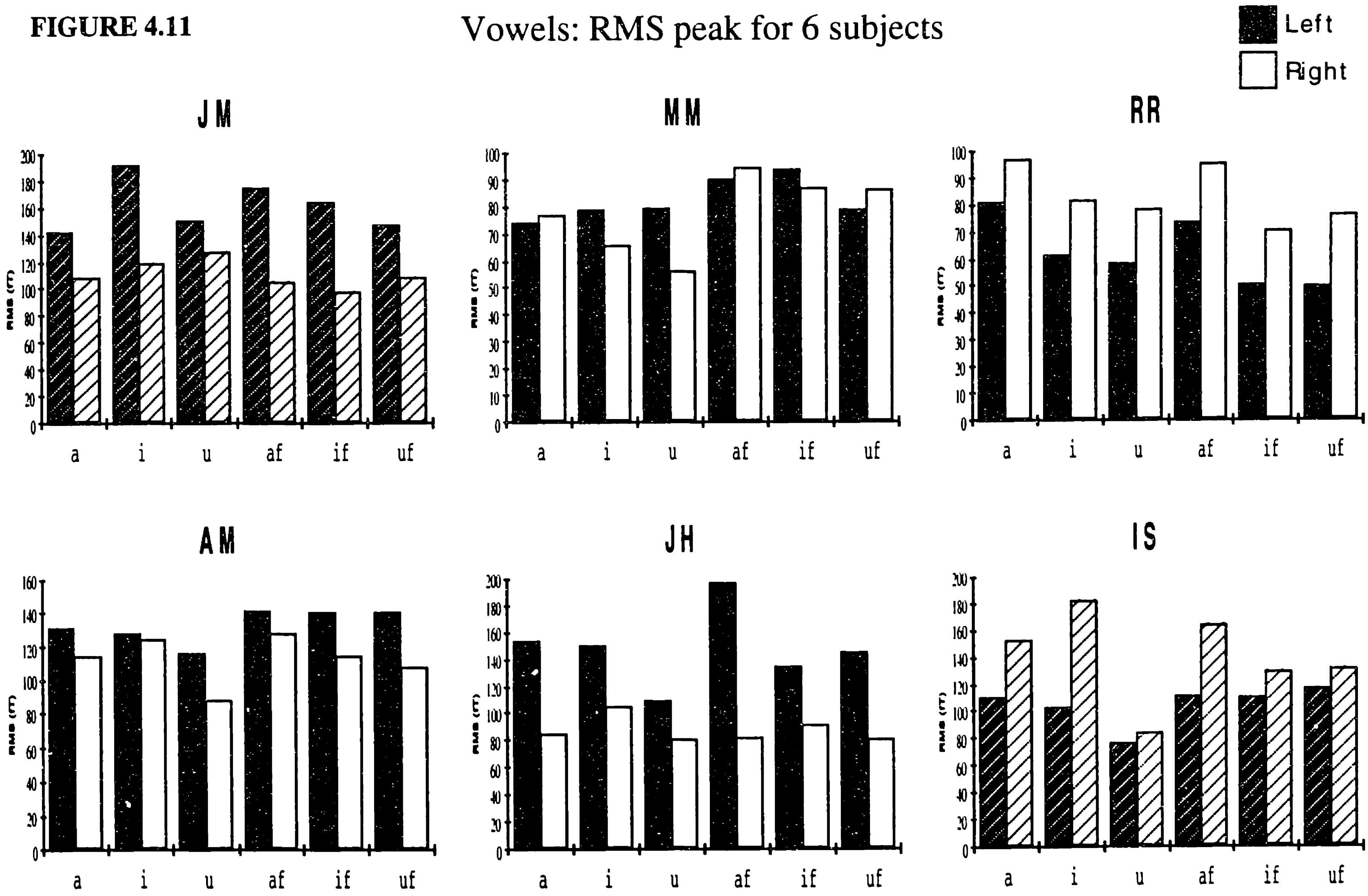


**Figure 4.12**

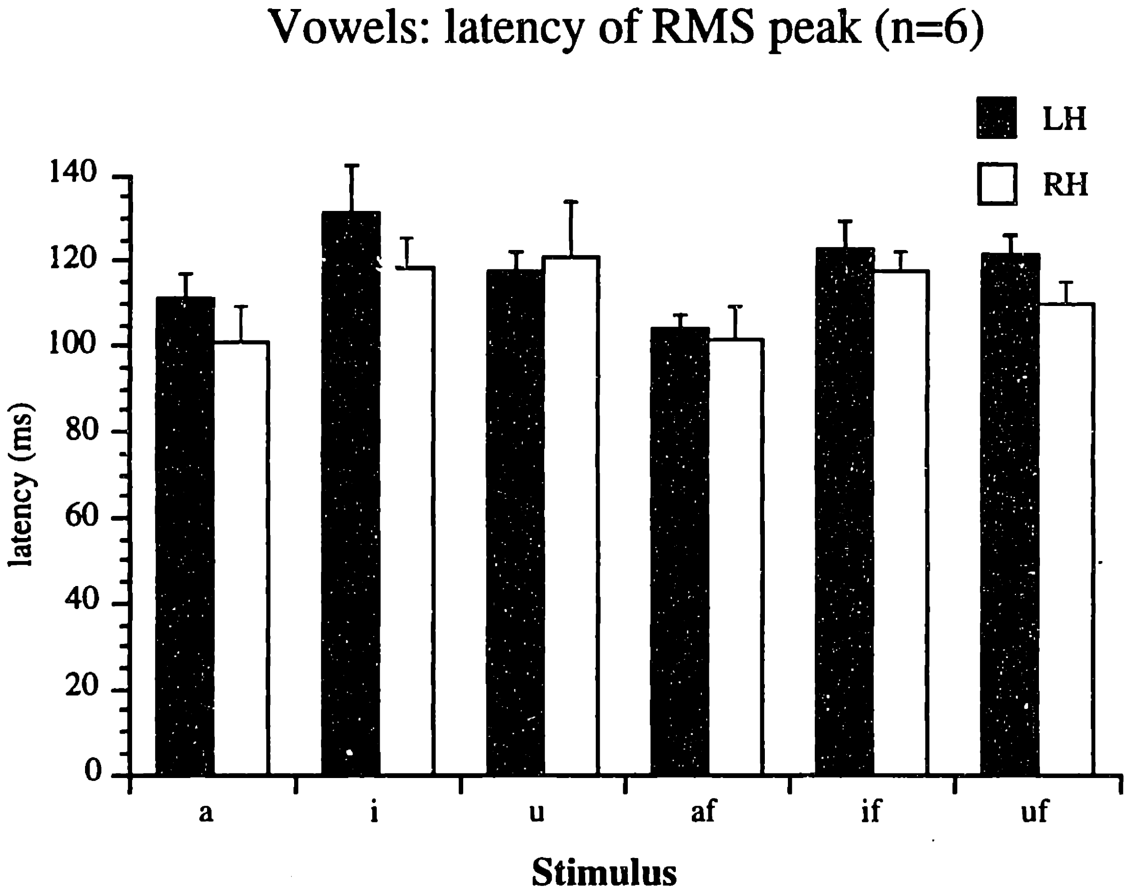

Vowels: latency of RMS peak (n=6)
LH
RH
140
120
100
80
60
40
20
0
latency (ms)
a
i
u
af
if
uf
Stimulus

FIGURE 4.13

Vowels: latency of RMS peak for 6 subjects

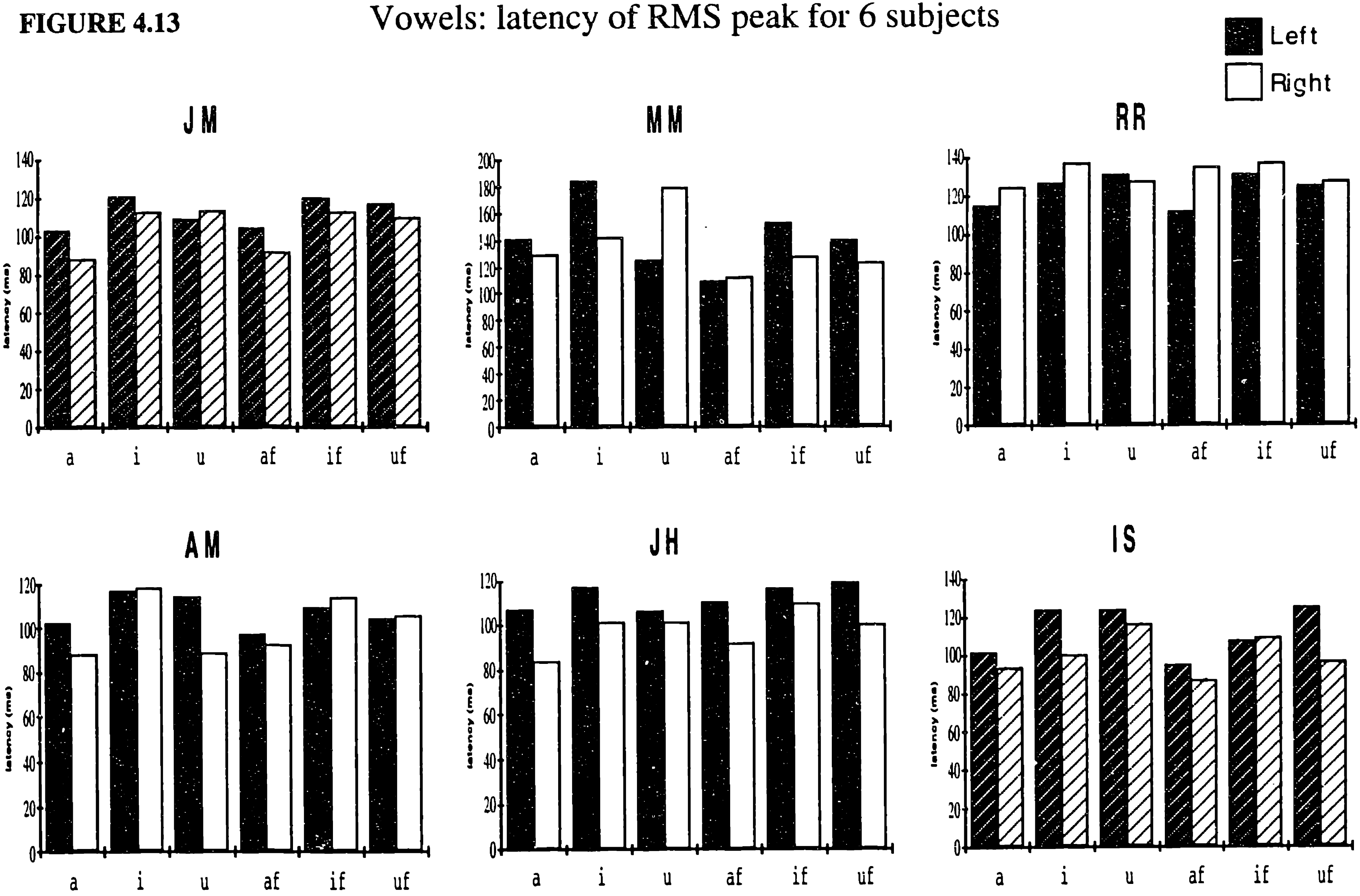


**FIGURE 4.14** Single dipole for vowel /a/(F0 = 100Hz or 200Hz) at the M100 RMS peak in both hemispheres (3 Ss.)

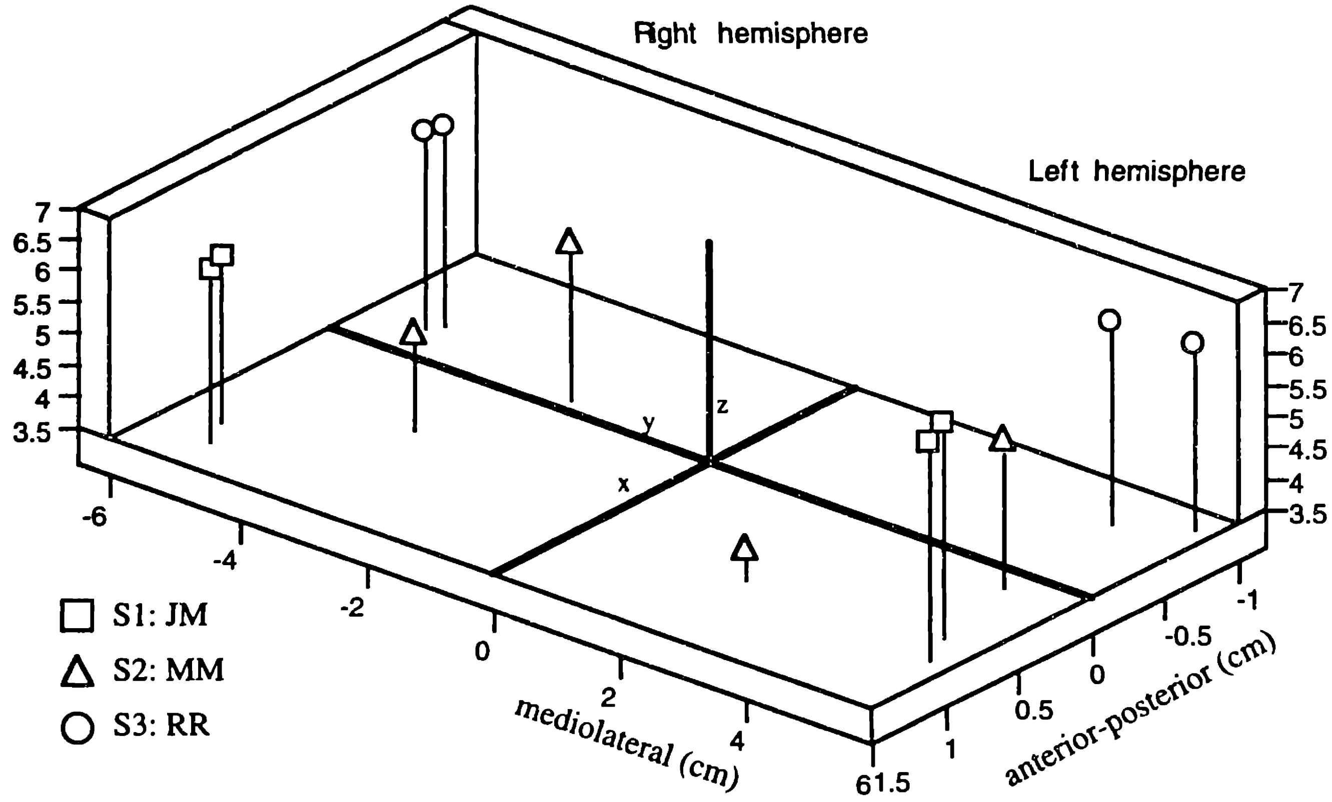


**FIGURE 4.15** Single dipole for vowel /a/(F0 = 100Hz or 200Hz) at the M100 RMS peak in both hemispheres (3 male Ss.)

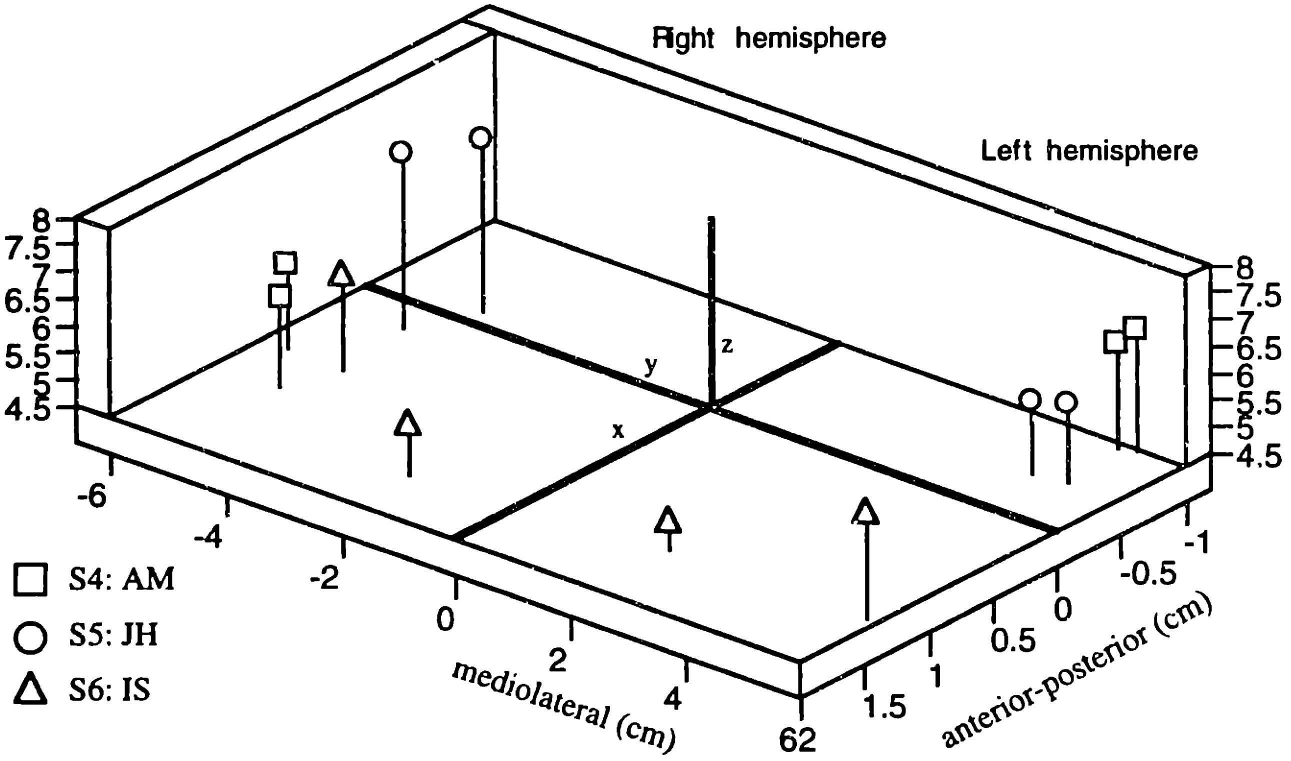
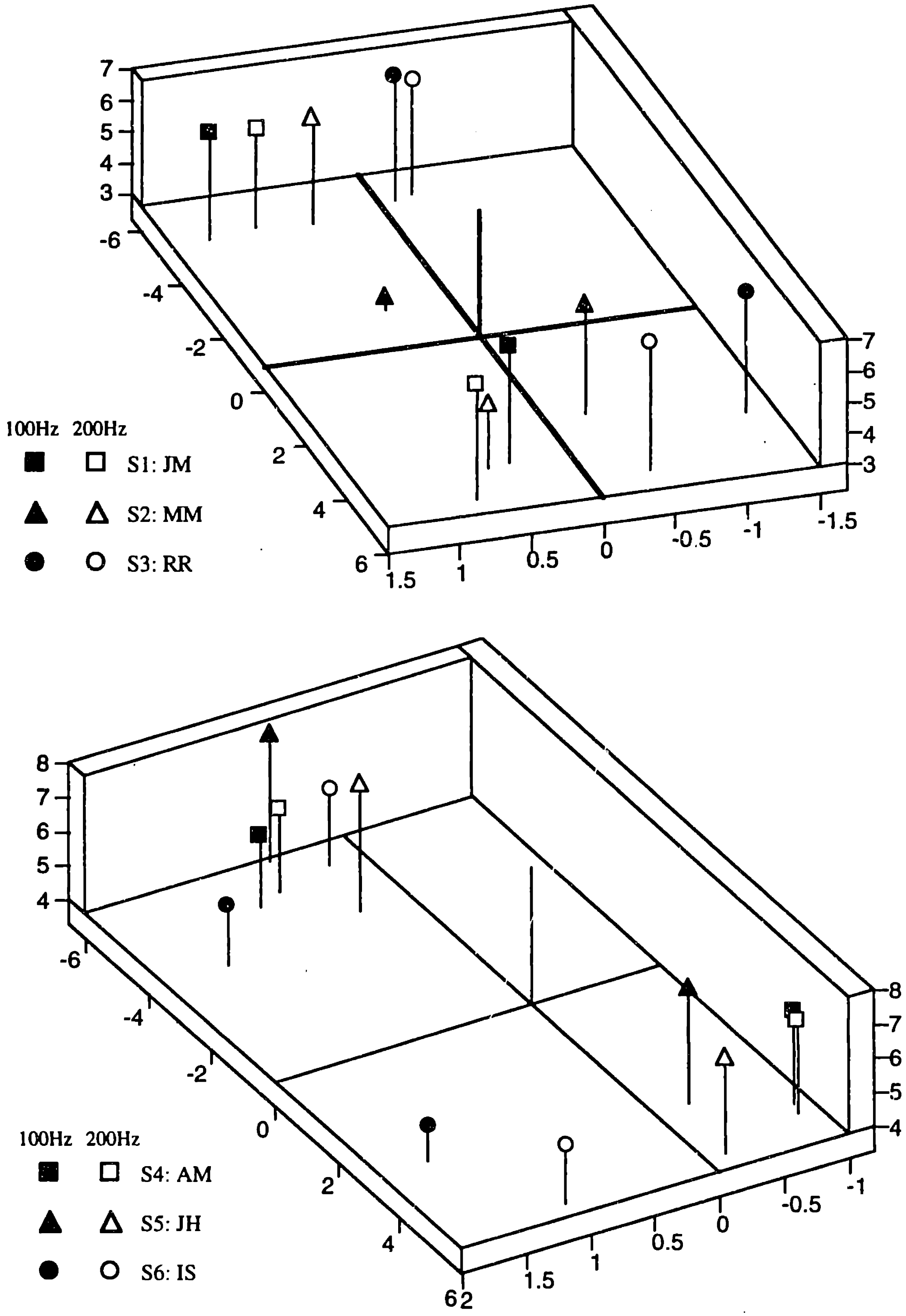
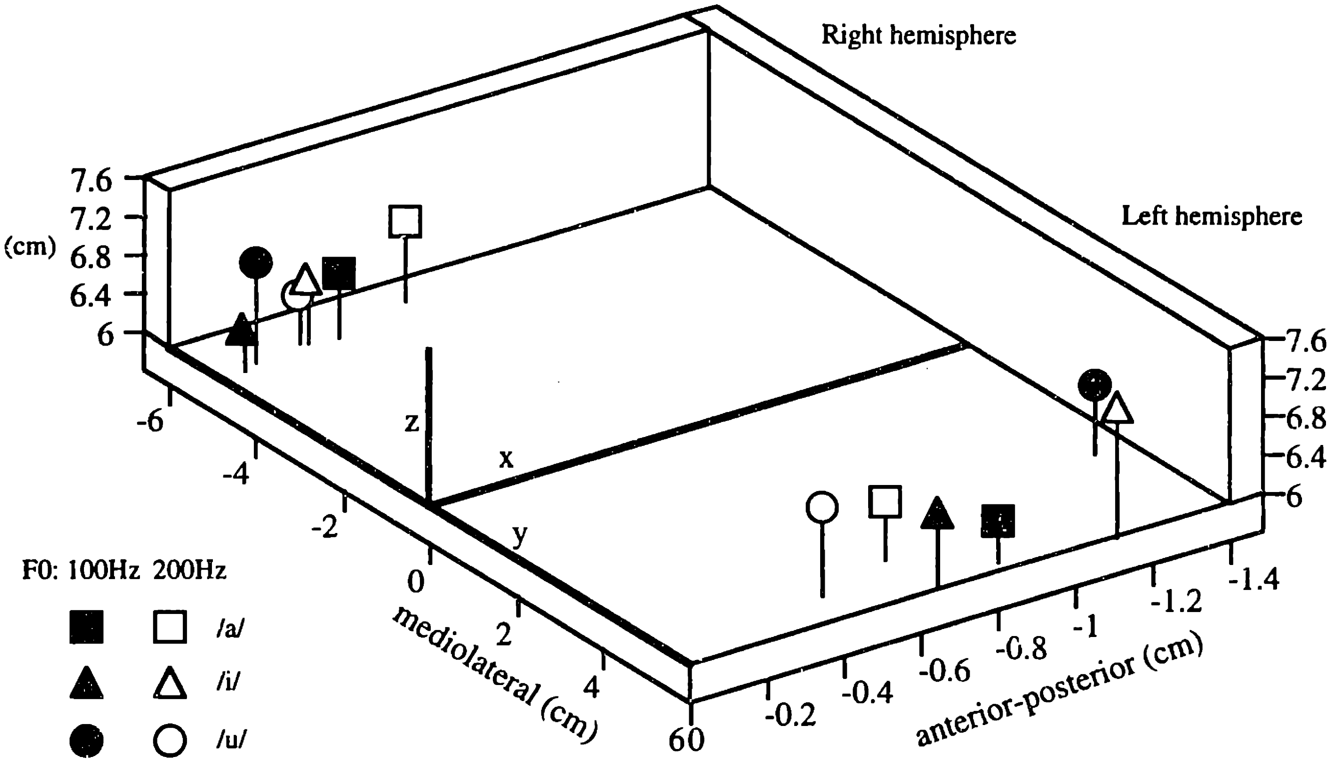


**Figure 4.16 Single dipoles for vowel /u/**



## Figure 4.17 Single dipoles for the three vowels /a i u/ at two F0 frequencies in one subject (S3: RR)



## Figure 4.18 Medio-lateral variation of single dipole around M100 (S4: AM, left hemisphere)

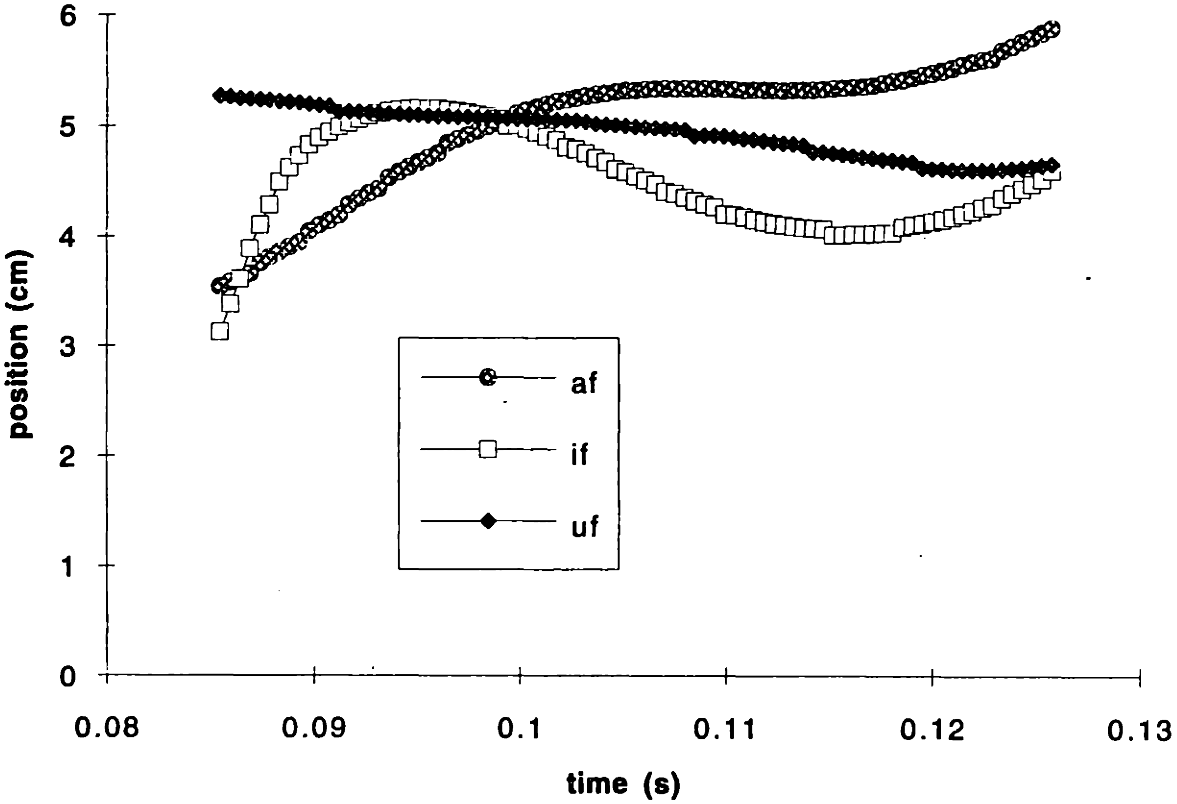
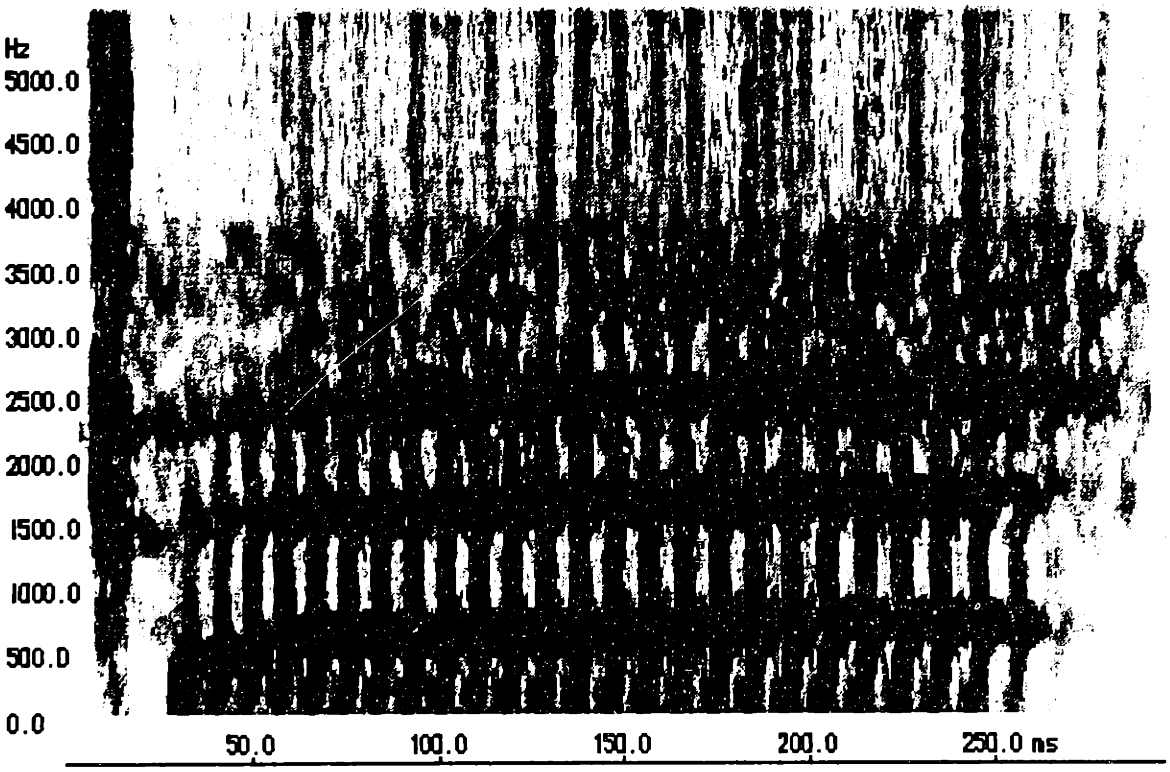


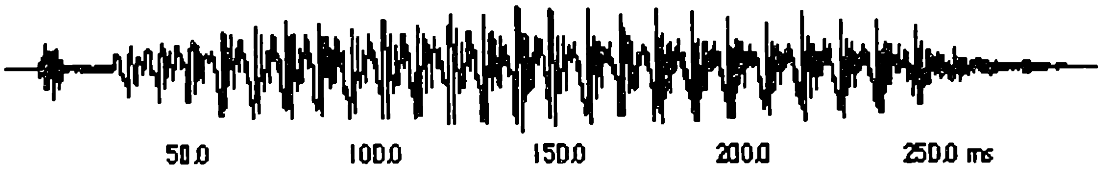
**FIGURE 4.19**

Syllable [bæ], 20ms VOT

Spectrogram



Waveform



Spectrum

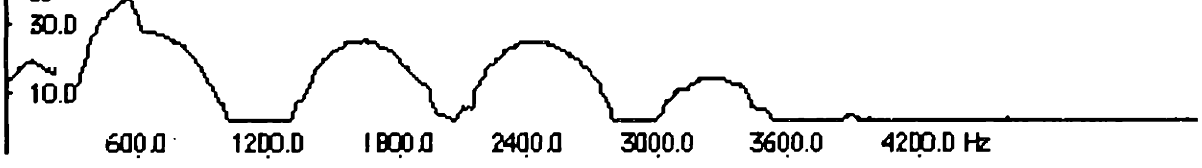
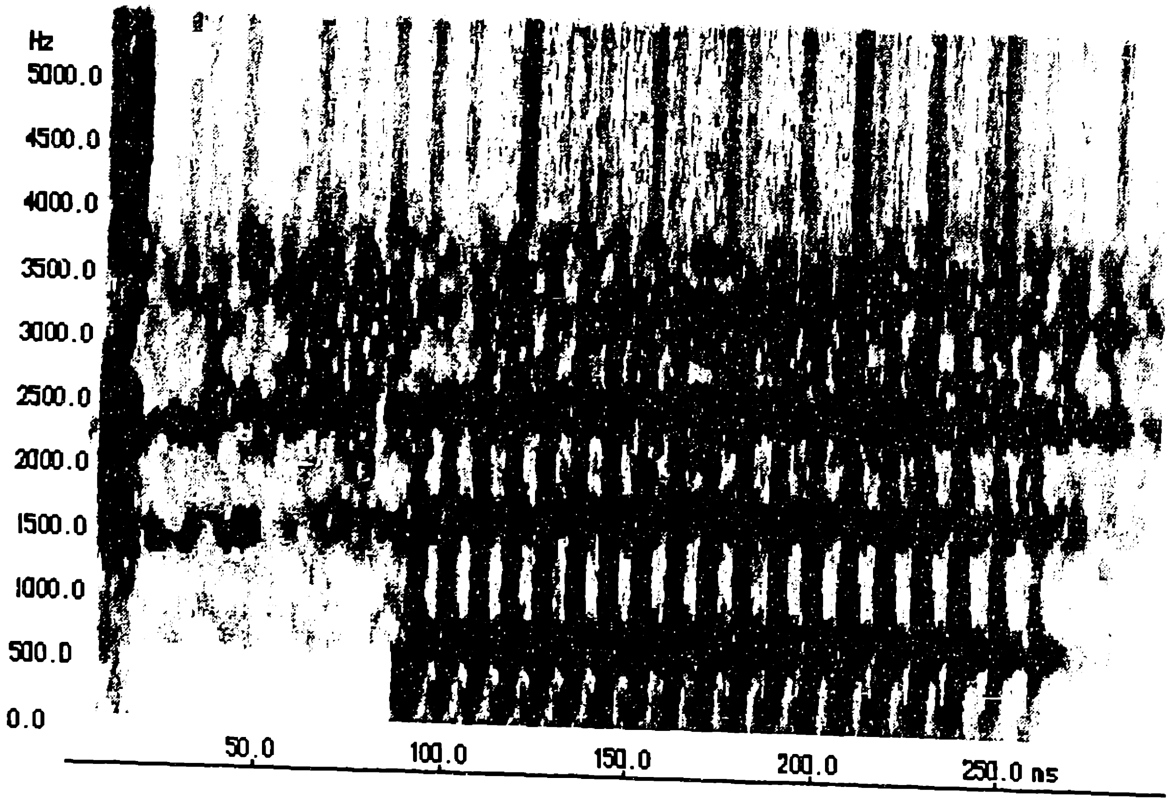


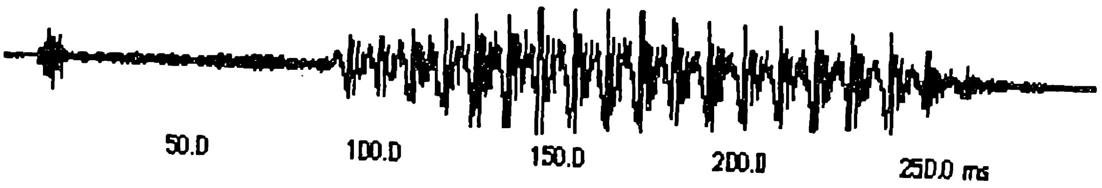
**FIGURE 4.20**

Syllable [pæ], 80ms VOT

Spectrogram



Waveform



Spectrum

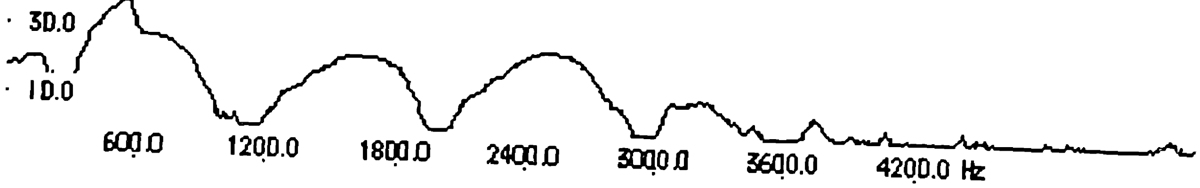
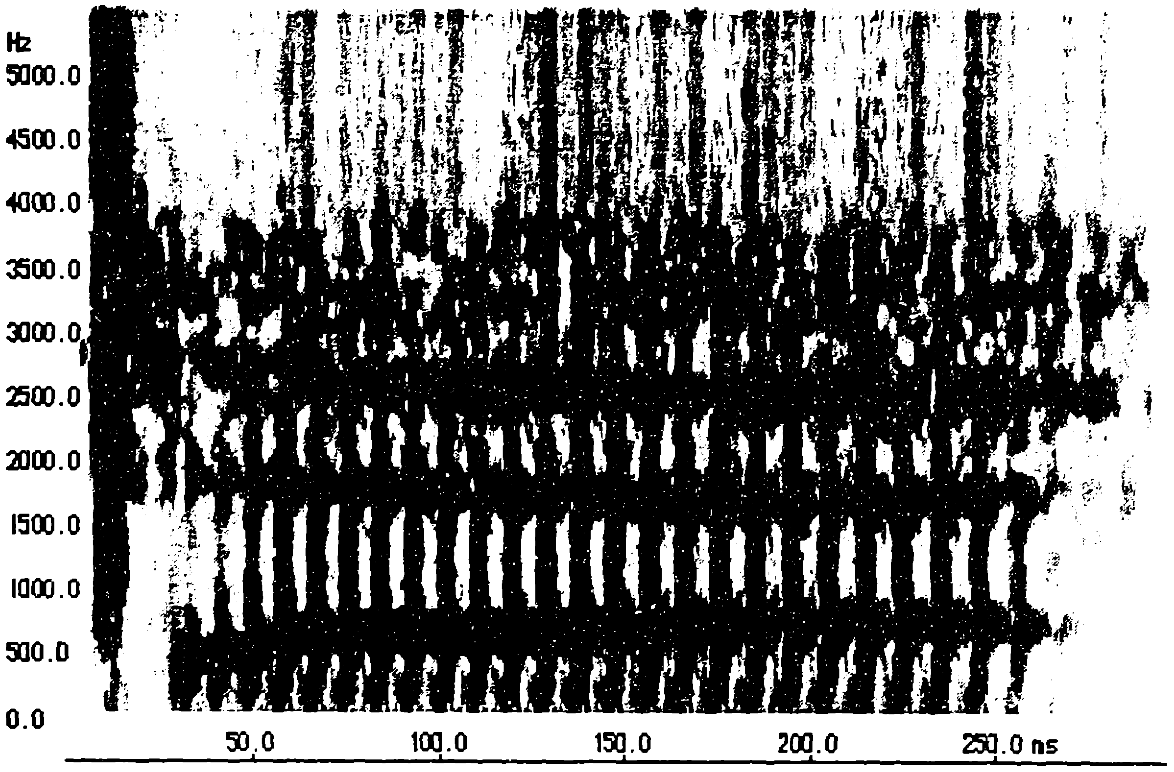


**FIGURE 4.21**

Syllable [dæ], 20ms VOT

Spectrogram



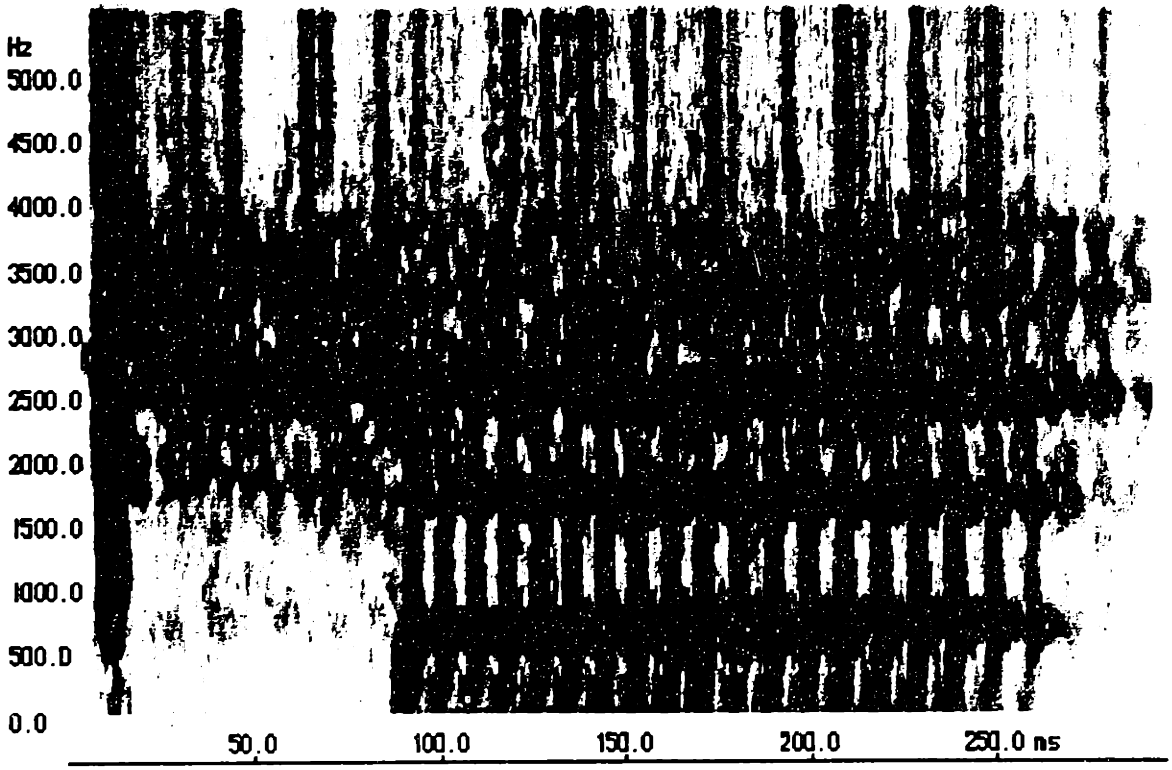
Waveform

Spectrum

**FIGURE 4.22**

Syllable [tæ], 80ms VOT

Spectrogram



Waveform

Spectrum

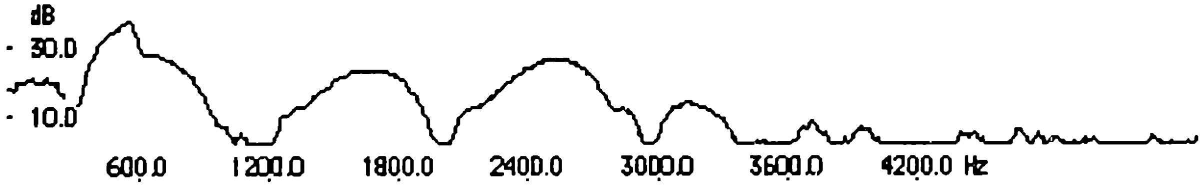


**Figure 4.23**

Passive BPDT: latency of RMS peak

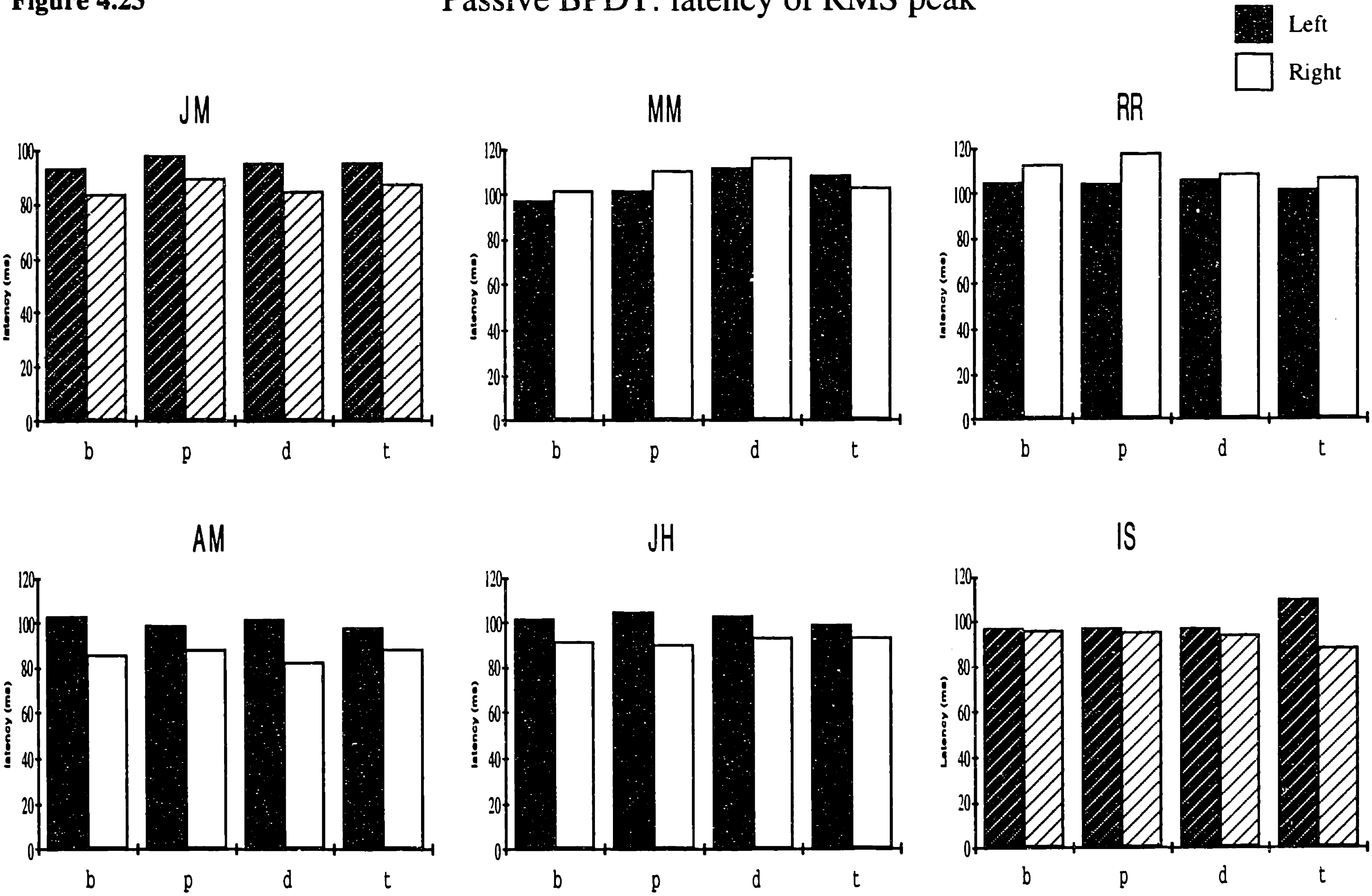


**Figure 4.24**

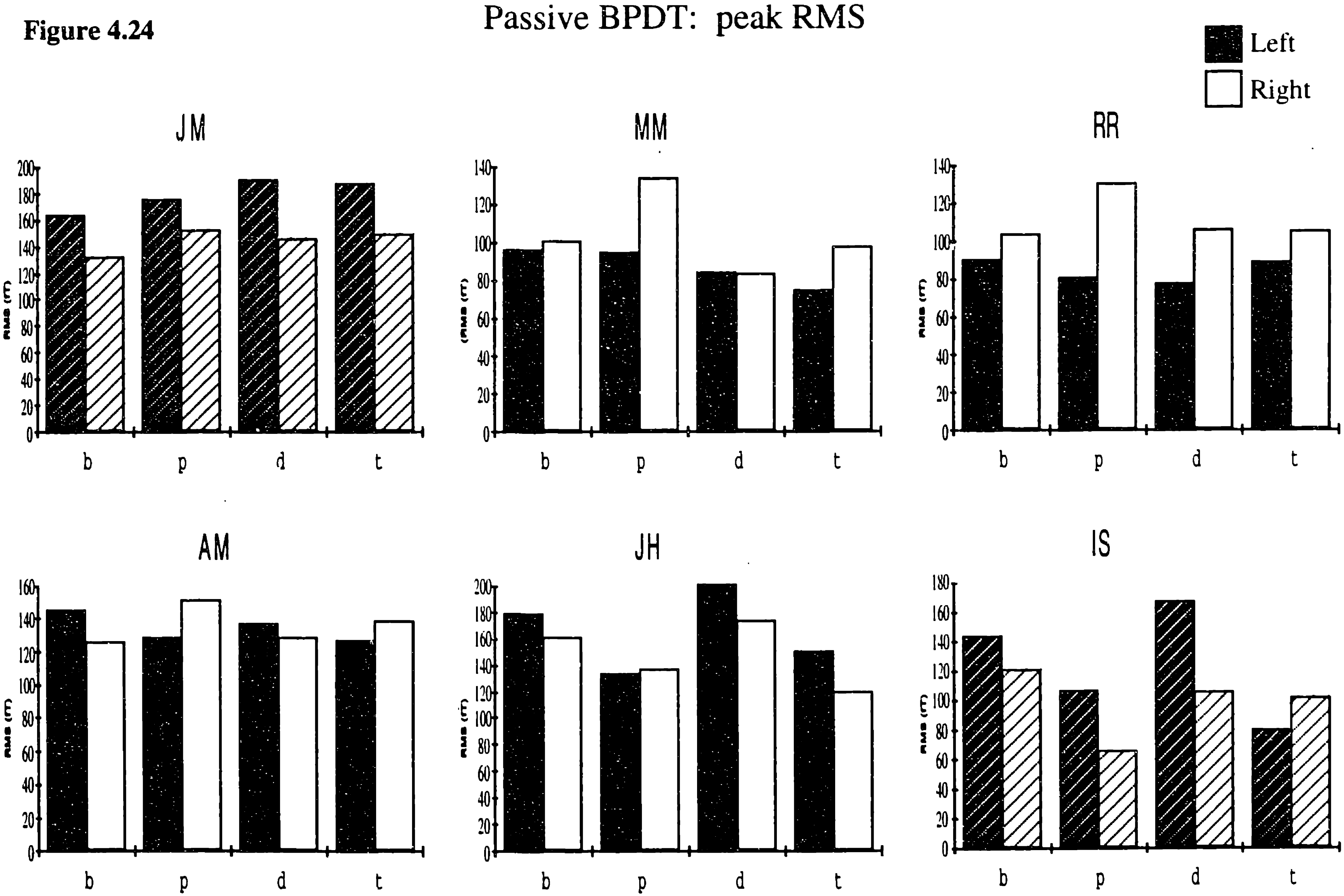

Passive BPDT: peak RMS

**FIGURE 4.25**

Reaction times in CV discrimination task (n=6)

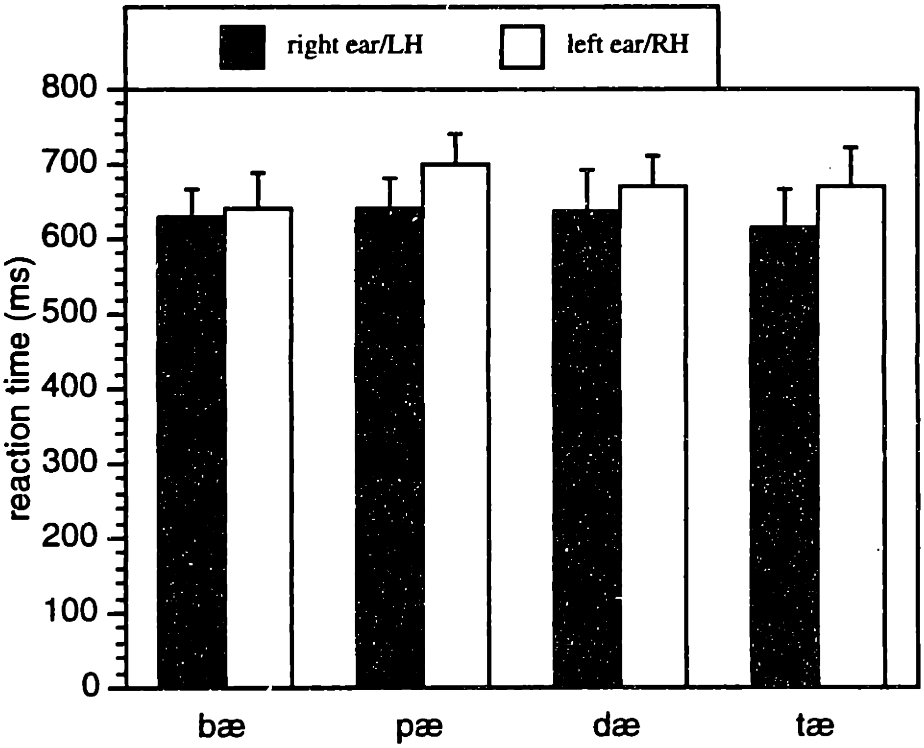


**Figure 4.26**

BPDT discrimination: peak RMS

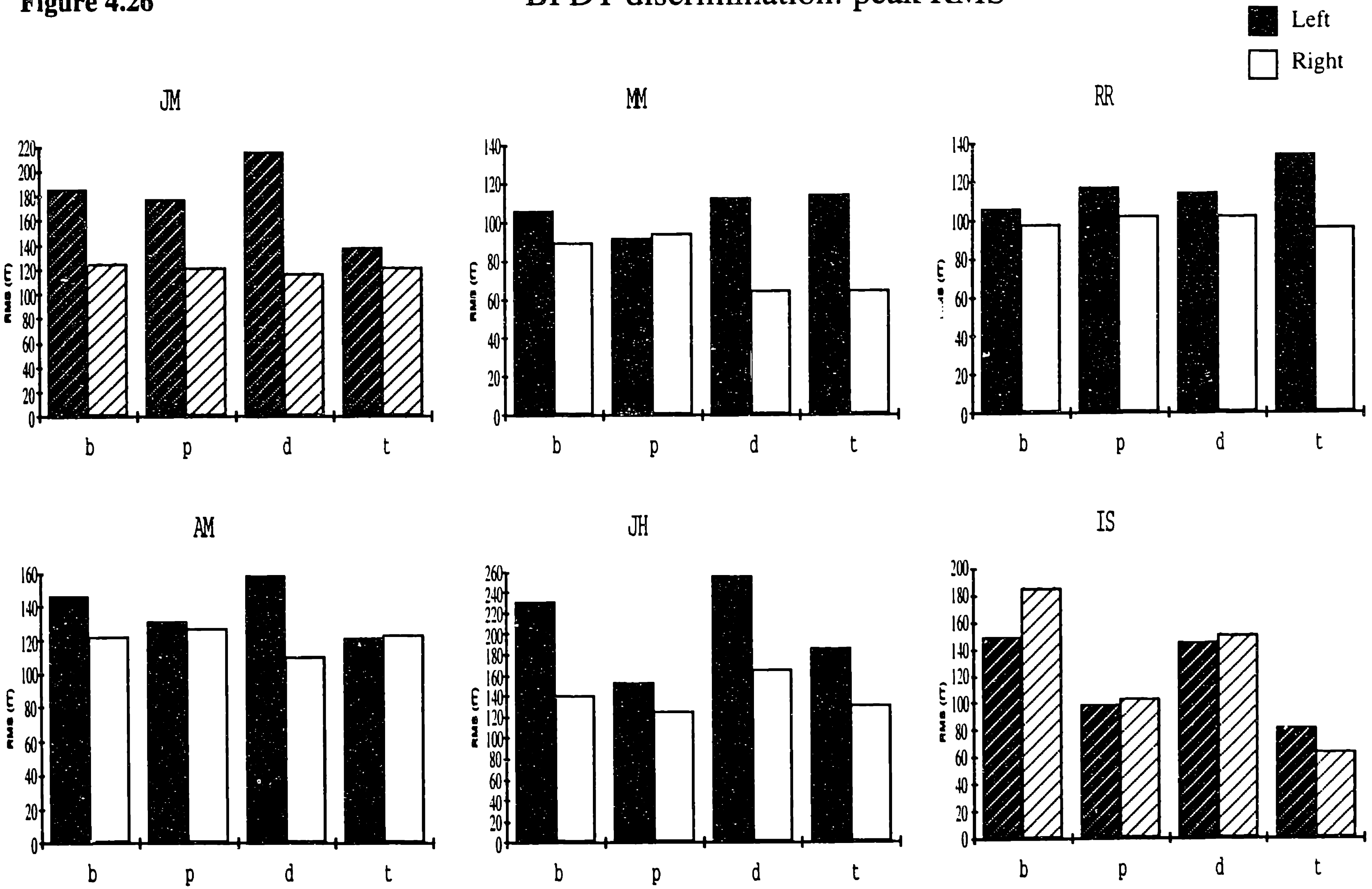


Figure 4.27 RMS difference values (task -passive)

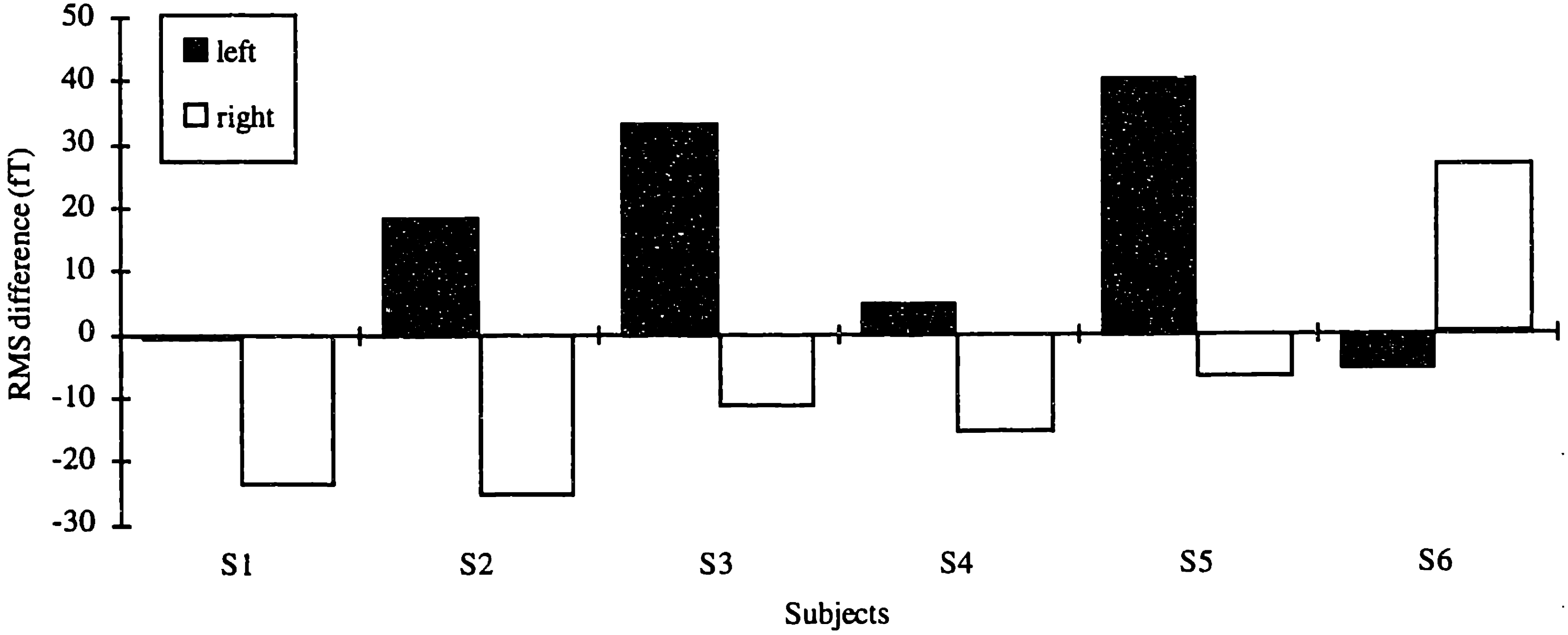


**Figure 4.28 Single equivalent current dipole location for syllables (passive & discriminate conditions) for one subject (AM)**

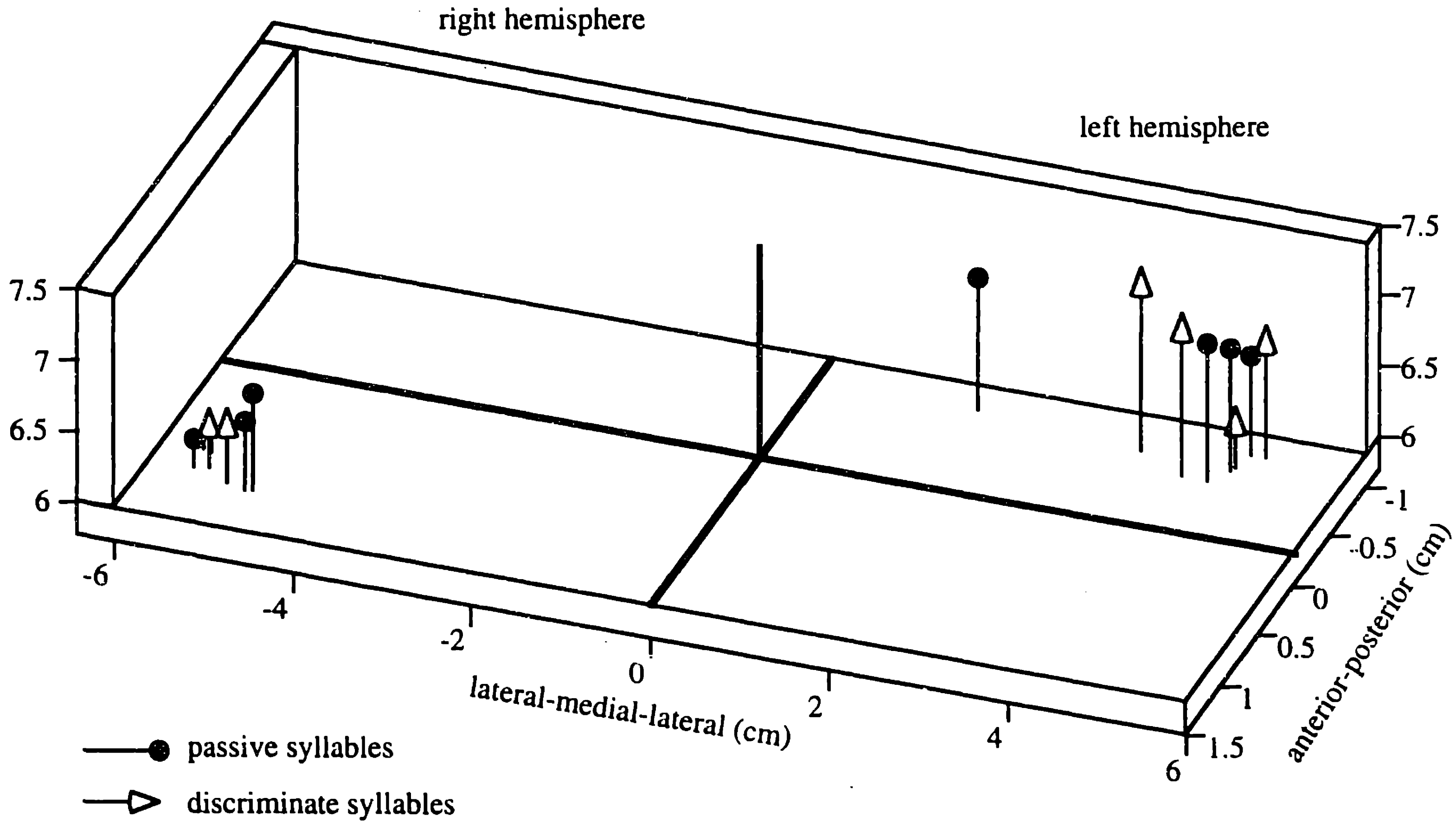


**Figure 4.29 Single equivalent current dipole location for syllables (passive & discriminate conditions) for one subject (JH)**

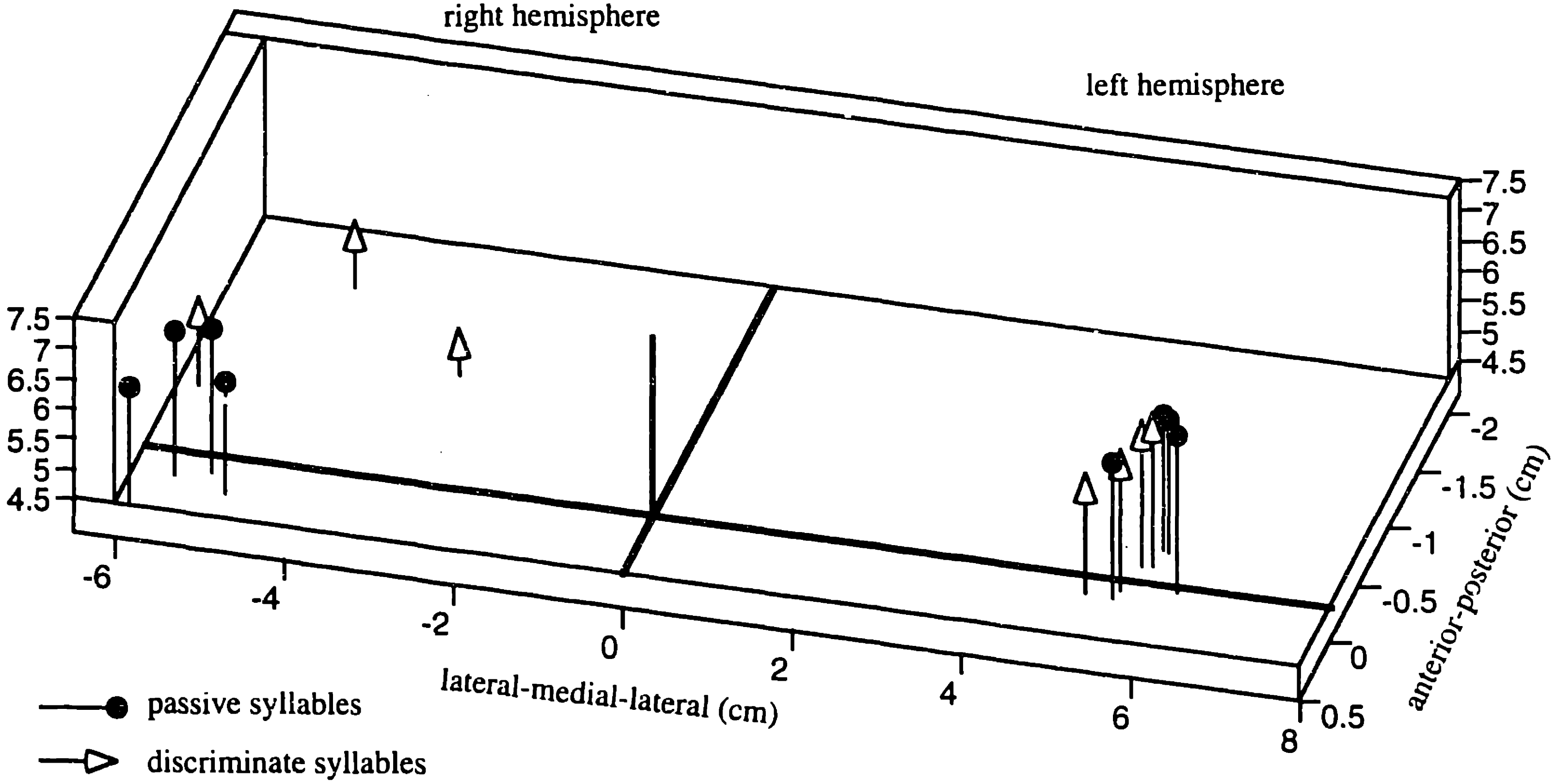


**FIGURE 4.30** **Syllable dipoles for subject IS**

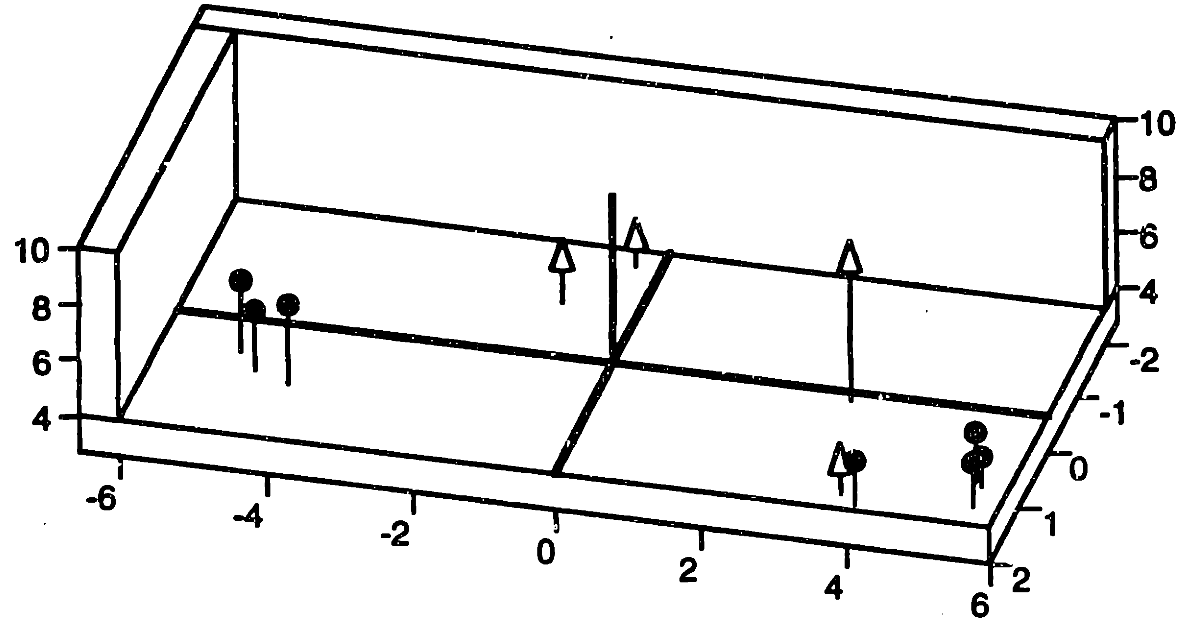


**FIGURE 4.31** **Syllable dipoles for subject JM**

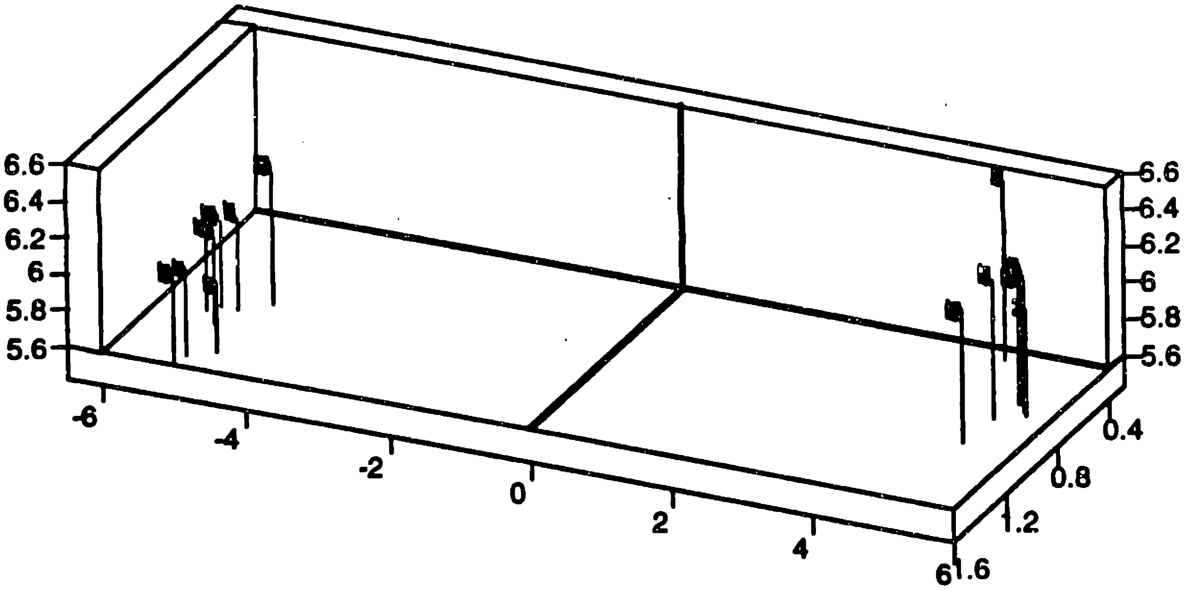


FIGURE 4.32 Syllable dipoles for subject MM

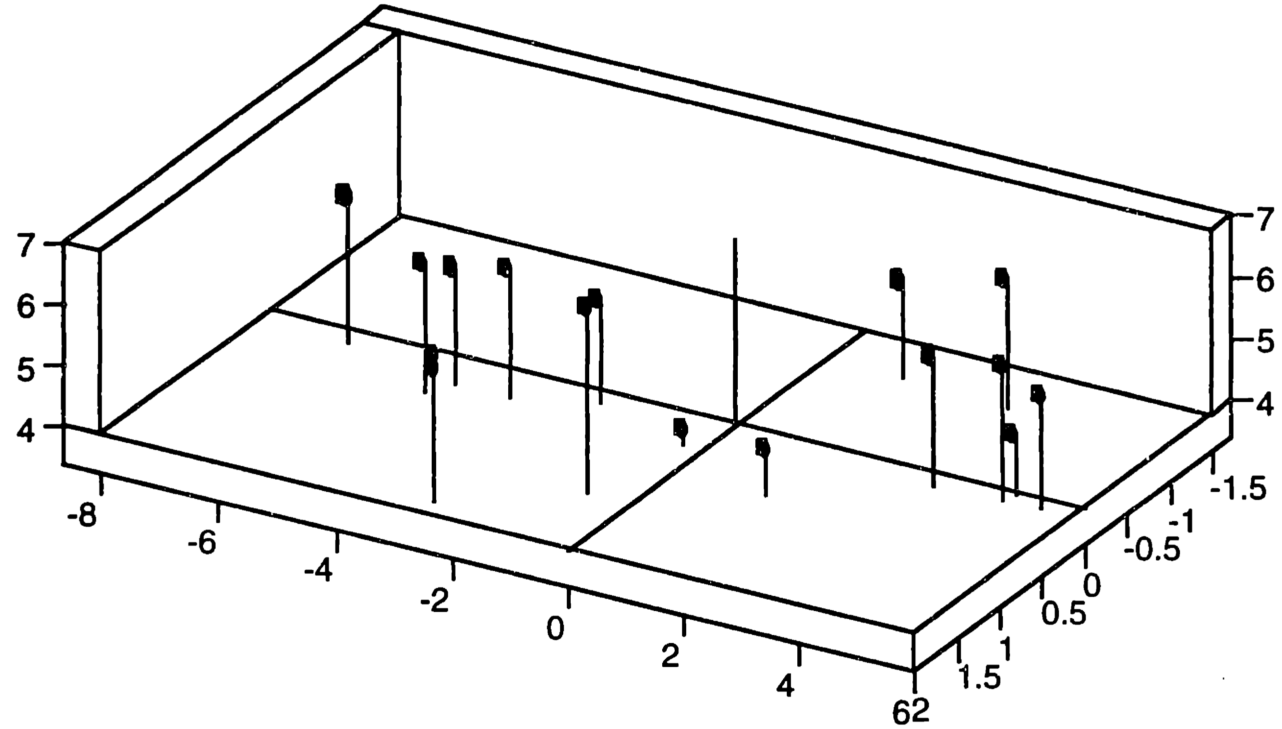
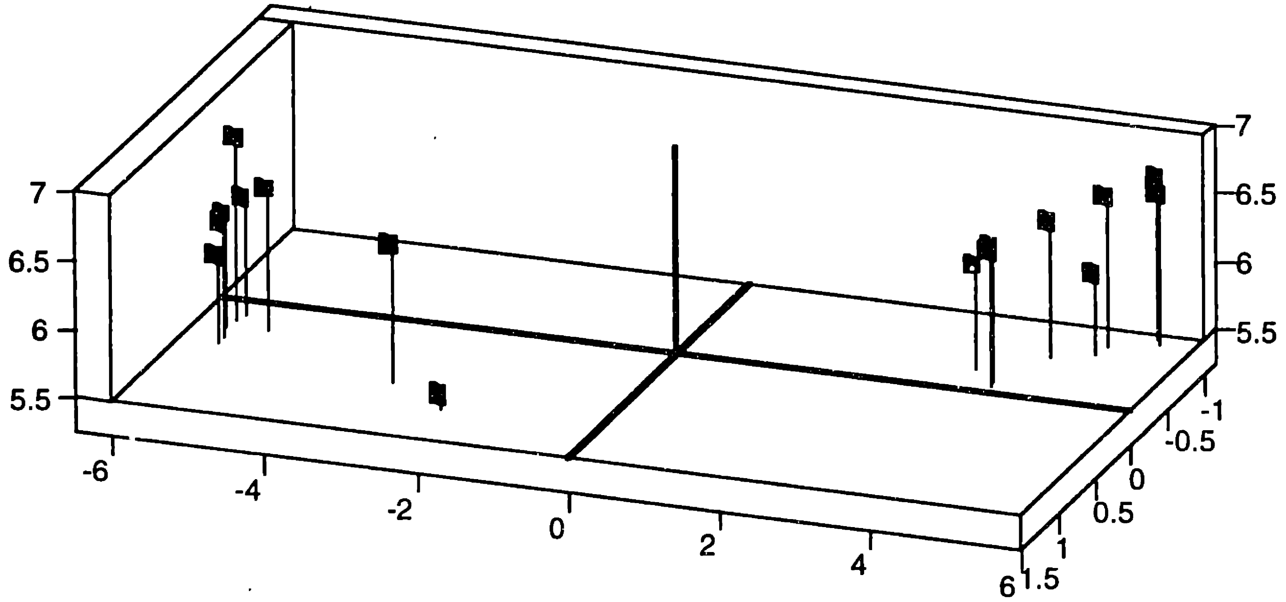


FIGURE 4.33 Syllable dipoles for subject RR



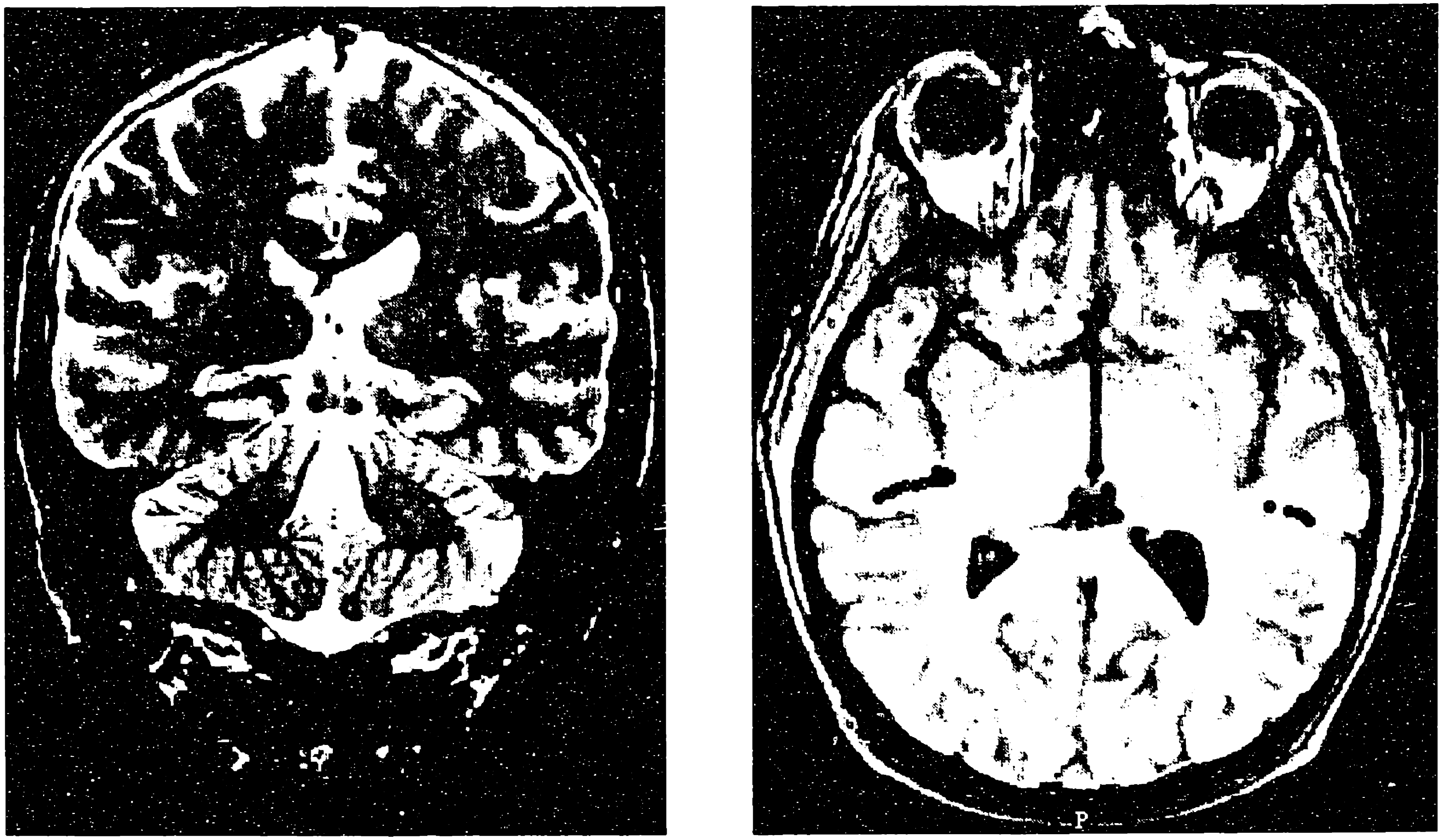


FIGURE 4.34 Dipole trajectory MR overlay, pure tones (Subject JH)

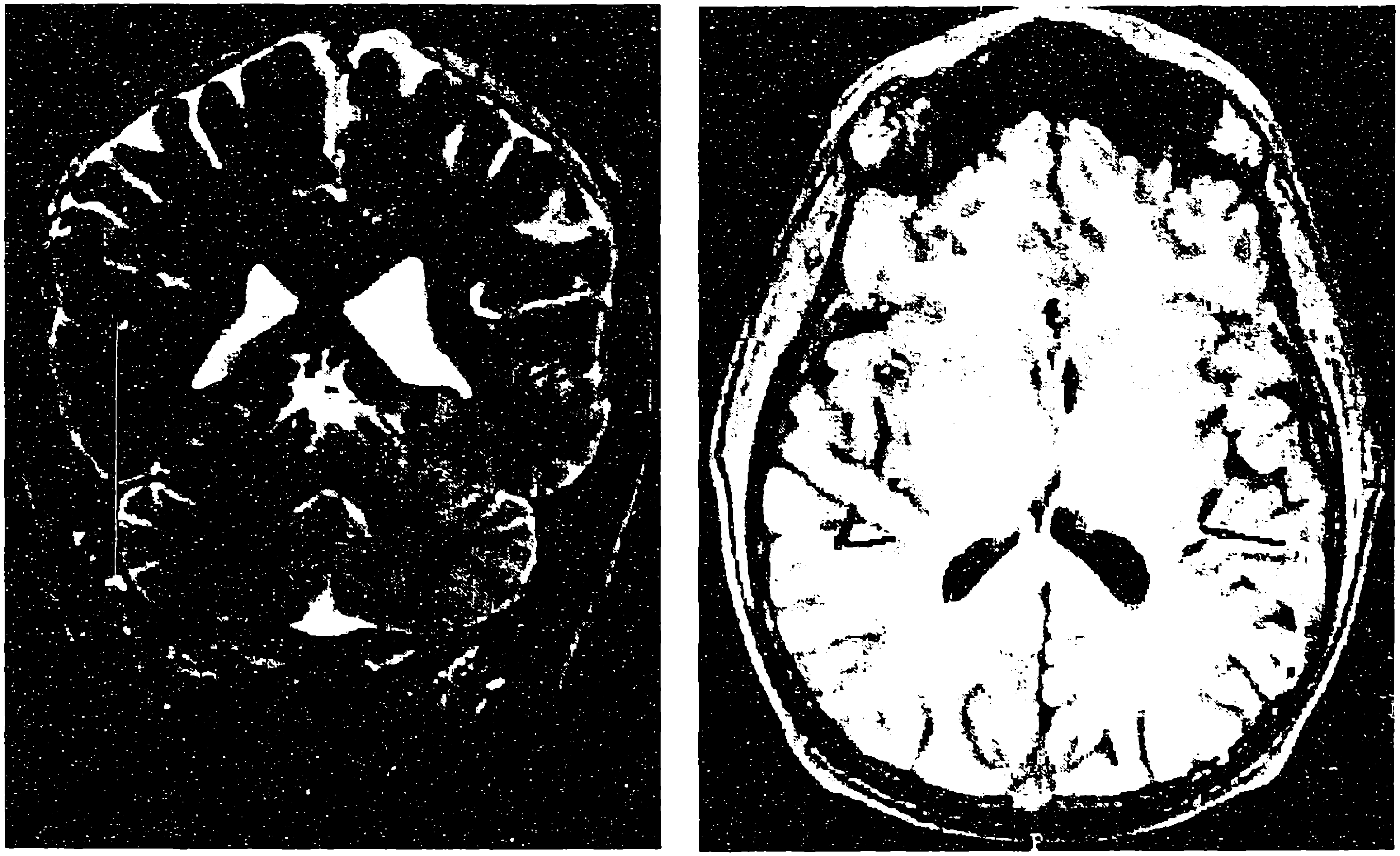


FIGURE 4.35 Dipole trajectory MR overlay, vowels /a/ (100 & 200 Hz F0) (Subject JH)

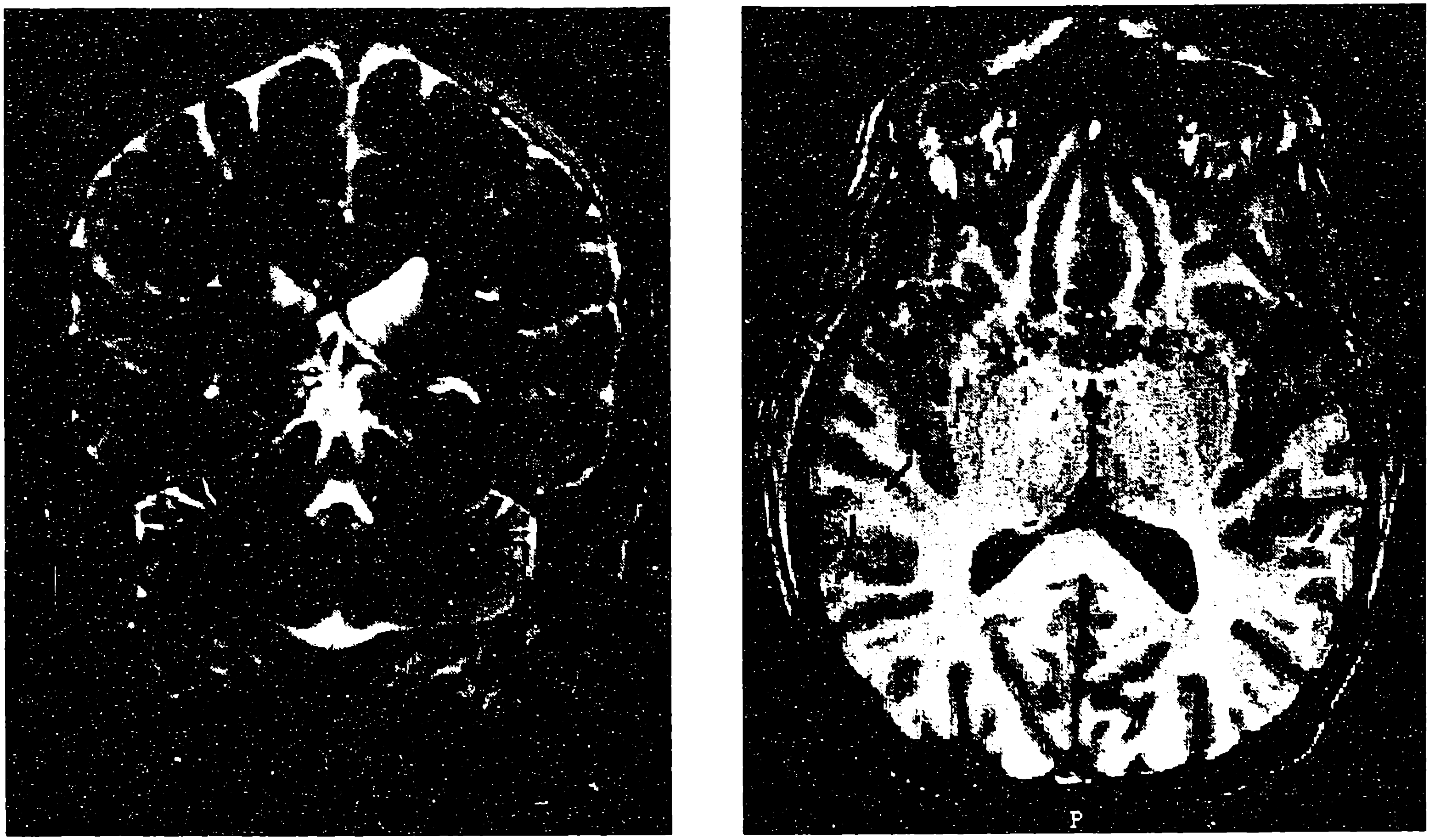


FIGURE 4.36 Dipole trajectory MR overlay, passive syllables (Subject JH)

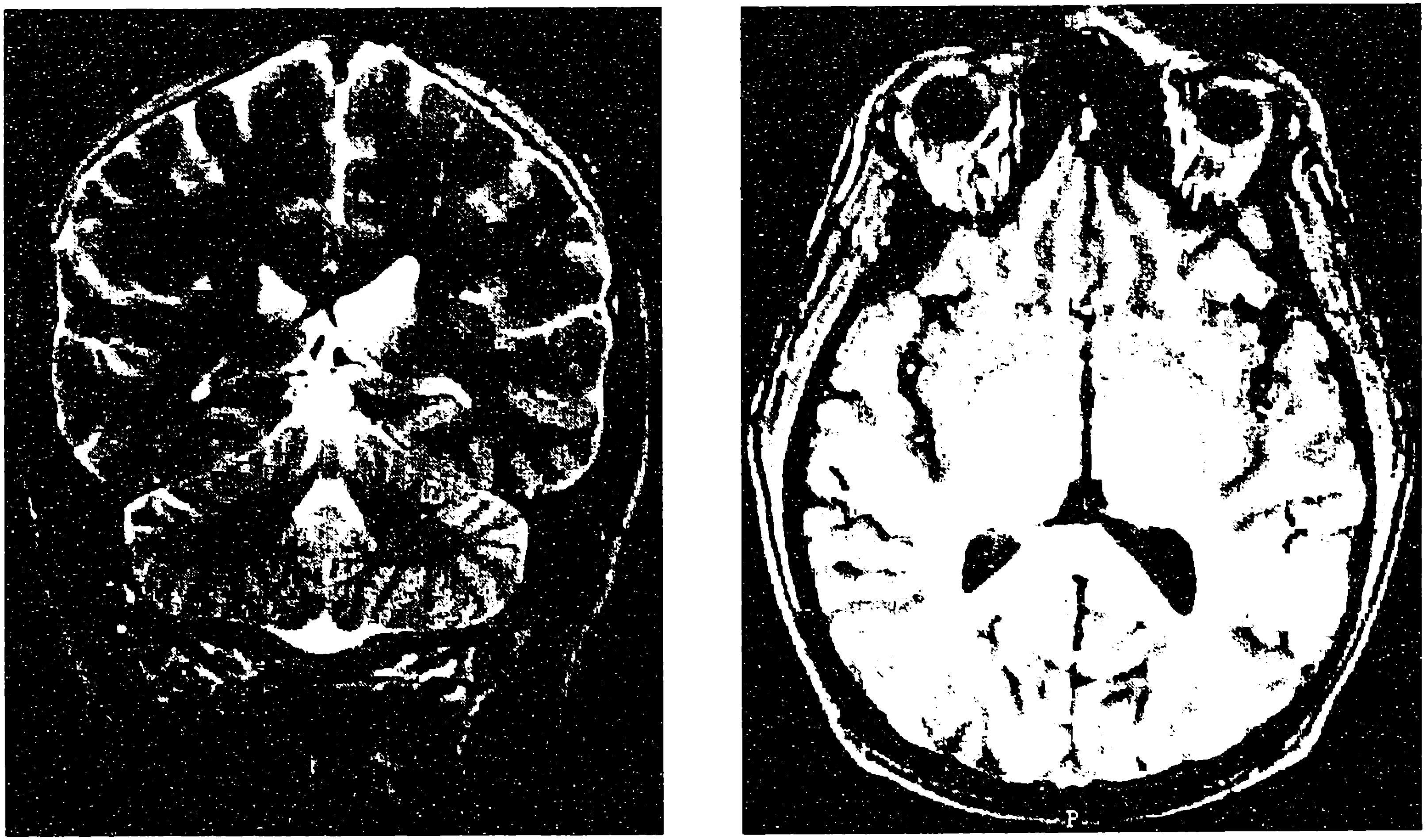


FIGURE 4.37 Dipole trajectory MR overlay, discriminated syllables (Subject JH)

## Figure Legends

FIGURE 4.1 The SQUID detectors, in this example a first-order gradiometer, are in a liquid-helium filled cryogenic dewar. Magnetic fields generated by intracranial sources are picked up by the detectors if the fields are orthogonal to the skull surface (the current dipole generating the field must be tangential to the head). The dewar can be positioned to record over the appropriate part of the head.

FIGURE 4.2 *Time-varying evoked magnetic field recorded in five channels.*

Time-varying evoked neuromagnetic field recorded from five (labeled) channels over the left temporal lobe in response to a vowel (/a/). The entire 600ms epoch is shown. Vertical display: 500fT/division. Note the polarity difference between the upper three and lower two channels, particularly prominent around 110ms (M100 or N1m response). The arrow marks the stimulus onset (300ms duration vowel).

FIGURE 4.3 *Sensor layout displays and contour map*

The auditory evoked magnetic field from one subject, elicited by the presentation of a vowel, is illustrated from four different perspectives.

**(a)** Sensor layout display of the time-varying evoked neuromagnetic field recorded simultaneously in 37 channels. The data shown are from a *single epoch* (i.e the evoked field from one presentation of the stimulus), recorded over the left auditory cortex of one subject, the same epoch from which five channels are displayed in Figure 4.2. The stimulus was the vowel /a/, with a fundamental frequency of 100Hz (male pitch) and a duration of 300ms. The horizontal scale is approximately 240ms/cm. The vertical display scale was set to 1pT/division (equivalent to 1000fT/div). The vertical marker indicates the stimulus onset.

**(b)** Sensor layout display of the *average* of 100 epochs collected from the same subject and the same dataset of evoked responses. The horizontal scale is as in (a). The vertical display scale is 500 fT/division. The vertical marker indicates the stimulus onset. The sensors on the upper right and lower left illustrate the prominent evoked response components. Large field deflections (of opposite polarities) at about 50ms, 100ms, and 200ms following stimulus onset are clear. These deflections represent the M50 (P50), M100 (N1m), and M200 (P2) classical auditory evoked response components. Although the evoked response components are visible in a single epoch, as seen in (a), signal averaging optimizes the signal-to-noise ratio.

**(c)** Sensor layout display of the data shown in (b), *low-pass filtered* from 1-20Hz (the data were acquired with a 1.0Hz high-pass cutoff). The evoked response to 100 presentations of /a/ was averaged, (b), and low-pass filtered (20dB attenuation at 25Hz). The vertical scale is 200fT/division.

**(d)** Isofield *contour map* for the data shown in (c) is plotted at a latency of 108ms (M100 peak). The contour interval is 10fT and the scale is 500fT. On one side (of the thick dividing line), the field lines emerge from the skull, on the other side they reenter the skull. The source parameters of the single dipole hypothesized to generate the field can be estimated by accounting for the changing field strength and orientation (in versus out) as visualized by a contour map.

Figure 4.4 Single dipoles computed for three pure tones presented to the contralateral ear overlaid on a subject's structural MR image. The dipoles were estimated to be in auditory association cortex. This represents the combination of MRI data and MEG data (Magnetic Source Imaging).

Figure 4.5 *M100 RMS peak data for six subjects over both hemispheres.* The stimulus was a 500Hz sinusoidal tone of 400ms duration. Black bars represent left hemisphere (LH) recordings, white bars represent right hemisphere (RH) recordings. The hatched bars are the data for the two left-handed subjects.

Figure 4.6 M100 latency for six subjects. Key as in Figure 4.5.

Figure 4.7 Single ECD for 500Hz tone at the M100 RMS peak measured independently in both hemispheres (6 subjects). Each plotted point is the single dipole chosen as representative for the tone condition for a given subject (6 subjects x 2 hemispheres). The positions of the dipoles are consistent with being on the supratemporal plane.

Figure 4.8 3 representations of the synthesized vowel /a/ with a 100Hz fundamental frequency F0. The *spectrogram* plots time (ms) over frequency (Hz). The dark horizontal bands are the formants (cf. Table 4.6), i.e. the frequency bands of increased energy. The vertical stripes reflect the glottal pulse (fundamental frequency) of 100Hz. The *waveform* depicts the overall stimulus envelope and crisply shows F0 (long vertical lines). The *spectrum* (Fast Fourier Transform) of the stimulus plots the frequency (abscissa) against amplitude (in dB on ordinate) and illustrates the energies at the formants (here: F1>F2>F3).

Figure 4.9 3 representations of the synthesized vowel /a/ with a 200Hz fundamental frequency F0. See legend for Figure 4.8.

Figure 4.10 Vowels: M100 RMS peak group data. LH = left hemisphere; RH = right hemisphere. a, i, and u: F0=100Hz. af, if, uf: F0=200Hz.

Figure 4.11 Vowels: M100 latency data for individual subjects and each stimulus. Subjects' initials are above each plot. The data for the two left-handed subjects are hatched. Abscissa: stimulus type (vowel); ordinate: RMS of field strength in femtoTesla (fT). The plot for each subject represents 1200 trial (2 hemispheres x 6 vowels x 100 trials/vowel).

Figure 4.12 Vowels: M100 latency of RMS peak group data. Legends as in Figure 4.10.

Figure 4.13 Vowels: M100 latencies for individual subjects and stimuli. Legends as in Figure 4.11.

Figure 4.14 Single dipole for two types of vowel /a/ (Fo=100Hz or 200Hz) at the M100 RMS peak in both hemispheres of the 3 female subjects (S1-S3). Each plotted point represents the single M100 dipole (see Methods for choice of single dipole) computed for the subject and stimulus in either the left or right hemispheres. Filled symbols are male F0 vowels, and open symbols are female F0 vowels. The square, triangle, and circle represent the 3 different subjects.

Figure 4.15 Single dipole for two types of vowel /a/ (Fo=100Hz or 200Hz) at the M100 RMS peak in both hemispheres of the 3 male subjects (S4-S6). Legend as in Figure 4.14.

Figure 4.16 Single dipoles for the vowel /u/. See Figure 4.14 for legend. All scales are as in Figure 4.14.

Figure 4.17 Single dipoles for the three vowel types at the two different F0 frequencies for one subject (S2: RR). Filled symbols represent F0=100Hz (male vowel) and open symbols represent F0=200Hz (female vowel). The squares, triangles, and circles represent the dipoles for a, i, and u respectively. The localizations are to auditory association cortex.

Figure 4.18 Variation in y-coordinate (mediolateral) of a single dipole computed for the passive presentation of the vowel /a/. The trajectories for /i/ and /u/ cross approximately at the RMS peak (100ms).

Figure 4.19 Spectrogram, waveform, and power spectrum for syllable /bæ/, with voice-onset time (VOT) 20ms. See legend for Figure 4.8.

Figure 4.20 Spectrogram, waveform, and power spectrum for syllable /pæ/, with voice-onset time (VOT) 80ms. See legend for Figure 4.8.

Figure 4.21 Spectrogram, waveform, and power spectrum for syllable /dæ/, with voice-onset time (VOT) 20ms. See legend for Figure 4.8.

Figure 4.22 Spectrogram, waveform, and power spectrum for syllable /tæ/, with voice-onset time (VOT) 20ms. See legend for Figure 4.8.

Figure 4.23 Passive syllables: M100 latency data for each subject and each stimulus. Hatched columns for the two left-handed subjects. Legend as in Figure 4.11.

Figure 4.24 Passive syllables: M100 RMS peak data for each subject and each syllable. Legend as in Figure 4.11.

Figure 4.25 Mean reaction time data for syllable discrimination task. Each column represents the mean across subjects for a given syllable presented to one ear. The black bars are the right ear presentation/left hemisphere recording condition and the white bars are the left ear presentation/right hemisphere recording condition. The apparent RT difference was not significant.

Figure 4.26 Discriminate syllables: M100 RMS peak for each subject and each syllable type. Legend as in Figure 4.11.

Figure 4.27 Task - notask M100 RMS differences in fT. Positive values (above abscissa) indicate that the values in the task condition were larger than the corresponding measurements in the passive condition. Negative values (below abscissa) indicate that the activation in the passive condition was greater than the discrimination condition for that hemisphere. Black bars represent the comparison for the left hemisphere, white bars represent the comparison for the right hemisphere. Plotted are the mean amplitude changes across syllables for each subject.

Figure 4.28 Single dipoles for syllables in two conditions for one subject (S4: AM). The filled circles represent the dipoles derived for the passive syllable condition, open triangles represent the dipoles for the syllable discrimination condition. The plotted data are from Table 4.14.

Figure 4.29 Single dipoles for syllables in two conditions for one subject (S5: JH). Legend as in Figure 4.28. The plotted data are summarized in Table 4.15.

Figure 4.30 Single dipoles for syllables in two conditions for one subject (S6: IS). Legend as in Figure 4.28. The plotted data are summarized in Table 4.19.

Figure 4.31 Single dipoles for syllables in two conditions for one subject (S1: JM). Legend as in Figure 4.28. The plotted data are summarized in Table 4.18.

Figure 4.32 Single dipoles for syllables in two conditions for one subject (S2: MM). Legend as in Figure 4.28. The plotted data are summarized in Table 4.16.

Figure 4.33 Single dipoles for syllables in two conditions for one subject (S3: RR). Legend as in Figure 4.28. The plotted data are summarized in Table 4.17.

Figure 4.34 Dipole trajectory around the M100 generated by 500Hz pure tone for S5: JH, overlaid on the subject's structural MR image. Two views are shown, coronal (T2-weighted) on the left and axial (T1-weighted) on the right. The left side of each image is the right hemisphere, by radiological convention. The interval of dipoles around the M100 was 128-148ms on the right and 126-132ms on the left. The MR slice was chosen at the 132ms. The dipoles computed for the tones were estimated to be in auditory association cortex by a neuroradiologist experienced with the auditory system (Dr. Albert Galaburda, Beth Israel Hospital, Boston, MA).

Figure 4.35 Dipole trajectory around the M100 generated by both types of the vowel /a/ for S5: JH, overlaid on the subject's structural MR image. In the left, the dipole trajectories were from 104-124ms (female a) and 101-121ms (male a). In the right, the plotted trajectories range from 77-97ms (female a) and 68-88ms (male a). On the left, the trajectories for the two types of /a/ completely overlap. In the right supratemporal cortex, the paths diverge. The dipoles were estimated to be in auditory association cortex.

Figure 4.36 Dipole trajectory around the M100 generated by passively presented syllables /bæ/ and /pæ/ for S5: JH, overlaid on the subject's structural MR image. The dipoles were estimated to be in auditory association cortex.

Figure 4.37 Dipole trajectory around the M100 generated by discriminated syllables /bæ/ and /pæ/ for S5: JH, overlaid on the subject's structural MR image. The dipoles were estimated to be in auditory association cortex.

**TABLE 2.2** *Reports on word deafness and probable lesion site*

| Report | Year | Lesion evaluated | Probable Lesion site | Bilateral lesion | Cortical lesion | Subcort. lesion | Nonling. agnosia |
|---|---|---|---|---|---|---|---|
| Wernicke | 1874 | | Lpost. temp. | - | ? | + | ? |
| Kussmaul | 1877 | | LSTG | - | + | ? | |
| Girodeau | 1882 | autopsy | temp. lobe | + | + | ? | ? |
| Wernicke & Friedlaender | 1883 | autopsy | sup. temp. lobe | + | + | ? | + |
| Lichtheim | 1885 | | Lpost temp lobe | - | ? | + | |
| Freud | 1891 | | aud. ctx. | + | + | ? | |
| Mills | 1891 | autopsy | temp. lobe | + | ? | ? | ? |
| Pick | 1892 | autopsy | temp. lobe | + | + | ? | probably |
| Freund | 1895 | | subcortical | ? | ? | + | |
| Ziehl | 1896 | | ? | ? | ? | ? | - |
| Bastian | 1897 | | Ltemp. lobe | - | ? | ? | |
| Dejerine& Serieux | 1898 | autopsy | temp. lobe | + | + | ? | ? |
| Liepmann (see also Liepmann & Storch, 1902) | 1898 | autopsy | Ltemp. subcort. | - | ? | + | very mild |
| Ballett | 1903 | autopsy | temp.lobe cort. & subcort. | + | + | + | ? |
| Bonvincini | 1905 | autopsy | temp. lobe | + | + | ? | mild |
| Barrett | 1910 | autopsy | temp. lobe | + | + | ? | mild |
| Henschen | 1919 | autopsy | post. STG post. MTG | + | + | ? | ? |
| Henneberg | 1926 | autopsy | Ltemp. subcort. | - | - | + | ? |
| Schuster & Taterka | 1926 | autopsy | Ltemp. subcort | - | ? | + | very mild |
| Wohlfahrt et | 1952 | autopsy | post. STG | + | + | ? | very mild |
| Ziegler | 1952 | autopsy | post. STG | + | + | ? | + |
| Klein & Harper | 1956 | | Ltemp | - | ? | + | ? |

**Table 2.2, continued**

| Report | Year | Lesion evaluated | Probable Lesion site | Bilateral lesion | Cortical lesion | Subcort. lesion | Nonling. agnosia |
|---|---|---|---|---|---|---|---|
| Krauland-Steinbereithner | 1956 | | temp. lobe | + | ? | ? | - |
| Richter | 1957 | | temp. lobe | + | ? | ? | - |
| Lemoyne & Mahoudeau | 1959 | autopsy | temp.lobe | + | + | ? | amusia |
| Jerger et | 1969 | scintiogr., arteriogr., EEG | post. sup temp. | + | + | ? | mild |
| Lhermitte et | 1972 | autopsy | temp. lobe | + | + | ? | + |
| Gazzaniga et | 1973 | clinical exam | Ltemp. lobe | - | ? | ? | mild |
| Kanshepolsky et. (see also Jerger et1972) | 1973 | arteriogr., autopsy | post STG | + | + | basal ganglia | - |
| Goldstein et | 1975 | EEG | ? | + | ? | ? | mild |
| Saffran et | 1976 | | left temp. | ? | ? | ? | - |
| Earnest et | 1977 | CT | temporo-parietal | + | + | + | + |
| Ulrich | 1978 | META-ANALYSIS | | | | | |
| Dalla et | 1979 | | post.peri-sylvian | + | + | | |
| Graham et | 1980 | CT | sup. temp. | + | + | ? | + |
| Hamanaka et | 1980 | autopsy | Lbas gang and subcort. | - | - | + | - |
| Michel et | 1980 | CT | STG, MTG Rparietal | + | + | ? | + |
| Parving et | 1980 | CT, scintigr., rCBF | STG&MTG mid&post | + | + | ? | + |
| Kamei et | 1981 | autopsy | LpostSTG, SMG, AG, white matter | - | + | + | - |
| Auerbach et | 1982 | CT, EEG | L: STG<br>R: Heschl's, STG | + | + | + | + |

## Table 2.2, continued

| Report | Year | Lesion evaluated | Probable Lesion site | Bilateral lesion | Cortical lesion | Subcort. lesion | Nonling. agnosia |
|---|---|---|---|---|---|---|---|
| Miceli | 1982 | CT | temp. lobe | + | + | ? | + |
| Sato et | 1982 | | temporo-parietal | + | + | | + |
| Stockert | 1982 | CT | sup. temp. | + | + | + | + |
| Coslett et | 1984 | CT | prim. aud ctx., STG | + | + | | - |
| Metz-Lutz & Dahl | 1984 | autopsy | Ltemporo-parietal, subcort. | - | + | + | - |
| Brick et | 1985 | CT, autopsy | R&L temp.white matter of STG | + | - | + | ? |
| Buchman et | 1986 | Patient1: CT, autopsy | Rtemp. Lpost temp.par | + | + | + | + |
| | | Patient2: CT | Rtemp. parietal Lpost. temp. | + | + | + | + |
| | | Patient3: EEG | R&L temp. | + | ? | ? | + |
| Kanter et | 1986 | CT, angiogr. | L post. STG | - | NA | NA | NA |
| Roberts et | 1987 | CT | Rtemp. parietal | - | + | + | + |
| Tanaka et | 1987 | CT, MRI | LpostSTG<br>RpostSTG<br>L&R SMG, AngG | + | + | + | + |
| Mendez & Geehan | 1988 | Patient1: CT | RSTG, LSTG | + | + | + | + |
| | | Patient2: CT | Rfronto par.-temp<br>Lpar.-temp | + | + | + | + |
| Yaqub et | 1988 | CT | LpostSTG, RSTG, RpostMTG | + | + | + | - |
| Wolberg et | 1990 | Patient.1: CT | sup temp. lobe | + | + | ? | amusia |
| | | Patient.2: CT | sup temp. lobe | + | + | ? | - |

## Table 2.2, continued

| Report | Year | Lesion evaluated | Probable Lesion site | Bilateral lesion | Cortical lesion | Subcort. lesion | Nonling. agnosia |
|---|---|---|---|---|---|---|---|
| Kazui et | 1990 | CT, MRI | aud. radiations | + | - | + | + |
| Kitayama et | 1990 | pneumoenc. angiogr. | temp. lobe | + | + | + | amusia |
| Phillips & Farmer | 1990 | META-ANALYSIS | | | | | |
| Mendez & Rosenberg | 1991 | CT | L STG (medial aspect) | - | + | + | - |
| Shindo et | 1991 | MRI | left subcort., temp. | - | ? | + | ? |
| Praamstra et | 1991 | CT | LpostSTG<br>RpostSTG | + | + | slight, if any | mild |
| Takahashi et | 1992 | CT, MRI, PET | L thalamus, aud. radiations | - | - | + | - |

**Key**: +affirmative/presence; -negative/absence; ?inconclusive/insufficient evidence. **Abbreviations**: L left; R right. temp.: temporal; sup.: superior; post.: posterior; cort.: cortical; subcort.: subcortical; STG: superior temporal gyrus; MTG: middle temporal gyrus; AG: angular gyrus; SMG: supramarginal gyrus; aud.ctx.: auditory cortex; par.: parietal.

**TABLE 3.2** *PET subtraction studies showing foci in the superior temporal gyrus*

| Study | Stimulus & Task | Stim. rate (word/min) | Area | Tailarach coordinates | | |
|---|---|---|---|---|---|---|
| | | | | x | y | z |
| Petersen 88 | passive words | 60 | Lsup temp | -46 | -10 | 14 |
| | (control: fixation) | | Rsup temp | 42 | -16 | 12 |
| | [no task] | | | | | |
| Wise 91 | nonwords, wordpairs | 40-60 | LpostSTG | -46 | -38 | 8 |
| | (control: rest) | | RpostSTG | 48 | -38 | 8 |
| | [no task & task] | | LmidSTG | -50 | -10 | -4 |
| | | | RmidSTG | 50 | -10 | -4 |
| Frith 91 | lexical decision | 24 | L 42, 22 | activation asymmetric (L>R) | | |
| | (control: rest) | | R 42, 22 | | | |
| | [task] | | | | | |
| Howard 92 | hear sound/say word | 40 | LSTG | -50 | -30 | 8 |
| | (control: see/say) | | RSTG | 52 | -18 | 8 |
| | [task] | | | | | |
| | word repetition | 40 | LSTG | -48 | -38 | -4 |
| | (control: hear/say) | | | | | |
| | [task] | | | | | |
| Price 92 | word list (nouns) | 10-90 | LpostSTG | -58 | -34 | 12 |
| | (control: rest) | | RpostSTG | 58 | -34 | 12 |
| | [no task] | | LmidSTG | -50 | -22 | 4 |
| | | | RmidSTG | 50 | -22 | 4 |
| Zatorre 92 | passive syllables | 14-18 | LpostSTG | -58 | -21 | 8 |
| | (control: noise) | | RmidSTG | 60 | -14 | 0 |
| | [no task] | | | | | |
| Mazoyer 93 | stories,Ss,words | ? | LSTG | | | |
| | (control: rest) | | RSTG | | | |
| | [no task] | | | | | |
| Demonet 94 | 4 tests with nonwords | 20 | LSTG | -58 | -8 | 4 |
| | (control: pure tone task) | | RSTG | 54 | -10 | 4 |
| | [task] | | | | | |
| Fiez (in press) | vowels, CV, CVC | app.30 | L 41, 42 | -50 | -18 | 11 |
| | (control: tones) | | R 41, 42 | 50 | -26 | 11 |
| | [no task & task] | | | | | |
| O'Leary (in pr.) | binaural words | ? | LpSTG | -57 | -31 | 7 |
| | (control: tone bursts) | | RpSTG | 51 | -11 | -3 |
| | [no task] | | | | | |

Key: Tailarach coordinates in mm (Tailarach & Tournoux,1988). x: left/right axis; y: superior/inferior axis; z: anterior/posterior axis. L - left; R - right; sup - superior; post - posterior; STG - superior temporal gyrus. 42, 22 - Brodmann's areas.

**TABLE 4.3** *TONES: M100 RMS PEAK AND LATENCY*

| | | RMS peak | | Lat. of RMS peak | |
|---|---|---|---|---|---|
| | Stimulus | Left | Right | Left | Right |
| S1: JM | 500Hz | 130.239 | 68.875 | 120.956 | 111.356 |
| S2: MM | 500Hz | 82.100 | 83.319 | 126.716 | 137.276 |
| S3: RR | | 64.857<br>(1500Hz) | 67.232<br>(500Hz) | 125.756 | 118.076 |
| S4: AM | 500Hz | 147.126 | 136.625 | 118.076 | 115.196 |
| S5: JH | 500Hz | 167.212 | 144.776 | 136.316 | 138.236 |
| S6: IS | 500Hz | 106.219 | 152.541 | 123.836 | 122.876 |

# TABLES 4.4a & 4.4b

**TABLE 4.4a** *500Hz tone: M100 peak RMS analysis of variance*

| FACTOR: | subject | gender | hemisphere |
|---|---|---|---|
| LEVELS: | 6 | 2 | 2 |
| TYPE : | RANDOM | BETWEEN | WITHIN |

| SOURCE | SS | df | MS | F | p | eta | $eta^2$ |
|---|---|---|---|---|---|---|---|
| mean | 152127.33 | 1 | 152127.33 | 332.10 | 0.000 *** | | |
| s/g | 1832.28 | 4 | 458.07 | | | | |
| gen | 10672.99 | 1 | 10672.99 | 23.30 | 0.008 ** | 0.97 | 0.96 |
| s/g | 1832.28 | 4 | 458.07 | | | | |
| hem | 164.16 | 1 | 164.16 | 0.245 | 0.647 | 0.44 | 0.19 |
| hs/g | 2679.9316 | 4 | 669.98 | | | | |
| gh | 421.91 | 1 | 421.91 | 0.630 | 0.472 | 0.62 | 0.38 |
| hs/g | 2679.93 | 4 | 669.98 | | | | |

**TABLE 4.4b** *500Hz tone: M100 latency of peak RMS analysis of variance*

| FACTOR: | subject | gender | hemisphere |
|---|---|---|---|
| LEVELS: | 6 | 2 | 2 |
| TYPE: | RANDOM | BETWEEN | WITHIN |

| SOURCE | SS | df | MS | F | p | eta | $eta^2$ |
|---|---|---|---|---|---|---|---|
| mean | 186170.36 | 1 | 186170.36 | 1063.19 | 0.000 *** | | |
| s/g | 700.41 | 4 | 175.10 | | | | |
| gen | 17.27 | 1 | 17.27 | 0.099 | 0.769 | 0.29 | 0.09 |
| s/g | 700.41 | 4 | 175.1040 | | | | |
| hem | 6.22 | 1 | 6.22 | 0.19 | 0.684 | 0.40 | 0.16 |
| hs/g | 129.63 | 4 | 32.4096 | | | | |
| gh | 1.92 | 1 | 1.92 | 0.059 | 0.820 | 0.23 | 0.05 |
| hs/g | 129.63 | 4 | 32.40 | | | | |

## TABLE 4.5
## M100 single dipole parameters for tone (500Hz)

| Subject | Latency (ms) | Xpos (cm) | Ypos (cm) | Zpos (cm) | Corr. | Conf. vol. (m^3) | Good. of fit | RMS (Tesla) |
|---|---|---|---|---|---|---|---|---|
| JM left | 120.96 | 0.73167 | 5.4869 | 5.7824 | 0.98842 | 1.00E-08 | 0.98155534 | 1.30E-13 |
| JM right | 112.32 | 1.5286 | -5.3204 | 6.406 | 0.99661 | 2.17E-08 | 0.99603379 | 6.88E-14 |
| MM left | 124.8 | 0.023359 | 5.6064 | 6.8496 | 0.98289 | 4.18E-08 | 0.98666984 | 8.18E-14 |
| MM right | 130.56 | 0.68784 | -4.5986 | 5.7408 | 0.97824 | 4.82E-09 | 0.96741641 | 8.02E-14 |
| RR left | 120.96 | 0.0009007 | 2.9294 | 6.9803 | 0.96048 | 1.28E-08 | 0.97610605 | 6.31E-14 |
| RR right | 120 | 0.28938 | -6.2324 | 6.0882 | 0.99113 | 3.00E-08 | 0.98330235 | 6.70E-14 |
| AM left | 118.08 | -0.99965 | 4.7034 | 6.7184 | 0.99838 | 4.41E-10 | 0.99636424 | 1.47E-13 |
| AM right | 115.2 | 0.6145 | -6 | 6.1837 | 0.99612 | 5.99E-09 | 0.98961556 | 1.37E-13 |
| JH left | 126.72 | 0.18714 | 5.7792 | 6.7822 | 0.98499 | 1.04E-08 | 0.98031551 | 1.48E-13 |
| JH right | 126.72 | 0.39752 | -5.1755 | 6.9198 | 0.9833 | 9.75E-09 | 0.96017104 | 1.28E-13 |
| IS left | 120.96 | 1.5601 | 5.1591 | 5.6444 | 0.99313 | 9.30E-08 | 0.98607469 | 1.06E-13 |
| IS right | 122.88 | 0.66743 | -4.8175 | 6.4273 | 0.97942 | 1.57E-08 | 0.97926354 | 1.53E-13 |

**TABLE 4.7** *VOWELS: M100 RMS peak and latency of RMS peak*

| | | RMS peak | | Lat. of RMS peak | |
|---|---|---|---|---|---|
| | | Left | Right | Left | Right |
| | Stimulus | | | | |
| S1: JM | a | 142.1 | 108.3 | 102.7 | 88.3 |
| | i | 191.2 | 118.22 | 120.5 | 112.8 |
| | u | 149.65 | 126.62 | 109 | 113.3 |
| | a(f) | 174.2 | 104.91 | 104.2 | 91.7 |
| | i(f) | 164.32 | 97.88 | 119.5 | 112.8 |
| | u(f) | 146.78 | 108.1 | 116.6 | 109.4 |
| S2: MM | a | 73.96 | 76.43 | 140.1 | 128.6 |
| | i | 78.71 | 65.33 | 184.3 | 141.6 |
| | u | 79.27 | 56.14 | 124.3 | 178.5 |
| | a(f) | 89.81 | 94.03 | 108.5 | 112.3 |
| | i(f) | 93.39 | 86.68 | 152.6 | 126.7 |
| | u(f) | 78.81 | 86.2 | 139.6 | 122.4 |
| S3: RR | a | 80.6 | 96.81 | 114.7 | 124.3 |
| | i | 61.55 | 81.17 | 126.2 | 136.8 |
| | u | 57.93 | 78.34 | 130.6 | 127.2 |
| | a(f) | 73.58 | 95.34 | 111.4 | 134.8 |
| | i(f) | 50.33 | 70.2 | 131 | 136.8 |
| | u(f) | 49.97 | 75.87 | 124.8 | 127.2 |
| S4: AM | a | 131.04 | 88.3 | 101.8 | 88.3 |
| | i | 128 | 117.6 | 116.6 | 117.6 |
| | u | 116.24 | 88.8 | 113.8 | 88.8 |
| | a(f) | 141.28 | 92.6 | 96.9 | 92.6 |
| | i(f) | 140.27 | 113.3 | 108.9 | 113.3 |
| | u(f) | 140.39 | 105.1 | 103.7 | 105.1 |
| S5: JH | a | 153.42 | 84.29 | 107 | 83.5 |
| | i | 149.9 | 104.67 | 117.1 | 101.7 |
| | u | 109.29 | 80.62 | 106.6 | 101.3 |
| | a(f) | 196.1 | 81.51 | 110.4 | 92.2 |
| | i(f) | 134.55 | 91.56 | 116.6 | 109.4 |
| | u(f) | 145.24 | 81 | 119 | 100.3 |
| S6: IS | a | 109.5 | 93.6 | 101.3 | 93.6 |
| | i | 102.46 | 99.8 | 123.4 | 99.8 |
| | u | 75.74 | 116.2 | 123.8 | 116.2 |
| | a(f) | 110.86 | 86.8 | 94.6 | 86.8 |
| | i(f) | 110.34 | 108.5 | 107.5 | 108.5 |
| | u(f) | 116.29 | 96.5 | 124.8 | 96.5 |

**TABLE 4.8** *Vowels: M100 peak RMS analysis of variance*

| FACTOR: | subject | hemisphere | fund.freq | phoneme |
|---|---|---|---|---|
| LEVELS: | 6 | 2 | 2 | 3 |
| TYPE: | RANDOM | WITHIN | WITHIN | WITHIN |

| SOURCE | SS | df | MS | F | p | | eta | $eta^2$ |
|---|---|---|---|---|---|---|---|---|
| mean | 848042 | 1 | 848042 | 105 | 0.000 | *** | | |
| s/ | 40031 | 5 | 8006 | | | | | |
| hem | 3201 | 1 | 3201 | 0.725 | 0.433 | | 0.35 | 0.13 |
| hs/ | 22087 | 5 | 4417 | | | | | |
| ffreq | 506 | 1 | 506 | 1.763 | 0.242 | | 0.51 | 0.26 |
| fs/ | 1436.0434 | 5 | 287.2087 | | | | | |
| hf | 275.9294 | 1 | 275.9294 | 1.832 | 0.234 | | 0.52 | 0.27 |
| hfs/ | 753.2474 | 5 | 150.6495 | | | | | |
| phon | 3719.3065 | 2 | 1859.6533 | 8.135 | 0.008 | ** | 0.79 | 0.62 |
| ps/ | 2286.0960 | 10 | 228.6096 | | | | | |
| hp | 29.6321 | 2 | 14.8161 | 0.044 | 0.957 | | 0.09 | 0.01 |
| hps/ | 3374.4442 | 10 | 337.4444 | | | | | |
| fp | 1796.7979 | 2 | 898.3990 | 4.163 | 0.048 | * | 0.67 | 0.45 |
| fps/ | 2158.2373 | 10 | 215.8237 | | | | | |
| hfp | 72.8066 | 2 | 36.4033 | 0.248 | 0.785 | | 0.22 | 0.05 |
| hfps/ | 1470.6228 | 10 | 147.0623 | | | | | |

Key: s - subjects; hem - hemisphere; ffreq - fundamental frequency; hf - hemisphere x fundamental; phone - phoneme type (vowel); hp - hemisphere x phoneme; fp - fundamental x phoneme; hfp - hemisphere x fundamental x phoneme. * $p<0.05$ ** $p<0.01$ *** $P<0.001$.

**TABLE 4.9** *Vowels: latency of RMS peak analysis of variance*

| FACTOR: | subject | hemisphere | fundfreq | phon |
|---|---|---|---|---|
| LEVELS: | 6 | 2 | 2 | 3 |
| TYPE : | RANDOM | WITHIN | WITHIN | WITHIN |

| SOURCE | SS | df | MS | F | p | eta | eta2 |
|---|---|---|---|---|---|---|---|
| mean | 951142.3100 | 1 | 951142.3100 | 388.0 | 0.000 *** | | |
| s/ | 12256.8361 | 5 | 2451.3672 | | | | |
| hem | 756.6050 | 1 | 756.6050 | 4.007 | 0.102 | 0.67 | 0.44 |
| hs/ | 944.1933 | 5 | 188.8387 | | | | |
| ffreq | 259.1606 | 1 | 259.1606 | 0.872 | 0.393 | 0.38 | 0.15 |
| fs/ | 1485.6813 | 5 | 297.1363 | | | | |
| hf | 1.3339 | 1 | 1.3339 | 0.024 | 0.883 | 0.069 | 0.004 |
| hfs/ | 279.7445 | 5 | 55.9489 | | | | |
| phon | 4135.8812 | 2 | 2067.9406 | 26.98 | 0.000*** | 0.92 | 0.84 |
| ps/ | 766.3656 | 10 | 76.6366 | | | | |
| hp | 65.4233 | 2 | 32.7116 | 0.183 | 0.835 | 0.19 | 0.03 |
| hps/ | 1784.0337 | 10 | 178.4034 | | | | |
| fp | 6.1678 | 2 | 3.0839 | 0.335 | 0.723 | 0.25 | 0.06 |
| fps/ | 92.1455 | 10 | 9.2146 | | | | |
| hfp | 485.0343 | 2 | 242.5172 | 1.926 | 0.196 | 0.53 | 0.28 |
| hfps/ | 1258.8822 | 10 | 125.8882 | | | | |

Key: see Table 4.8

## TABLE 4.10
## Single dipole parameters for six subjects' responses to the vowel /a/

| Subject | Latency (s) | x pos (m) | y pos (m) | z pos (m) | Corr. | Conf. vol. (m^3) | Good. of fit | Rms (Tesla) |
|---|---|---|---|---|---|---|---|---|
| JMlefta | 0.1032 | 0.0094739 | 0.05636 | 0.067881 | 0.97583 | 1.64E-10 | 0.95125 | 1.42E-13 |
| JMleftaf | 0.10128 | 0.0072664 | 0.053127 | 0.066743 | 0.97663 | 1.13E-10 | 0.95331 | 1.72E-13 |
| JMrighta | 0.088321 | 0.012572 | -0.049935 | 0.060462 | 0.97278 | 5.18E-10 | 0.93979 | 1.08E-13 |
| JMrightaf | 0.091681 | 0.010692 | -0.052579 | 0.059986 | 0.95992 | 5.03E-10 | 0.89987 | 1.05E-13 |
| | | | | | | | | |
| MMlefta | 0.1392 | 0.0082176 | 0.024435 | 0.03837 | 0.87014 | 9.72E-09 | 0.83267 | 7.39E-14 |
| MMleftaf | 0.10464 | 0.0017501 | 0.05017 | 0.056388 | 0.94404 | 8.97E-10 | 0.89063 | 8.90E-14 |
| MMrighta | 0.1248 | 0.0058088 | -0.033595 | 0.048333 | 0.90171 | 4.66E-09 | 0.81035 | 7.61E-14 |
| MMrightaf | 0.11472 | -0.00095285 | -0.024406 | 0.058611 | 0.93675 | 3.18E-09 | 0.8876 | 9.37E-14 |
| | | | | | | | | |
| RRlefta | 0.11952 | -0.008311 | 0.056933 | 0.062877 | 0.91066 | 8.33E-10 | 0.77925 | 7.93E-14 |
| RRleftaf | 0.11136 | -0.0064427 | 0.048118 | 0.06539 | 0.92944 | 1.32E-09 | 0.86708 | 7.36E-14 |
| RRrighta | 0.12624 | -0.0025737 | -0.05092 | 0.065312 | 0.95774 | 7.66E-10 | 0.9193 | 9.67E-14 |
| RRrightaf | 0.13104 | -0.0032411 | -0.04956 | 0.065245 | 0.95299 | 9.10E-10 | 0.91195 | 9.44E-14 |
| | | | | | | | | |
| AMlefta | 0.099362 | -0.0094424 | 0.049238 | 0.062834 | 0.97592 | 6.31E-10 | 0.95161 | 1.30E-13 |
| AMleftaf | 0.10176 | -0.0097299 | 0.0522 | 0.065989 | 0.97715 | 4.19E-10 | 0.95388 | 1.38E-13 |
| AMrighta | 0.090721 | 0.011786 | -0.04921 | 0.059033 | 0.96656 | 1.19E-09 | 0.93794 | 1.14E-13 |
| AMrightaf | 0.093601 | 0.0081462 | -0.055863 | 0.058783 | 0.97651 | 7.43E-10 | 0.9528 | 1.28E-13 |
| | | | | | | | | |
| JHlefta | 0.11184 | -0.0049354 | 0.050927 | 0.058348 | 0.9664 | 3.01E-10 | 0.93914 | 1.48E-13 |
| JHleftaf | 0.11424 | -0.0044069 | 0.045662 | 0.057204 | 0.96296 | 1.41E-10 | 0.94102 | 1.92E-13 |
| JHrighta | 0.078721 | -0.0015053 | -0.043239 | 0.07544 | 0.92537 | 1.09E-09 | 0.89141 | 8.34E-14 |
| JHrightaf | 0.087361 | 0.0025482 | -0.048102 | 0.075218 | 0.91867 | 1.57E-09 | 0.85055 | 8.05E-14 |
| | | | | | | | | |
| ISlefta | 0.1032 | 0.013794 | 0.057616 | 0.062819 | 0.97337 | 5.36E-10 | 0.94756 | 1.09E-13 |
| ISleftaf | 0.089281 | 0.014103 | 0.023881 | 0.048786 | 0.9315 | 4.09E-09 | 0.89985 | 1.09E-13 |
| ISrighta | 0.091681 | 0.01573 | -0.017611 | 0.053328 | 0.95 | 9.58E-10 | 0.95264 | 1.52E-13 |
| ISrightaf | 0.088801 | 0.0082887 | -0.046074 | 0.061877 | 0.99809 | 1.63E-10 | 0.99644 | 1.63E-13 |

**TABLE 4.11**
M100 RMS peak and latency for passive syllable and discriminate syllable conditions

| | | **3a: Stops - passive** | | | | **3b: Stops - discriminate** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | RMS peak | | Lat. of RMS peak | | RMS peak | | Lat. of RMS peak | |
| | | Left | Right | Left | Right | Left | Right | Left | Right |
| | Stim | | | | | | | | |
| S1: JM | bæ | 164.2 | 132.08 | 93.1 | 83.5 | 184.89 | 125.07 | 95 | 83.5 |
| | pæ | 175.2 | 152.7 | 97.9 | 89.3 | 176.8 | 120.94 | 98.8 | 88.3 |
| | dæ | 189.9 | 145.62 | 95 | 84.5 | 214.57 | 116.81 | 98 | 85.4 |
| | tæ | 187.1 | 149.01 | 95 | 87.3 | 136.97 | 121.3 | 98 | 86.3 |
| S2: MM | bæ | 96.15 | 100.67 | 96.95 | 101.76 | 105.52 | 89.57 | 122 | 108.5 |
| | pæ | 94.42 | 133.6 | 101.76 | 110.4 | 91.7 | 94.08 | 100.8 | 101.7 |
| | dæ | 84.54 | 83.5 | 111.36 | 116.1 | 112.6 | 64.2 | 119 | 107.5 |
| | tæ | 74.87 | 98.04 | 108.48 | 102.7 | 114.07 | 64.78 | 120.9 | 112.3 |
| S3: RR | bæ | 90.27 | 103.64 | 104.6 | 112.3 | 105.82 | 97.82 | 107.5 | 102.7 |
| | pæ | 80.34 | 130.28 | 103.7 | 117.1 | 116.61 | 101.83 | 113.3 | 107.5 |
| | dæ | 78.03 | 105.57 | 105.6 | 108.47 | 113.93 | 102.47 | 103.7 | 101.7 |
| | tæ | 88.51 | 105.1 | 101.7 | 106.5 | 134.13 | 96.52 | 112.3 | 96 |
| S4: AM | bæ | 145.08 | 125.73 | 102.7 | 85.4 | 146.51 | 121.87 | 105.6 | 89.3 |
| | pæ | 128.57 | 151.61 | 98.8 | 88.3 | 131.38 | 126.54 | 100.8 | 92.1 |
| | dæ | 136.64 | 128.49 | 101.7 | 82.5 | 157.65 | 110.2 | 104.6 | 92.1 |
| | tæ | 126.72 | 138.35 | 97.9 | 88.3 | 121.07 | 122.42 | 104.6 | 89.3 |
| S5: JH | bæ | 178.97 | 160.9 | 101.7 | 91.2 | 231.22 | 140.66 | 100.8 | 85.4 |
| | pæ | 133.58 | 135.95 | 104.6 | 90.2 | 152.04 | 125.24 | 102.72 | 83.5 |
| | dæ | 201.19 | 173.5 | 102.7 | 93.1 | 255.98 | 165.64 | 100.8 | 88.3 |
| | tæ | 150.32 | 119.92 | 98.8 | 93.1 | 185.11 | 130.82 | 96 | 82.5 |
| S6: IS | bæ | 143.31 | 121.31 | 96.9 | 96 | 149.03 | 184.76 | 98.8 | 118.1 |
| | pæ | 107.07 | 65.34 | 96.9 | 95 | 98.78 | 102.67 | 100.7 | 122.9 |
| | dæ | 167.37 | 106.13 | 96.9 | 94 | 145.2 | 150.38 | 98.8 | 129.6 |
| | tæ | 80.43 | 101.8 | 109.4 | 88.3 | 81.73 | 63.86 | 94.1 | 113.3 |

**TABLE 4.12 Syllables: M100 RMS peak analysis of variance**

| FACTOR: | subject | hemisphere | task | voice | place |
|---|---|---|---|---|---|
| LEVELS: | 6 | 2 | 2 | 2 | 2 |
| TYPE: | RANDOM | WITHIN | WITHIN | WITHIN | WITHIN |

| SOURCE | SS | df | MS | F | p | eta | eta$^2$ |
|---|---|---|---|---|---|---|---|
| mean | 1569634.93 | 1 | 1569634.93 | 120.4 | 0.000 *** | | |
| s/ | 65145.2629 | 5 | 13029.0526 | | | | |
| hem | 7646.9401 | 1 | 7646.9401 | 4.423 | 0.089 | 0.68 | 0.47 |
| hs/ | 8644.6962 | 5 | 1728.9392 | | | | |
| task | 181.8852 | 1 | 181.8852 | 0.349 | 0.580 | 0.25 | 0.065 |
| ts/ | 2606.0215 | 5 | 521.2043 | | | | |
| ht | 3612.7784 | 1 | 3612.7784 | 3.933 | 0.104 | 0.66 | 0.44 |
| hts/ | 4593.2676 | 5 | 918.6535 | | | | |
| voice | 7261.7154 | 1 | 7261.7154 | 2.334 | 0.187 | 0.56 | 0.32 |
| vs/ | 15555.5811 | 5 | 3111.1162 | | | | |
| hv | 1900.6836 | 1 | 1900.6836 | 9.039 | 0.030* | 0.80 | 0.64 |
| hvs/ | 1051.3408 | 5 | 210.2682 | | | | |
| tv | 1113.9800 | 1 | 1113.9800 | 8.700 | 0.032* | 0.79 | 0.63 |
| tvs/ | 640.2450 | 5 | 128.0490 | | | | |
| htv | 36.2604 | 1 | 36.2604 | 0.096 | 0.769 | 0.14 | 0.02 |
| htvs/ | 1884.1431 | 5 | 376.8286 | | | | |
| place | 49.9682 | 1 | 49.9682 | 0.117 | 0.747 | 0.15 | 0.022 |
| ps/ | 2143.1408 | 5 | 428.6282 | | | | |
| hp | 885.7351 | 1 | 885.7351 | 12.51 | 0.017* | 0.84 | 0.71 |
| hps/ | 353.8473 | 5 | 70.7695 | | | | |
| tp | 0.7740 | 1 | 0.7740 | 0.002 | 0.963 | 0.022 | 0.000 |
| tps/ | 1624.9697 | 5 | 324.9939 | | | | |
| htp | 142.8864 | 1 | 142.8864 | 1.288 | 0.308 | 0.45 | 0.20 |
| htps/ | 554.7903 | 5 | 110.9581 | | | | |
| vp | 414.0872 | 1 | 414.0872 | 6.669 | 0.049* | 0.76 | 0.57 |
| vps/ | 310.4440 | 5 | 62.0888 | | | | |
| hvp | 20.9066 | 1 | 20.9066 | 0.065 | 0.808 | 0.11 | 0.012 |
| hvps/ | 1598.6010 | 5 | 319.7202 | | | | |
| tvp | 2.0242 | 1 | 2.0242 | 0.050 | 0.831 | 0.09 | 0.009 |
| tvps/ | 200.8177 | 5 | 40.1635 | | | | |
| htvp | 30.8720 | 1 | 30.8720 | 0.134 | 0.730 | 0.16 | 0.025 |
| htvps/ | 1153.9632 | 5 | 230.7926 | | | | |

## TABLE 4.14
## M100 single dipole parameters for syllables in passive and discriminate conditions (subject AM)

| Subj. & stim | Latency (s) | x pos (m) | y pos (m) | z pos (m) | Corr. | Conf. vol. (m^3) | Good. of fit | Rms (Tesla) |
|---|---|---|---|---|---|---|---|---|
| AMleft.b | 0.10464 | -0.0067548 | 0.047759 | 0.06432 | 0.9807 | 3.84E-10 | 0.96183 | 1.44E-13 |
| AMleft.p | 0.094077 | -0.0078179 | 0.018088 | 0.069401 | 0.96006 | 5.89E-09 | 0.93192 | 1.26E-13 |
| AMleft.d | 0.10752 | -0.0080207 | 0.048289 | 0.067042 | 0.97123 | 6.01E-10 | 0.93969 | 1.32E-13 |
| AMleft.t | 0.10272 | -0.0082012 | 0.049879 | 0.069274 | 0.97245 | 5.63E-10 | 0.94257 | 1.25E-13 |
| AMleft.dis.b | 0.10464 | -0.0063396 | 0.047413 | 0.068622 | 0.99432 | 4.37E-10 | 0.98479 | 1.46E-13 |
| AMleft.dis.p | 0.1056 | -0.0051035 | 0.042994 | 0.071574 | 0.99382 | 1.06E-09 | 0.987 | 1.29E-13 |
| AMleft.dis.d | 0.10848 | -0.0052256 | 0.045878 | 0.069642 | 0.98905 | 5.14E-10 | 0.97497 | 1.55E-13 |
| AMleft.dis.t | 0.1008 | -0.0066966 | 0.037154 | 0.07302 | 0.97595 | 2.47E-09 | 0.95132 | 1.20E-13 |
| AMright.b | 0.087357 | 0.0084834 | -0.054533 | 0.060872 | 0.96912 | 8.91E-10 | 0.93813 | 1.25E-13 |
| AMright.p | 0.092157 | 0.011121 | -0.04845 | 0.064858 | 0.97412 | 5.04E-10 | 0.95075 | 1.50E-13 |
| AMright.d | 0.086397 | 0.0068264 | -0.054654 | 0.060313 | 0.97818 | 7.67E-10 | 0.95552 | 1.27E-13 |
| AMright.t | 0.093117 | 0.010737 | -0.050823 | 0.065187 | 0.97702 | 4.93E-10 | 0.95641 | 1.36E-13 |
| AMright.dis.b | 0.091197 | 0.00817 | -0.055621 | 0.060547 | 0.96696 | 8.56E-10 | 0.93342 | 1.21E-13 |
| AMright.dis.p | 0.094077 | 0.0097879 | -0.053644 | 0.064145 | 0.97111 | 6.24E-10 | 0.94392 | 1.26E-13 |
| AMright.dis.d | 0.092157 | 0.010065 | -0.055219 | 0.061963 | 0.96523 | 9.28E-10 | 0.93319 | 1.10E-13 |
| AMright.dis.t | 0.093117 | 0.011132 | -0.047612 | 0.06686 | 0.96912 | 7.68E-10 | 0.9435 | 1.20E-13 |

## TABLE 4.15
## M100 single dipole parameters for syllables in passive and discriminate conditions (subject JH)

| Subj. & stim | Latency (s) | x pos (m) | y pos (m) | z pos (m) | Corr. | Conf. vol. (m^3) | Good. of fit | Rms (Tesla) |
|---|---|---|---|---|---|---|---|---|
| JHleft.b | 0.10272 | 0.0015423 | 0.055253 | 0.067695 | 0.97179 | 1.36E-10 | 0.94307 | 1.79E-13 |
| JHleft.p | 0.1008 | 0.0002925 | 0.061799 | 0.070822 | 0.96804 | 2.19E-10 | 0.92858 | 1.32E-13 |
| JHleft.d | 0.10176 | -0.001485 | 0.056572 | 0.06923 | 0.98716 | 8.70E-11 | 0.97415 | 2.01E-13 |
| JHleft.t | 0.098877 | -0.0026819 | 0.058745 | 0.067382 | 0.97121 | 1.96E-10 | 0.94201 | 1.50E-13 |
| | | | | | | | | |
| JHleft.dis.b | 0.095997 | 0.00073106 | 0.055636 | 0.068468 | 0.96984 | 6.12E-11 | 0.93867 | 2.32E-13 |
| JHleft.dis.p | 0.097917 | -0.0029818 | 0.058171 | 0.06753 | 0.96088 | 2.26E-10 | 0.91364 | 1.49E-13 |
| JHleft.dis.d | 0.096957 | 0.0014753 | 0.052047 | 0.065273 | 0.9739 | 6.17E-11 | 0.93748 | 2.52E-13 |
| JHleft.dis.t | 0.095037 | -0.0014396 | 0.057826 | 0.070873 | 0.96403 | 1.04E-10 | 0.91882 | 1.85E-13 |
| | | | | | | | | |
| JHright.b | 0.089277 | 0.003047 | -0.048521 | 0.06377 | 0.97392 | 2.70E-10 | 0.94071 | 1.60E-13 |
| JHright.p | 0.092157 | 0.00097548 | -0.051426 | 0.066246 | 0.97381 | 3.52E-10 | 0.94775 | 1.36E-13 |
| JHright.d | 0.095037 | 0.0020424 | -0.055186 | 0.069044 | 0.95565 | 1.28E-10 | 0.91118 | 1.73E-13 |
| JHright.t | 0.091197 | 0.0012418 | -0.051134 | 0.068687 | 0.95091 | 5.08E-10 | 0.90543 | 1.20E-13 |
| | | | | | | | | |
| JHright.dis.b | 0.083517 | 0.004921 | -0.058437 | 0.064604 | 0.94601 | 2.45E-10 | 0.87847 | 1.41E-13 |
| JHright.dis.p | 0.083517 | -0.0055843 | -0.057675 | 0.059877 | 0.95388 | 3.93E-10 | 0.90517 | 1.25E-13 |
| JHright.dis.d | 0.083517 | -0.009402 | -0.029585 | 0.050197 | 0.91682 | 1.85E-09 | 0.84851 | 1.63E-13 |
| JHright.dis.t | 0.087357 | -0.015513 | -0.046236 | 0.055887 | 0.93329 | 5.83E-10 | 0.88216 | 1.26E-13 |

# TABLE 4.16

## M100 single dipole parameters for syllables in passive and discrimination condition (subject MM)

| | Latency ms | x pos m | y pos m | z pos m | Corr. | Conf. vol. m^3 | Good. of fit | Rms (T) |
|---|---|---|---|---|---|---|---|---|
| MMleftb | 101 | -0.010342 | 0.013473 | 0.055573 | 0.90075 | 6.60E-09 | 0.90656 | 9.53E-14 |
| MMleftp | 106 | -0.0097989 | 0.032315 | 0.060749 | 0.92561 | 1.57E-09 | 0.90765 | 9.24E-14 |
| MMleftd | 111 | 0.00025525 | 0.048476 | 0.049452 | 0.95155 | 1.19E-09 | 0.90836 | 8.45E-14 |
| MMleftt | 113 | 0.001167 | 0.054414 | 0.058006 | 0.91957 | 1.38E-09 | 0.83785 | 7.39E-14 |
| MMleft.dis.b | 117 | 0.007656 | 0.016264 | 0.047112 | 0.95287 | 5.73E-09 | 0.91321 | 1.05E-13 |
| MMleft.dis.p | 105 | 0.0012999 | 0.047736 | 0.061553 | 0.95245 | 9.60E-10 | 0.91004 | 9.09E-14 |
| MMleft.dis.d | 114 | 0.0016834 | 0.036349 | 0.060753 | 0.96006 | 7.40E-10 | 0.92562 | 1.11E-13 |
| MMleft.dis.t | 115 | 0.0040208 | -0.0032324 | 0.042149 | 0.95626 | 1.43E-08 | 0.93199 | 1.13E-13 |
| MMrightb | 101 | 0.0053611 | -0.046044 | 0.060661 | 0.96159 | 7.39E-10 | 0.92161 | 1.01E-13 |
| MMrightp | 115 | 0.0034169 | -0.034326 | 0.060441 | 0.97598 | 7.03E-10 | 0.95244 | 1.32E-13 |
| MMrightd | 115 | 0.0033678 | -0.043801 | 0.058457 | 0.94758 | 1.67E-09 | 0.8959 | 8.35E-14 |
| MMrightt | 98 | 0.001595 | -0.02133 | 0.056683 | 0.95209 | 4.03E-09 | 0.91068 | 9.60E-14 |
| MMright.dis.b | 104 | 0.012518 | -0.0075195 | 0.069646 | 0.79719 | 9.69E-09 | 0.78626 | 8.85E-14 |
| MMright.dis.p | 103 | 0.011936 | -0.034804 | 0.052956 | 0.88721 | 1.64E-09 | 0.81013 | 9.40E-14 |
| MMright.dis.d | 112 | 0.0015254 | -0.064909 | 0.063362 | 0.88731 | 1.03E-09 | 0.78754 | 6.39E-14 |
| MMright.dis.t | 107 | 0.018161 | -0.025718 | 0.063425 | 0.75108 | 7.36E-09 | 0.68474 | 6.36E-14 |

# TABLE 4.17

## M100 single dipole parameters for syllables in passive and discrimination conditions (subject RR)

| | Latency ms | x pos m | y pos m | z pos m | Corr, | Conf. vol. m^3 | Good. of fit | Rms (T) |
|---|---|---|---|---|---|---|---|---|
| RRleftb | 108 | -0.0029699 | 0.036769 | 0.062357 | 0.93415 | 1.04E-09 | 0.88239 | 8.91E-14 |
| RRleftp | 101 | -0.0096478 | 0.054087 | 0.066392 | 0.93956 | 8.81E-10 | 0.8708 | 7.98E-14 |
| RRleftd | 107 | -0.0091563 | 0.054988 | 0.065524 | 0.93904 | 8.52E-10 | 0.88028 | 7.79E-14 |
| RRleftt | 96 | -0.0068383 | 0.048728 | 0.060533 | 0.94178 | 1.02E-09 | 0.88017 | 8.55E-14 |
| | | | | | | | | |
| RRleft.dis.b | 106 | -0.0078715 | 0.049144 | 0.065662 | 0.9671 | 4.42E-10 | 0.93575 | 1.06E-13 |
| RRleft.dis.p | 108 | -0.00094874 | 0.040752 | 0.064553 | 0.95347 | 4.68E-10 | 0.91403 | 1.16E-13 |
| RRleft.dis.d | 108 | -0.0015937 | 0.040281 | 0.064434 | 0.95864 | 5.38E-10 | 0.9224 | 1.12E-13 |
| RRleft.dis.t | 107 | -0.005707 | 0.043859 | 0.064555 | 0.97743 | 3.45E-10 | 0.95427 | 1.32E-13 |
| | | | | | | | | |
| RRrightb | 107 | 0.0037127 | -0.056026 | 0.063085 | 0.96165 | 4.94E-10 | 0.91849 | 1.01E-13 |
| RRrightp | 114 | 0.0032228 | -0.050883 | 0.065022 | 0.97274 | 3.26E-10 | 0.94583 | 1.30E-13 |
| RRrightd | 106 | 0.0051048 | -0.054853 | 0.06314 | 0.96253 | 5.03E-10 | 0.92294 | 1.05E-13 |
| RRrightt | 101 | 0.001898 | -0.055013 | 0.063299 | 0.96107 | 4.88E-10 | 0.91785 | 1.03E-13 |
| | | | | | | | | |
| RRright.dis.b | 100 | 0.0079342 | -0.030046 | 0.06476 | 0.90359 | 1.64E-09 | 0.84149 | 9.76E-14 |
| RRright.dis.p | 105 | 0.0025349 | -0.055684 | 0.068201 | 0.95657 | 4.41E-10 | 0.91458 | 1.01E-13 |
| RRright.dis.d | 103 | 0.010637 | -0.021197 | 0.055632 | 0.85931 | 2.56E-09 | 0.82243 | 1.02E-13 |
| RRright.dis.t | 91 | 0.005812 | -0.054819 | 0.061188 | 0.93062 | 6.79E-10 | 0.83126 | 9.48E-14 |

# TABLE 4.18

## M100 dipole parameters for syllables in passive and discrimination conditions (subject JM)

| | Latency ms | x pos m | y pos m | z pos m | Corr. | Conf. vol. m^3 | Good.of fit | Rms (T) |
|---|---|---|---|---|---|---|---|---|
| JMleftb | 92 | 0.0047868 | 0.046885 | 0.065803 | 0.97951 | 1.73E-10 | 0.95068 | 1.65E-13 |
| JMleftp | 97 | 0.0076448 | 0.054255 | 0.063232 | 0.98964 | 9.63E-11 | 0.97765 | 1.75E-13 |
| JMleftd | 94 | 0.0090151 | 0.052837 | 0.063602 | 0.98968 | 7.90E-11 | 0.97939 | 1.90E-13 |
| JMleftt | 95 | 0.0111 | 0.052377 | 0.0628 | 0.98734 | 8.49E-11 | 0.96298 | 1.87E-13 |
| | | | | | | | | |
| JMleft.dis.b | 11 | -0.023132 | -0.02047 | -0.069817 | 0.4985 | 0.0323 | 0.25622 | 3.02E-14 |
| JMleft.dis.p | 101 | 0.008048 | 0.055403 | 0.063221 | 0.98734 | 8.41E-11 | 0.97478 | 1.77E-13 |
| JMleft.dis.d | 95 | 0.0076387 | 0.053779 | 0.062415 | 0.99161 | 5.43E-11 | 0.98297 | 2.13E-13 |
| JMleft.dis.t | 96 | 0.0083911 | 0.056193 | 0.061453 | 0.97491 | 1.87E-10 | 0.95038 | 1.36E-13 |
| | | | | | | | | |
| JMrightb | 84 | 0.011337 | -0.05205 | 0.060822 | 0.97457 | 2.51E-10 | 0.92522 | 1.32E-13 |
| JMrightp | 87 | 0.012503 | -0.050656 | 0.060796 | 0.98618 | 1.61E-10 | 0.9711 | 1.53E-13 |
| JMrightd | 87 | 0.010439 | -0.046382 | 0.063279 | 0.96852 | 2.81E-10 | 0.91089 | 1.44E-13 |
| JMrightt | 85 | 0.011174 | -0.04993 | 0.06104 | 0.98277 | 1.87E-10 | 0.96063 | 1.49E-13 |
| | | | | | | | | |
| JMright.dis.b | 84 | 0.015059 | -0.049979 | 0.060308 | 0.98064 | 3.10E-10 | 0.96082 | 1.25E-13 |
| JMright.dis.p | 88 | 0.015743 | -0.050565 | 0.060484 | 0.97631 | 3.20E-10 | 0.95363 | 1.21E-13 |
| JMright.dis.d | 85 | 0.011705 | -0.05314 | 0.060249 | 0.97113 | 3.43E-10 | 0.91637 | 1.17E-13 |
| JMright.dis.t | 86 | 0.014411 | -0.046971 | 0.05934 | 0.97109 | 3.94E-10 | 0.94396 | 1.21E-13 |

## TABLE 4.19

## M100 dipole parameters for syllables in passive and discrimination condition (subject IS)

| | Latency | x pos | y pos | z pos | Correlation | Volume | GoF | Rms |
|---|---|---|---|---|---|---|---|---|
| | ms | m | m | m | units | m^3 | units | T |
| ISleftb | 101 | 0.012428 | 0.055029 | 0.058365 | 0.97202 | 3.05E-10 | 0.92854 | 1.42E-13 |
| ISleftp | 94 | 0.0061192 | -0.027718 | 0.053007 | 0.93912 | 4.92E-10 | 0.90742 | 1.06E-13 |
| ISleftd | 101 | 0.019014 | 0.040968 | 0.056304 | 0.97491 | 2.81E-10 | 0.95851 | 1.65E-13 |
| ISleftt | 112 | 0.016663 | 0.056294 | 0.055793 | 0.94376 | 1.67E-09 | 0.89659 | 8.02E-14 |
| | | | | | | | | |
| ISleft.dis.b | 93 | 0.014133 | 0.038382 | 0.040575 | 0.94923 | 7.17E-10 | 0.93345 | 1.47E-13 |
| ISleft.dis.p | 105 | 0.0012404 | 0.033184 | 0.097792 | 0.82925 | 4.88E-09 | 0.89019 | 9.78E-14 |
| ISleft.dis.d | 93 | 0.017155 | 0.038211 | 0.051264 | 0.95636 | 6.68E-10 | 0.93025 | 1.43E-13 |
| ISleft.dis.t | 95 | 0.013056 | 0.056057 | 0.050932 | 0.91943 | 1.70E-09 | 0.84858 | 8.17E-14 |
| | | | | | | | | |
| ISrightb | 99 | 0.0057641 | -0.049249 | 0.06597 | 0.96557 | 4.54E-10 | 0.91161 | 1.21E-13 |
| ISrightp | 97 | 0.010027 | -0.040869 | 0.067983 | 0.90952 | 4.71E-09 | 0.83155 | 6.51E-14 |
| ISrightd | 98 | 0.0086219 | -0.046156 | 0.061902 | 0.96529 | 7.06E-10 | 0.9368 | 1.05E-13 |
| ISrightt | 88 | -0.020517 | 0.019726 | -0.079641 | 0.91648 | 9.59E-05 | 0.88465 | 1.02E-13 |
| | | | | | | | | |
| ISright.dis.b | 123 | -0.015155 | 0.0031145 | 0.039196 | 0.93428 | 5.19E-09 | 0.97642 | 1.82E-13 |
| ISright.dis.p | 117 | 0.012618 | 0.067034 | 0.034442 | 0.93436 | 3.23E-06 | 0.92127 | 9.63E-14 |
| ISright.dis.d | 124 | -0.017735 | -0.0039448 | 0.056642 | 0.88064 | 4.14E-09 | 0.95363 | 1.47E-13 |
| ISright.dis.t | 117 | -0.0098503 | -0.010918 | 0.06212 | 0.8152 | 2.54E-08 | 0.7854 | 6.27E-14 |

TABLE 4.20

```
FACTOR:        sub        hem        f0       phone       DATA
LEVELS:          6          2         2           3         72
TYPE  :     RANDOM     WITHIN    WITHIN      WITHIN       DATA

SOURCE                  SS      df                MS          F        p
================================================================
mean       848042.5628       1       848042.5628    105.922   0.000 ***
s/          40031.4680       5         8006.2936

hem         11567.9655       1        11567.9655      4.215   0.095
hs/         13721.1004       5         2744.2201

f0            506.4155       1          506.4155      1.763   0.242
fs/          1436.0434       5          287.2087

hf             38.4127       1           38.4127      0.194   0.678
hfs/          990.7641       5          198.1528

phone        3719.3065       2         1859.6533      8.135   0.008 **
ps/          2286.0960      10          228.6096

hp            607.4072       2          303.7036      1.086   0.374
hps/         2796.6691      10          279.6669

fp           1796.7979       2          898.3990      4.163   0.048 *
fps/         2158.2373      10          215.8237

hfp           379.4192       2          189.7096      1.630   0.244
hfps/        1164.0102      10          116.4010
```

TABLE 4.21

```
FACTOR:        sub        hem         f0      phone       DATA
LEVELS:          4          2          2          3         48
TYPE  :     RANDOM     WITHIN     WITHIN     WITHIN       DATA

SOURCE                SS     df             MS         F      p
===============================================================
mean         463172.1279      1    463172.1279    58.825  0.005 **
s/            23620.9959      3      7873.6653

hem            3203.7839      1      3203.7839     0.918  0.409
hs/           10475.4326      3      3491.8109

f0              641.3065      1       641.3065     2.108  0.242
fs/             912.7215      3       304.2405

hf               56.7024      1        56.7024     0.264  0.643
hfs/            644.0550      3       214.6850

phone          2429.2346      2      1214.6173    10.133  0.012 *
ps/             719.1952      6       119.8659

hp              162.3222      2        81.1611     0.313  0.742
hps/           1555.3248      6       259.2208

fp              474.8697      2       237.4349     2.779  0.140
fps/            512.5789      6        85.4298

hfp              30.5449      2        15.2725     0.185  0.836
hfps/           496.4905      6        82.7484
```

TABLE 4.22

```
FACTOR:         sub       hem        f0     phone       DATA
LEVELS:           6         2         2         3         72
TYPE  :      RANDOM    WITHIN    WITHIN    WITHIN       DATA

SOURCE                 SS     df              MS          F       p
====================================================================
mean          951142.3100      1     951142.3100    388.005   0.000 ***
s/             12256.8361      5       2451.3672

hem              101.2938      1        101.2938      0.317   0.598
hs/             1599.5045      5        319.9009

f0               259.1606      1        259.1606      0.872   0.393
fs/             1485.6813      5        297.1363

hf                 0.0672      1          0.0672      0.001   0.974
hfs/             281.0112      5         56.2022

phone           4135.8812      2       2067.9406     26.984   0.000 ***
ps/              766.3656     10         76.6366

hp               160.4877      2         80.2439      0.475   0.635
hps/            1688.9692     10        168.8969

fp                 6.1678      2          3.0839      0.335   0.723
fps/              92.1455     10          9.2146

hfp              165.0544      2         82.5272      0.523   0.608
hfps/           1578.8622     10        157.8862
```

TABLE 4.23

```
FACTOR:          sub        hem        f0       phone       DATA
LEVELS:            4          2         2           3         48
TYPE  :       RANDOM     WITHIN    WITHIN      WITHIN       DATA

SOURCE                SS      df              MS          F       p
===================================================================
mean         676352.6142       1     676352.6142    199.156   0.001 ***
s/            10188.2889       3       3396.0963

hem             276.4800       1        276.4800      1.030   0.385
hs/             805.1817       3        268.3939

f0              228.8134       1        228.8134      0.497   0.532
fs/            1381.8519       3        460.6173

hf                4.4408       1          4.4408      0.049   0.838
hfs/            269.2709       3         89.7570

phone          2642.7014       2       1321.3507     18.360   0.003 **
ps/             431.8150       6         71.9692

hp              106.0549       2         53.0275      0.197   0.827
hps/           1617.7387       6        269.6231

fp               12.7616       2          6.3808      0.597   0.580
fps/             64.1083       6         10.6847

hfp             322.1016       2        161.0508      0.850   0.473
hfps/          1136.5517       6        189.4253
```

TABLE 4.24

```
FACTOR:          sub         hem         task        voice       place        DATA
LEVELS:            6           2            2            2           2          96
TYPE  :       RANDOM      WITHIN       WITHIN       WITHIN      WITHIN        DATA

SOURCE                        SS       df                  MS           F         p
======================================================================================
mean            1569634.9364        1     1569634.9364     120.472    0.000 ***
s/                65145.2629        5       13029.0526

hem                5154.7498        1        5154.7498       2.314    0.189
hs/               11136.8864        5        2227.3773

task                181.8852        1         181.8852       0.349    0.580
ts/                2606.0215        5         521.2043

ht                 7525.6869        1        7525.6869      55.307    0.001 ***
hts/                680.3591        5         136.0718

voice              7261.7154        1        7261.7154       2.334    0.187
vs/               15555.5811        5        3111.1162

hv                 1761.1351        1        1761.1351       7.394    0.042 *
hvs/               1190.8893        5         238.1779

tv                 1113.9800        1        1113.9800       8.700    0.032 *
tvs/                640.2450        5         128.0490

htv                 324.6498        1         324.6498       1.017    0.359
htvs/              1595.7537        5         319.1507

place                49.9682        1          49.9682       0.117    0.747
ps/                2143.1408        5         428.6282

hp                  573.7929        1         573.7929       4.309    0.093
hps/                665.7895        5         133.1579

tp                    0.7740        1           0.7740       0.002    0.963
tps/               1624.9697        5         324.9939

htp                  12.9213        1          12.9213       0.094    0.771
htps/               684.7554        5         136.9511

vp                  414.0872        1         414.0872       6.669    0.049 *
vps/                310.4440        5          62.0888

hvp                 328.0422        1         328.0422       1.270    0.311
hvps/              1291.4654        5         258.2931

tvp                   2.0242        1           2.0242       0.050    0.831
tvps/               200.8177        5          40.1635

htvp                607.7250        1         607.7250       5.265    0.070
htvps/              577.1102        5         115.4220
```

TABLE 4.25

```
FACTOR:        sub       hem       task      voice     place       DATA
LEVELS:          4         2          2          2         2          64
TYPE  :     RANDOM    WITHIN     WITHIN     WITHIN    WITHIN        DATA

SOURCE                  SS     df              MS         F        p
=================================================================
mean           978380.6324      1     978380.6324    59.240   0.005 **
s/              49546.8675      3      16515.6225

hem              2678.4505      1       2678.4505     1.650   0.289
hs/              4868.9664      3       1622.9888

task              337.4108      1        337.4108     0.719   0.459
ts/              1407.2653      3        469.0884

ht               6149.1076      1       6149.1076    39.957   0.008 **
hts/              461.6775      3        153.8925

voice            1560.3483      1       1560.3483     0.623   0.488
vs/              7519.4061      3       2506.4687

hv               1573.2128      1       1573.2128     6.498   0.084
hvs/              726.3554      3        242.1185

tv                191.2343      1        191.2343     6.460   0.085
tvs/               88.8047      3         29.6016

htv                 2.0556      1          2.0556     0.008   0.933
htvs/             743.1332      3        247.7111

place               8.4754      1          8.4754     0.014   0.912
ps/              1775.5715      3        591.8572

hp                882.4614      1        882.4614     9.957   0.051
hps/              265.8927      3         88.6309

tp                449.7049      1        449.7049    11.930   0.041 *
tps/              113.0813      3         37.6938

htp                78.6548      1         78.6548     1.179   0.357
htps/             200.1551      3         66.7184

vp                128.6806      1        128.6806     7.672   0.070
vps/               50.3188      3         16.7729

hvp               310.7729      1        310.7729     4.956   0.112
hvps/             188.1119      3         62.7040

tvp                60.8205      1         60.8205     2.820   0.192
tvps/              64.6958      3         21.5653

htvp               68.9523      1         68.9523     1.351   0.329
htvps/            153.1552      3         51.0517
```

TABLE 4.26

```
FACTOR:        sub        hem        task       voice      place        DATA
LEVELS:          6          2           2          2          2           96
TYPE  :     RANDOM     WITHIN      WITHIN     WITHIN     WITHIN         DATA

SOURCE                 SS      df             MS          F        p
==================================================================
mean          958980.2789       1    958980.2789   1153.934    0.000 ***
s/              4155.2644       5       831.0529

hem             1488.6900       1      1488.6900     12.214    0.017 *
hs/              609.4420       5       121.8884

task             192.4968       1       192.4968      1.284    0.309
ts/              749.6177       5       149.9235

ht               638.6019       1       638.6019      5.285    0.070
hts/             604.2145       5       120.8429

voice              2.2143       1         2.2143      0.122    0.741
vs/               90.5703       5        18.1141

hv                29.1721       1        29.1721      2.362    0.185
hvs/              61.7637       5        12.3527

tv                17.8365       1        17.8365      1.737    0.245
tvs/              51.3472       5        10.2694

htv                8.0273       1         8.0273      0.516    0.505
htvs/             77.7826       5        15.5565

place              2.0768       1         2.0768      0.041    0.847
ps/              252.0661       5        50.4132

hp                 3.4884       1         3.4884      0.201    0.673
hps/              86.7006       5        17.3401

tp                 0.2166       1         0.2166      0.101    0.763
tps/              10.7018       5         2.1404

htp                5.0692       1         5.0692      0.455    0.530
htps/             55.7604       5        11.1521

vp                40.1968       1        40.1968      5.382    0.068
vps/              37.3418       5         7.4684

hvp                0.4401       1         0.4401      0.017    0.903
hvps/            133.1806       5        26.6361

tvp                5.2641       1         5.2641      0.078    0.791
tvps/            335.5762       5        67.1152

htvp               9.0405       1         9.0405      4.404    0.090
htvps/            10.2647       5         2.0529
```

TABLE 4.27

```
FACTOR:          sub        hem        task      voice      place       DATA
LEVELS:            4          2           2          2          2         64
TYPE  :       RANDOM     WITHIN      WITHIN     WITHIN     WITHIN       DATA

SOURCE                SS      df               MS          F        p
=====================================================================
mean         657376.3708       1      657376.3708    746.889    0.000 ***
s/             2640.4591       3         880.1530

hem             831.1689       1         831.1689      4.315    0.129
hs/             577.8769       3         192.6256

task             10.0331       1          10.0331      0.113    0.759
ts/             266.6001       3          88.8667

ht              222.7557       1         222.7557      4.799    0.116
hts/            139.2543       3          46.4181

voice            13.1225       1          13.1225      0.929    0.406
vs/              42.3709       3          14.1236

hv                5.8322       1           5.8322      0.508    0.527
hvs/             34.4223       3          11.4741

tv                4.7198       1           4.7198      0.370    0.586
tvs/             38.2627       3          12.7542

htv               3.6864       1           3.6864      0.147    0.727
htvs/            75.0072       3          25.0024

place             3.4782       1           3.4782      0.042    0.851
ps/             250.6134       3          83.5378

hp               12.8702       1          12.8702      0.577    0.503
hps/             66.9077       3          22.3026

tp                0.0004       1           0.0004      0.000    0.989
tps/              5.3517       3           1.7839

htp               1.5438       1           1.5438      1.211    0.352
htps/             3.8245       3           1.2748

vp               11.1556       1          11.1556      1.257    0.344
vps/             26.6168       3           8.8723

hvp              24.8751       1          24.8751      1.631    0.291
hvps/            45.7593       3          15.2531

tvp              52.3452       1          52.3452      0.749    0.451
tvps/           209.7287       3          69.9096

htvp              9.6565       1           9.6565      3.231    0.170
htvps/            8.9674       3           2.9891
```

TABLE 4.30

```
FACTOR:        sub        hem       voice       place        DATA
LEVELS:          6          2           2           2          48
TYPE  :     RANDOM     WITHIN      WITHIN      WITHIN        DATA

SOURCE                 SS      df             MS          F        p
==============================================================
mean         768011.8618        1    768011.8618    112.978    0.000 ***
s/            33989.5480        5      6797.9096

hem             111.8130        1       111.8130      0.089    0.777
hs/            6247.2546        5      1249.4509

voice          1343.6596        1      1343.6596      0.949    0.375
vs/            7075.7919        5      1415.1584

hv              286.7495        1       286.7495      1.123    0.338
hvs/           1276.4220        5       255.2844

place            19.1521        1        19.1521      0.064    0.810
ps/            1490.8119        5       298.1624

hp              379.4627        1       379.4627     16.255    0.010 *
hps/            116.7195        5        23.3439

vp              237.0073        1       237.0073      8.958    0.030 *
vps/            132.2826        5        26.4565

hvp             914.3805        1       914.3805      5.555    0.065
hvps/           823.0857        5       164.6171
```

TABLE 4.31

```
FACTOR:          sub        hem      voice      place       DATA
LEVELS:            4          2          2          2         32
TYPE  :       RANDOM     WITHIN     WITHIN     WITHIN       DATA

SOURCE                SS      df              MS          F      p
===============================================================
mean         471189.9220       1     471189.9220     66.385  0.004 **
s/            21293.4997       3       7097.8332

hem             355.4444       1        355.4444      0.460  0.546
hs/            2316.2206       3        772.0735

voice           329.5385       1        329.5385      0.263  0.643
vs/            3752.1142       3       1250.7047

hv              844.5022       1        844.5022     13.114  0.036 *
hvs/            193.1882       3         64.3961

place           290.8268       1        290.8268      0.976  0.396
ps/             893.9882       3        297.9961

hp              217.1007       1        217.1007      5.833  0.095
hps/            111.6611       3         37.2204

vp              183.2177       1        183.2177      8.378  0.063
vps/             65.6043       3         21.8681

hvp             336.2474       1        336.2474     11.138  0.044 *
hvps/            90.5636       3         30.1879
```

TABLE 4.32

```
FACTOR:         sub        hem      voice      place       DATA
LEVELS:           6          2          2          2         48
TYPE  :      RANDOM     WITHIN     WITHIN     WITHIN       DATA

SOURCE                 SS     df             MS         F      p
================================================================
mean          801804.9598      1    801804.9598   118.745  0.000 ***
s/             33761.7364      5      6752.3473

hem            12568.6237      1     12568.6237    11.282  0.020 *
hs/             5569.9910      5      1113.9982

voice           7032.0358      1      7032.0358     3.855  0.107
vs/             9120.0341      5      1824.0068

hv              1799.0354      1      1799.0354     5.956  0.059
hvs/            1510.2210      5       302.0442

place             31.5901      1        31.5901     0.069  0.803
ps/             2277.2986      5       455.4597

hp               207.2516      1       207.2516     0.840  0.401
hps/            1233.8254      5       246.7651

vp               179.1041      1       179.1041     2.363  0.185
vps/             378.9791      5        75.7958

hvp               21.3867      1        21.3867     0.102  0.762
hvps/           1045.4899      5       209.0980
```

TABLE 4.33

```
FACTOR:       sub       hem      voice     place     DATA
LEVELS:         4         2          2         2       32
TYPE  :    RANDOM    WITHIN     WITHIN    WITHIN     DATA

SOURCE             SS     df            MS          F      p
=================================================================
mean      507528.1212      1   507528.1212     51.334  0.006 **
s/         29660.6331      3     9886.8777

hem         8472.1137      1     8472.1137      8.432  0.062
hs/         3014.4233      3     1004.8078

voice       1422.0441      1     1422.0441      1.106  0.370
vs/         3856.0966      3     1285.3655

hv           730.7663      1      730.7663      1.718  0.281
hvs/        1276.3005      3      425.4335

place        167.3535      1      167.3535      0.505  0.529
ps/          994.6646      3      331.5549

hp           744.0155      1      744.0155      6.298  0.087
hps/         354.3867      3      118.1289

vp             6.2835      1        6.2835      0.382  0.581
vps/          49.4102      3       16.4701

hvp           43.4778      1       43.4778      0.520  0.523
hvps/        250.7035      3       83.5678
```

Adams, R., & Victor, M. (1989). Principles of neurology (4 ed.). New York: McGraw-Hill.

Albert, M., & Bear, D. (1974). Time to understand. A case study of word deafness with reference to the role of time in auditory comprehension. Brain, 97, 373-384.

Albert, M., Sparks, R., von Stockert, T., & Sax, D. (1972). A case study of auditory agnosia: linguistic and nonlinguistic processing. Cortex, 8, 427-443.

Auerbach, S., Allard, T., Naeser, M., Alexander, M., & Albert, M. (1982). Pure word deafness: analysis of a case with bilateral lesions and a defect at the prephonemic level. Brain, 105, 271-300.

Awh, E., Schumacher, E., Lauber, E., Jonides, J., Smith, E., & Koeppe, R. (1995). Modality independence of verbal working memory: PET evidence. In Cognitive Neuroscience Society Annual Meeting, San Francisco:

Baddeley, A. (1992). Working memory. Science, 255, 556-559.

Ballet, G. (1903). Un cas de surdite verbale. Revue Neurologique, 11, 685-688.

Barrett, A. (1910). A case of pure word deafness with autopsy. Journal of Nervous and Mental Disease, 37, 73-92.

Bartlett, E. J., Brown, J. W., Wolf, A. P., & Brodie, J. D. (1987). Correlations between glucose metabolic rates in brain regions of healthy male adults at rest and during language stimulation. Brain and Language, 32, 1-18.

Bastian, H. (1897). Some problems in connexion with aphasia and other speech defects. Lancet, 1, 933-942.

Baynes, K. (1990). Language and reading in the right hemisphere: highways or byways of the brain? Journal of Cognitive Neuroscience, 2(3), 159-179.

Binder, J., Rao, S., Hammeke, T., Frost, J., Bandettini, P., Jesmanowicz, A., & Hyde, J. (1995). Lateralized human brain language systems demonstrated by task subtraction functional magnetic resonance imaging. Archives of Neurology, 52, 593-601.

Binder, J., Rao, S., Hammeke, T., Yetkin, F., Jesmanowicz, A., Bandettini, P., Wong, E., Estkowski, L., Goldstein, M., Haughton, V., & Hyde, J. (1994). Functional magnetic resonance imaging of human auditory cortex. Annals of Neurology, 35, 662-672.

Boatman, D., Lesser, R., & Gordon, B. (in press). Auditory speech perception in the left temporal lobe: an electrical interference study. Brain & Language.

Bonvincini, G. (1905). Über subcorticale sensorische Aphasie. Jahrbücher für Psychiatrie und Neurologie, 26, 126-229.

Bramwell, E. (1927). A case of cortical deafness. Brain, 50, 579-580.

Brick, J., Frost, J., Schochet, S., Gutmann, L., & Crosby, T. (1985). Pure word deafness: CT localization of the pathology. Neurology, 35, 441-442.

Broca, P. (1861). Remarques sur le siege de la faculte de la parole articulee, suives d'une observacion d'aphemie (perte de parole). Bulletin de la Societe d'Anatomie, 36, 330-357.

Buchman, A., Garron, D., Trost-Cardamone, J., Wichter, M., & Schwartz, M. (1986). Word deafness: one hundred years later. J Neurol Neurosurg Psychiatry, 49(5), 489-499.

Cansino, S., Williamson, S., & Karron, D. (1994). Tonotopic organization of human auditory cortex. Brain Research.

Caplan, D. (1992). Language: structure, processing, and disorders. Cambridge: MIT Press.

Caplan, D. (1994). Language and the brain. In M. Gernsbacher (Eds.), Handbook of Psycholinguistics (pp. 1023-1053). New York: Academic Press.

Caplan, D., Gow, D., & Makris, N. (1995). Analysis of lesions by MRI in stroke patients with acoustic-phonetic processing deficits. Neurology, 45, 293-298.

Caplan, D., & Waters, G. (1994). Articulatory length and phonological similarity in span tasks: a reply to Baddeley and Andrade. Quarterly Journal of Experimental Psychology, 47(4), 1055-1062.

Caplan, D., & Waters, G. (1995). On the nature of phonological output planning processes involved in verbal rehearsal: evidence from aphasia. Brain & Language, 48(2), 191-220.

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

Chocolle, R., Chedru, F., Botte, M., Chain, F., & Lhermitte, F. (1975). Etude psychoacoustique d'un cas de 'surdite corticale'. Neuropsychologia, 13, 163-172.

Chomsky, N. (1959). A review of Skinner's Verbal Behavior. Language, 35(1), 26-58.

Chomsky, N. (1986). Knowledge of Language. New York: Praeger.

Chomsky, N. (1992). A minimalist program for linguistic theory. MIT Occasional Papers in Linguistics, 1, 1-67.

Clarey, J., Barone, P., & Imig, T. (1992). Physiology of thalamus and cortex. In A. Popper & R. Fay (Eds.), The Mammalian Auditory Pathway: Neurophysiology (pp. 232-334). New York: Springer-Verlag.

Cohen, R., Semple, W., Gross, M., King, A., & Nordahl, T. (1992). Metabolic brain pattern of sustained auditory discrimination. Experimental Brain Research, 92, 165-172.

Colombo, M., D'Amato, M., Rodman, H., & Gross, C. (1990). Auditory association cortex lesions impair auditory short-term memory in monkeys. Science, 247, 336-339.

Coslett, H., Brashear, H., & Heilman, K. (1984). Pure word deafness after bilateral primary auditory cortex infarcts. Neurology, 34, 347-352.

Creutzfeldt, O., Ojemann, G., & Lettich, E. (1989). Neuronal activity in the human lateral temporal lobe. Experimental Brain Research, 77, 451-475.

Cuffin, B. N. (1990). Effects of Head Shape on EEG's and MEG's. IEEE Transaction on Biomedical Engineering, 37(1), 52.

Cuffin, B. N. (1993). Effects of Local Variations in Skull and Scalp Thickness on EEG's and MEG's. IEEE Transaction on Biomedical Engineering, 40(1), 42-48.

Cuffin, B. N. (1995). A Method for Localizing EEG Sources in Realistic Head Models. IEEE Transaction on Biomedical Engineering, 42(1), 68-71.

Dale, A., & Sereno, M. (1993). Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. Journal of Cognitive Neuroscience, 5(2), 162-176.

Dalla, P., Spinnler, H., & Vallar, G. (1979). Pure word deafness and bilateral posterior perisylvian softenings. Arch. Suisses Neurol. Neurochi. Psychiatry, 125, 47-58.

Dejerine, J., & Serieux, P. (1898). Un cas de surdite verbale pure terminee par aphasie sensorielle. Revue de Psychiatrie, 2, 7-11.

Démonet, J., Price, C., Wise, R., & Frackowiak, R. (1994). Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects. Neuroscience Letters, 182, 25-28.

Démonet, J., Price, C., Wise, R., & Frackowiak, R. (1994). A PET study of cognitive strategies in normal subjects during language tasks: influence of phonetic ambiguity and sequence processing o phoneme monitoring. Brain, 117, 671-682.

Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A., & Frackowiak, R. (1992). The anatomy of phonological and semantic processing in normal subjects. Brain, 115, 1753-1768.

Denes, G., & Semenza, C. (1975). Auditory modality-specific anomia: evidence from a case of pure word deafness. Cortex, 11, 401-411.

Di Giovanni, M., D'Alessandro, G., Baldini, S., Cantalupi, D., & Bottacchi, E. (1992). Clinical and neuroradiological findings in a case of pure word deafness. Ital J Neurol Sci, 13, 507-510.

Doyle, P., & Holland, A. (1982). Clinical management of a patient with pure word deafness. In R. Brookshire (Eds.), Clinical aphasiology (pp. 138-146).

Dronkers, N., Redfern, B., & Ludy, C. (1995). Lesion localization in chronic Wernicke's aphasia. In Cognitive Neuroscience Society Annual Meeting, San Francisco, CA.

Dronkers, N., Shapiro, J., Redfern, B., & Knight, R. (1992). The role of Broca's area in Broca's aphasia. Journal of Clinical and Experimental Neuropsychology, 14, 52-53.

Duhamel, J. R., & Poncet, M. (1986). Deep dysphasia in a case of phonemic deafness: role of the right hemisphere in auditory language comprehension. Neuropsychologia, 24, 769-779.

Earnest, M., Monroe, P., & Yarnell, M. (1977). Cortical deafness: demonstration of the pathologic anatomy by CT scan. Neurology, 27, 1172-1175.

Elberling, C., Bak, C., Kofoed, B., Lebech, J., & Saermark, K. (1982). Auditory magnetic fields: source location and tonotopical organization in the right hemisphere of the human brain. Scand Audiol, 11, 61-65.

Elliott, D., & Trahiotis, C. (1972). Cortical lesions and auditory discrimination. Psychological Bulletin, 77(3), 198-222.

Elliott, L. (1994). Functional brain imaging and hearing. Journal of the Acoustic Society of America, 96(3), 1397-1408.

Eulitz, C., Diesch, E., Pantev, C., Hampson, S., & Elbert, T. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. The Journal of Neuroscience, 15(4), 2748-2755.

Evans, E. (1992). Auditory processing of complex sounds: an overview. Philosophical Transactions of the Royal Society of London, Biology, 336, 295-306.

Fiez, J., Tallal, P., Raichle, M., Katz, W., & Petersen, S. (in press). Studies of auditory and phonological processing: effects of stimulus characteristics and task demands. Journal of Cognitive Neuroscience.

Fodor, J. (1983). The modularity of mind. Cambridge MA: MIT Press.

Forster, K., & Bednall, E. (1976). Terminating and exhaustive search in lexical access. Memory & Cognition, 4, 53-61.

Fox, P. T., Mintun, M. A., Reiman, E. M., & Raichle, M. E. (1988). Enhanced detection of focal brain responses using inter-subject averaging and change-distribution analysis of subtracted PET images. Journal of Cerebral Blood Flow and Metabolism, 8, 642-653.

Freud, S. (1891). Zur Auffassung der Aphasien. Wien: Deuticke.

Freund, C. (1895). Labyrinth der Taubheit und Sprachtaubheit. Wiesbaden: Bergmann.

Frith, C. D., Friston, K. J., Liddle, P. F., & Frackowiak, R. S. J. (1991). A PET study of word finding. Neuropsychologia, 29(12), 1137-1148.

Fujii, T., Fukatsu, R., Watabe, S., Ohnuma, A., Teramura, K., Kimura, I., Saso, S., & Kogure, K. (1990). Auditory sound agnosia without aphasia following a right temporal lobe lesion. Cortex, 26, 263-268.

Gallistel, C. (1990). The Organization of Learning. Cambridge, MA: MIT Press.

Gazzaniga, M. (Ed.). (1995). The Cognitive Neurosciences. Cambridge, MA: MIT Press.

Gazzaniga, M., Glass, A., Sarno, M., & Posner, J. (1973). Pure word deafness and hemispheric dynamics: a case history. Cortex, 9, 136-143.

Gazzaniga, M. S. (1983). Right hemisphere language following brain bisection: a twenty year perspective. American Psychologist, 38, 525-547.

Geschwind, N. (1970). The organization of language and the brain. Science, 170, 940-944.

Geschwind, N. (1979). Language and the brain. Scientific American, 241, 180-199.

Geschwind, N., & Galaburda, A. (1987). Cerebral Lateralization. Cambridge MA: MIT Press.

Gevins, A. (1993). High-resolution EEG enters imaging arena. Diagnostic Imaging, 78.

Giraudeau, D. (1882). Note sur un cas de surdite cerebrale. Revue de Medicine, 2, 448-452.

Goldstein, M. (1974). Auditory agnosia for speech ('pure word deafness'): a historical review with current implications. Brain and Language, 1, 195-204.

Goldstein, M., Brown, M., & Hollander, J. (1975). Auditory agnosia and cortical deafness: analysis of a case with three-year followup. Brain and Language, 2, 324-332.

Gordon, W., & Naeser, M. (1977). Frequency discrimination deficits in three word-deaf patients with bilateral temporal lobe lesions. In Academy of Aphasia Metting, . Montreal, Canada:

Graham, J., Greenwood, R., & Lecky, B. (1980). Cortical deafness: a case report and review of the literature. J. Neurological Sciences, 48, 35-49.

Halliday, D., & Resnick, R. (1988). Fundamentals of Physics. New York: Wiley.

Hamanaka, H., Asano, K., Morimune, S., & Seko, K. (1980). Ein Fall von reiner Worttaubheit ohne akustische Agnosie. Studia Phonologica, 14, 16-24.

Hari, R. (1991). Activation of the human auditory cortex by speech sounds. Acta Otolaryngol (Stockh), Suppl. 491, 132-138.

Hari, R., Pelizzone, M., Mäkelä, J., Hällström, J., Leinonen, L., & Lounasama, O. (1987). Neuromagnetic responses of the human auditory cortex to on- and offsets of noise bursts. Audiology, 26, 31-43.

Hayashi, M. (1993). Auditory neuromagnetic fields evoked by spectral transition of syllables. Journal of Robotics and Mechatronics, 5(4), 409-412.

Hayashi, M., Mashiko, T., Imada, T., & Odaka, K. (1992). Brain magnetic fields evoked by five japanese vowels. ICA, 14, 2-3.

Heffner, H., & Heffner, R. (1984). Temporal lobe lesions and perception of species-specific vocalizations by macaques. Science, 226, 75-76.

Heffner, H., & Heffner, R. (1986). Effect of unilateral and bilateral auditory cortex lesions on the discrimination of vocalizations by Japanese macaques. Journal of Neurophysiology, 56(3), 683-701.

Heffner, H., & Heffner, R. (1986). Hearing loss in Japanese macaques following bilateral auditory cortex lesions. Journal of Neurophysiology, 55(2), 256-271.

Heffner, H., & Heffner, R. (1989). Cortical deafness cannot account for the inability of Japanese maxaques to discriminate species-specific vocalizations. Brain and Language, 36(2), 275-285.

Heintze, H., Mangun, G., Burchert, W., Hinrichs, H., Scholz, M., Münte, T., Gös, A., Scherg, M., Johannes, S., Hundeshagen, H., Gazzaniga, M., & Hillyard, S. (1994). Combined spatial and temporal imaging of brain activity during visual selective attention in humans. Nature, 372, 543-546.

Hemphill, R., & Stengel, E. (1940). A study on pure word deafness. Journal of Neurology and Psychiatry, 3, 251-262.

Hinke, R., Hu, X., Stillman, A., Kim, S., Merkle, H., Salmi, R., & Ugurbil, K. (1993). Functional magnetic resonance imaging of Broca's area during internal speech. NeuroReport, 4, 675-678.

Horenstein, S., & Bealmear, K. (1973). Concerning the nature of word deafness. Trans. Am. Neurol. Assoc., 98, 264-267.

Howard, D., Patterson, K., Wise, R., Brown, W., Friston, K., Weiller, C., & Frackowiak, R. (1992). The cortical localization of the lexicons: positron emission tomography evidence. Brain, 115, 1769-1782.

Hubel, D. (1988). Eye, brain, and vision. New York: Freeman.

Hugdahl, K. (1992). Brain lateralization: dichotic studies. In G. Adelman (Eds.), The Encyclopedia of Neuroscience Boston: Birkhäuser.

Hupfer, K., Jürgens, U., & Ploog, D. (1977). The effect of superior temporal lesions on the recognition of species-specific calls in the squirrel monkey. Experimental Brain Research, 30, 75-87.

Ikemura, Y. (1983). The development of the study of pure word deafness. Psychiatry, 25, 351-361.

Iversen, S., & Mishkin, M. (1973). Comparison of superior temporal and inferior prefrontal lesions on auditor and non-auditory tasks in rhesus monkeys. Brain Research, 55, 355-367.

Jackendoff, R. (1994). Patterns in the Mind. New York: HarperCollins.

Jerger, J., Lovering, L., & Wertz, M. (1972). Auditory disorder following bilateral temporal lobe insult: report of a case. J. Speech and Hearing Disorders, 37(4), 523-535.

Jerger, J., Weikers, N., Sharbrough, F., & Jerger, S. (1969). Bilateral lesions of the temporal lobe: a case study. Acta Oto-Laryngologica(Suppl. 258), 1-51.

Jones, J., & Dinolt, R. (1952). Pure word deafness. Laryngoscope, 62, 194-203.

Kamei, H., Nakane, K., Nishimaru, K., Shiraishi, K., & Matuo, M. (1981). Pure word deafness after cerebrovascular disease: a case report. Clinical Neurology (Tokyo), 21, 402-408.

Kandel, E., Schwartz, J., & T., J. (Ed.). (1991). Principles of Neural Science (3 ed.). New York: Elsevier.

Kanshepolsky, J., Kelley, J., & Waggener, J. (1973). A cortical auditory disorder: clinical, audiologic, and pathologic aspects. Neurology, 23, 699-705.

Kanter, S. L., Day, A. L., Heilman, K. M., & Gonzalez-Rothi, L. J. (1986). Pure word deafness: a possible explanation of transient deteriorations after extracranial-intracranial bypass grafting. Neurosurgery, 18(2), 186-189.

Kaukoranta, E., Hari, R., & Lounasmaa, O. (1987). Responses of the human auditory cortex to vowel onset after fricative consonants. Experimental Brain Research, 69, 19-23.

Kazui, S., Naritomi, H., Sawada, T., Inoue, N., & Okuda, J.-I. (1990). Subcortical auditory agnosia. Brain and Language, 38, 476-487.

Kirshner, H., & Webb, W. (1981). Selective involvement of the auditory-verbal modality in an acquired communication disorder: benefit from sign language therapy. Brain and Language, 13, 161-170.

Kirshner, H., Webb, W., & Duncan, G. (1981). Word deafness in Wernicke's aphasia. J. Neurol. Neurosurg. Psychiatry, 44, 197-201.

Kitayama, I., Yamazaki, K., Shibahara, K., & Nomura, J. (1990). Pure word deafness with possible transfer of language dominance. Jpn J Psychiatry Neurol, 44(3), 577-584.

Klein, D., Milner, B., Zatorre, R., Meyer, E., & Evans, A. (1995). The neural substrates underlying word generation: a bilingual functional-imaging study. Proceedings of the National Academy of Sciences of the USA, 92, 2899-2903.

Klein, R., & Harper, J. (1956). The problem of agnosia in light of a case of pure word deafness. J. Ment. Sci, 102, 112-120.

Kleist, K. (1962). Sensory aphasia and amusia. New York: Pergamon.

Knopman, D., Rubens, A., Klassen, A., Meyer, M., & Niccum, N. (1980). Regional cerebral blood flow patterns during verbal and nonverbal auditory activation. Brain and Language, 9, 93-112.

Kolb, B., & Wishaw, I. (1990). Fundamentals of Human Neuropsychology (3 ed.). New York: Freeman.

Krauland-Steinbereithner, F., & Weingarten, K. (1956). Beitrag zur Klinik der reinen Worttaubheit. Winer Zeitschrift für Nervenheilkunde, 12, 452.

Kuhl, P. (1994). Learning and representation in speech and language. Current Opinion in Neurobiology, 4, 812-822.

Kuriki, S., & Murase, M. (1989). Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. Experimental Brain Research, 77, 127-134.

Kuriki, S., Okita, Y., & Hirata, Y. (1995). Source analysis of magnetic field responses from the human auditory cortex elicited by short speech sounds. Experimental Brain Research, 104, 144-152.

Kushner, M., Rosenquist, A., Alavi, A., Rosen, B., & Dann, R. (1987). Cerebral activation by nonmeaningful monaural verbal auditory stimulation. Brain Research, 409, 79-87.

Kussmaul, A. (1877). Disturbances of speech. In H. von Ziemssen (Eds.), Cyclopedia of the practice of medicine (pp. 581-875). New York: William Wood.

Larsen, B., Skinhøj, E., & Lassen, N. (1978). Variations in regional cortical blood flow in the right and left cerebral hemispheres during automatic speech. Brain, 101, 192-209.

Lauter, J., Herscovitch, P., Formby, C., & Raichle, M. (1985). Tonotopic organization in human auditory cortex revealed by positron emission tomography. Hearing Research, 20, 199-205.

LeGros Clark, W., & Russell, W. (1938). Cortical deafness without aphasia. Brain, 61, 373-383.

Lemoyne, J., & Mahoudeau, D. (1959). A propos un cas diagnosie auditive pure avec surdite cortiale associee a une dysphonie functionelle. Annales D'otolaryngologie, 76, 293-310.

Lenneberg, E. (1967). Biological Foundations of Language. New York: Wiley.

Lewine, J., & Orrison, W. (1995). Magnetoencephalography and magnetic source imaging. In W. Orrison, J. Lewine, J. Sanders, & M. Hartshorne (Eds.), Functional Brain Imaging St. Louis: Mosby.

Lhermitte, F., Chain, F., Escourolle, R., Ducarne, B., Pillon, P., & Chedru, F. (1971). Etude des troubles perceptifs auditifs dans les lesions temporales bilaterales. Revue neurologique, 124, 329.

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. Cognition, 21, 1-36.

Lichtheim, L. (1885). On aphasia. Brain, 7, 433-484.

Liepmann, H. (1898). Ein Fall von reiner Sprachtaubheit. In C. Wernicke (Eds.), Psychiatrische Abhandlungen Breslau: Schletter:

Lund, E., Spliid, P., Andersen, E., & Bojsen-Moller, A. (1986). Vowel perception: a neuroradiological study of the perception of vowels in the human cortex. Brain and Language, 25, 191-211.

Maddock, R., Buonocore, M., Lisheng, G., & Hordahl, T. (1994). Cortical activation responses to the auditory presentation of words. Society for Neuroscience Abstracts, 20, 352.

Mäkelä, J., Ahonen, A., Hämäläinen, M., Hari, R., Ilmoniemi, R., Kajola, M., Knuutila, J., Lounasmaa, O., McEvoy, L., Salmelin, R., Salonen, O., Sams, M., Simola, J., Tesche, C., & Vasama, J. (1993). Functional differences between auditory cortices of the two hemispheres revealed by whole-head neuromagnetic recordings. Human Brain Mapping, 1, 48-56.

Mäkelä, J., Hari, R., & Leinonen, L. (1988). Magnetic responses of the human auditory cortex to noise/square wave transitions. Electroencephalography and Clinical Neurophysiology, 69, 423-430.

Marr, D. (1982). Vision. New York: Freeman.

Masterson, R. (1992). Role of the central auditory system in hearing: the new direction. Trends in Neurosciences, 15(8), 280-285.

Mattingly, I., & Liberman, A. (1988). Specialized perceiving systems for speech and other biologically significant sounds. In G. Edelman, W. Gall, & W. Cowan (Eds.), Auditory Function: Neurobiological Bases of Hearing (pp. 775-793). New York: John Wiley and Sons.

Mazoyer, B. M., Dehaene, S., Tzourio, N., Frak, J., Murayama, N., Cohen, L., Levrier, O., Salamon, G., Syrota, A., & Mehler, J. (1993). The cortical representation of speech. Journal of Cognitive Neuroscience, 5(4), 467-479.

Mazziotta, J., Phelps, M., Carson, R., & Kuhl, D. (1982). Tomographic mapping of human cerebral metbolism: auditory stimulation. Neurology, 32, 921-937.

McLaughlin, T., Steinberg, B., Christensen, B., Law, I., Parving, A., & Friberg, L. (1992). Potential language and attentional networks revealed through factor analysis of rCBF data measured with SPECT. Journal of Cerebral Blood Flow and Metabolism, 12, 535-545.

Mehler, J., Dupoux, E., Pallier, C., & Dehaene-Lambertz, G. (1994). Cross-linguistic approaches to speech processing. Current Opinion in Neurobiology, 4, 171-176.

Mendez, M., & Geehan, G. (1988). Cortical auditory disorders: clinical and psychoacoustic features. J Neurol Neurosurg Psychiatry, 51(1), 1-9.

Mendez, M., & Rosenberg, S. (1991). Word deafness mistaken for Alzheimer's Disease: differential characteristics. Journal of the American Geriatrics Society, 39, 209-211.

Metz-Lutz, M., & Dahl, E. (1984). Analysis of word comprehension in a case of pure word deafness. Brain and Language, 23, 13-25.

Miceli, G. (1982). The processing of speech sounds in a patient with cortical auditory disorder. Neuropsychologia, 20, 5-20.

Michel, F., Peronnet, F., & Schott, B. (1980). A case of cortical deafness: clinical and electrophysiological data. Brain and Language, 10, 367-377.

Miller, J. (1995). Speech perception: from signal to word. Annual Review of Psychology, 46, 467-492.

Miller, J., & Jusczyk, P. (1989). Seeking the neurobiological bases of speech perception. Cognition, 33, 111-137.

Mills, C. (1891). On the localisation of the auditory centre. Brain, 14, 465-472.

Motomura, N., Yamadori, A., Mori, E., & Tamaru, F. (1986). Auditory agnosia: analysis of a case with bilateral subcortical lesions. Brain, 109, 379-391.

Mott, F. (1907). Bilateral lesion of the auditory cortical centre: complete deafness and aphasia. Br. Med. J., 2, 310-315.

Mountcastle, V. (1978). An organizing principle for cerebral function: the unit module and the distributed system. In F. O. Schmitt (Eds.), The Mindful Brain (pp. 7-50). Cambridge: MIT Press.

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

Näätänen, R. (1992). Attention and Brain Function. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Näätänen, R., Ilmoniemi, R., & Alho, K. (1994). Magnetoencephalography in studies of human cognitive brain function. Trends in Neurosciences, 17(9), 389-395.

Näätänen, R., & Picton, T. (1987). The N1 wave of the human electric and magnetic response to sound: a review and analysis of the component structure. Psychophysiology, 24(4), 375-425.

Nagafuchi, M., Iinuma, K., Yamamoto, K., & Kitahara, T. (1993). Diazepam therapy of verbal auditory agnosia. Brain and Language, 45, 180-188.

Nagafuchi, M., & Suzuki, J. (1973). Auditory agnosia due to incision of splenium corporis callosi. Acta Otolaryngol, 76, 109-113.

Nakasato, N., Fujita, S., Seki, K., Kawamura, T., Matani, A., Tamura, I., Fujiwara, S., & Yoshimoto, T. (1995). Functional localization of bilateral auditory cortices using an MRI-linked whole head magnetoencephalography (MEG) system. Electroencephalography and Clinical Neurophysiology, 94, 183-190.

Nishioka, H., Takeda, Y., Koba, T., Yano, J., Ohiwa, Y., J, H., & Ito, H. (1993). [A case of cortical deafness with bilateral putaminal hemorrhage]. No Shinkei Geka, 21(3), 269-272.

Nishizawa, Y., Olsen, T., Larsen, B., & Lassen, N. (1982). Left-right cortical asymmetries of regional cerebral blood flow during listening to words. J. Neurophysiology, 48, 458-466.

Notoya, M., Suzuki, S., Furukawa, M., & Enokido, H. (1991). A case of pure word deafness associated with Landau-Kleffner syndrome: a long-term study of auditory disturbance. Auris Nasus Larynx, 18, 297-305.

O'Leary, D., Andreasen, N., Hurtig, R., Hichwa, R., Watkins, G., Boles, G., Ponto, L., Rogers, M., & Kirchner, P. (in press). A positron emission tomography study of binaurally and dichotically presented stimuli: effects of language and directed attention. Brain and language.

Ojemann, G. (1991). The cortical organization of language. Journal of Neuroscience, 11(8), 2281-2287.

Okada, S., Hanada, M., Hattori, H., & Shoyama, T. (1963). A`case of pure word deafness (About the relation between auditory perception and recognition of speech sound). Studia Phonologica, 14, 58-65.

Oppenheimer, D., & Newcombe, F. (1978). Clinical and anatomic findings in a case of auditory agnosia. Arch. Neurol., 35, 712-719.

Orrison, W., Lewine, J., Sanders, J., & Hartshorne, M. (Ed.). (1995). Functional Brain Imaging. St. Louis: Mosby.

Pantev, C., Eulitz, C., Elbert, T., & Hoke, M. (1994). The auditory evoked sustained field: origin an frequency dependence. Electroencephalography and Clincal Neurophysiology, 90, 82-90.

Pantev, C., Hoke, M., Lehnertz, K., & Lütkenhöner, B. (1989). Neuromagnetic evidence of an amplitopic organization of the human auditory cortex. Electroencephalography and clinical Neurophysiology, 72, 225-231.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G., & Wittkowski, W. (1988). Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields. Electroencephalography and Clinical Neurphysiology, 69, 160-170.

Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Fahrendorf, G., & Stöber, U. (1990). Identification of sources of brain neuronal activity with high spatiotemporal resolution through combination of neuromagnetic source localization (NMSL) and magnetic resonance imaging (MRI). Electroencephalography and clinical Neurophysiology, 75, 173-184.

Pantev, C., Hoke, M., Lütkenhöner, B., & Lehnertz, K. (1989). Tonotopic Organization of the auditory cortex: pitch versus frequency representation. Science, 246, 246-248.

Parving, A., Salomon, G., Elberling, C., Larson, B., & Lassen, N. (1980). Middle components of the auditory evoked response in bilateral temporal lobe lesions. Scand. Audiol., 9, 161-167.

Paulesu, E., Frith, C. D., & Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. Nature, 362, 342-345.

Peretz, I., Kolinsky, R., Tramo, M., Labrecque, R., Hublet, C., Demeurisse, G., & Belleville, S. (1994). Functional dissociations following bilateral lesions of auditory cortex. Brain, 117, 1283-1301.

Petersen, S., Fox, P., Posner, M., Mintun, M., & Raichle, M. (1988). Positron emission tomographic studies of the cortical anatomy of single-word processing. Nature, 331, 585-589.

Petersen, S., Fox, P., Posner, M., Mintun, M., & Raichle, M. (1989). Positron emission tomographic studies of the processing of single words. Journal of Cognitive Neuroscience, 1(2), 153-170.

Phillips, D., & Farmer, M. (1990). Acquired word deafness, and the temporal grain of sound representation in the primary auditory cortex. Behavioural Brain Research, 40, 85-94.

Pick, A. (1892). Beiträge zur Lehre von den Störungen der Sprache. Archiv für Psychiatrie und Nervenkrankheiten, 23, 896-918.

Pinker, S. (1994). The Language Instinct. New York: Morrow.

Plante, E. (1991b). MRI findings in the parents and siblings of specifically language-impaired boys. Brain & Language, 41, 67-80.

Plante, E., Swisher, L., Vance, R., & Rapcsak, S. (1991a). MRI findings in boys with specific language impairment. Brain & Language, 41, 52-66.

Poeppel, D. (in press). A critical review of PET studies of phonological processing. Brain & Language.

Poeppel, D., & Johnson, S. (in press). Neuroimaging studies of language should connect with (psycho)linguistic theories. Behavioral and Brain Sciences.

Popper, A., & Fay, R. (Ed.). (1992). The Mammalian Auditory Pathway: Neurophysiology. New York: Springer.

Posner, M., & Raichle, M. (1994). Images of Mind. New York: Freeman.

Praamstra, P., Hagoort, P., Maassen, B., & Crul, T. (1991). Word deafness and auditory cortical function: a case history and hypothesis. Brain, 114, 1197-1225.

Price, C., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K., & Frackowiak, R. (1992). Regional response differences within the human auditory cortex when listening to words. Neuroscience Letters, 146, 179-182.

Rapin, I. (1985). Cortical deafness, auditory agnosia, and word deafness: how distinct are they? Hum. Commun. Can., 9, 29-37.

Reinhold, M. (1950). A case of auditory agnosia. Brain, 73, 203-223.

Remez, R. (1994). A guide to research on the perception of speech. In M. Gernsbacher (Eds.), Handbook of Psycholinguistics New York: Academic Press.

Richter, H. (1957). Akustischer Funktionswandel bei Wortaubheit. Archive für Psychiatrie und Nervenkrankheiten, 196, 99.

Rif, J., Hari, R., Hämäläinen, M., & Sams, M. (1991). Auditory attention affects two different areas in the human supratemporal cortex. Electroencephalography and Clinical Neurophysiology, 79, 464-472.

Roberts, M., Sandercock, P., & Ghadiali, E. (1987). Pure word deafness and unilateral right temporo-parietal lesions: a case report. J Neurol Neurosurg Psychiatry, 50(12), 1708-1709.

Roland, P. (1993). Brain Activation. New York: Wiley-Liss.

Roland, P., Larsen, B., Lassen, N., & Skinhøj, E. (1980). Supplementary motor area and other cortical areas in the organization of voluntary movements in man. J. Neurophysiology, 43, 118-136.

Roland, P., Skinhøj, E., & Lassen, N. (1981). Focal activations of human cerebral cortex during auditory discrimination. Journal of Neurophysiology, 54(6), 1139-1151.

Romani, G., Williamson, S., & Kaufman, L. (1982). Tonotopic organization of the human auditory cortex. Science, 216, 1330-1340.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. Philosophical Transaction of the Royal Society of London, Biology, 336, 367-373.

Rugg, M. (1983). The relationship between evoked potentials and lateral asymmetries of processing. In A. Gaillard & W. Ritter (Eds.), Tutorials in ERP Research: Endogenous Components (pp. 369-431). Amsterdam: North-Holland.

Rumelhart, D., & McClelland, J. (Ed.). (1986). Parallel distributed processing. Cambridge, MA: MIT Press.

Saffran, E., Marin, O., & Yeni-Komshian, G. (1976). An analysis of speech perception in word deafness. Brain and Language, 3, 209-228.

Sams, M., Reijo, A., Hämäläinen, M., Hari, R., Lounasmaa, O., Lu, S., & Simola, J. (1991). Seeing speech: visual information from lip movements modifies activity in the human auditory cortex. Neuroscience Letters, 127, 141-145.

Sato, M., Yasui, N., & Isobe, I. (1982). A case of pure word deafness and auditory agnosia associated with bilateral temporo-parietal lesions. Brain and Nerve (Tokyo), 34, 939-945.

Scherg, M., Vajsar, J., & Picton, T. (1989). A source analysis of the late human auditory evoked potentials. Journal of Cognitive Neuroscience, 1(4), 336-355.

Schlaug, G., Jäncke, L., Huang, Y., & Steinmetz, H. (1995). In vivo evidence of structural brain asymmetry in musicians. Science, 267, 699-701.

Schuster, P., & Taterka, H. (1926). Beitrag zur Anatomie der reinen Worttaubheit. Zeitschrift für die gesamte Neurologie und Psychiatrie, 105, 494-538.

Shaywitz, B., Shaywitz, S., Pugh, K., Constable, R., Skudlarski, P., Fulbright, R., Bronen, R., Fletcher, J., Shankweiler, D., Katz, L., & Gore, J. (1995). Sex differences in the functional organization of the brain for language. Nature, 373, 607-609.

Shindo, M., Kaga, K., & Tanaka, Y. (1991). Speech discrimination and lip reading in patients with word deafness or auditory agnosia. Brain and Language, 40, 153-161.

Shoumaker, R., Ajax, E., & Schenkenberg, T. (1977). Pure word deafness (auditory verbal agnosia). Dis. Nerv. Syst., 38, 213-299.

Steinschneider, M., Schroeder, C., Arezzo, J., & Vaughan, H. (1994). Speech-evoked activity in primary auditory cortex: effects of voice onset time. Electroencephalogr Clin Neurophysiol, 92(1), 30-43.

Steinschneider, M., Schroeder, C., Arezzo, J., & Vaughn, H. (1995). Physiologic correlates of the voice onset time boundary in primary auditory cortex (A1) of the awake monkey: temporal response patterns. *Brain and Language*, *48*, 326-340.

Steinschneider, M., Schroeder, C. E., Arezzo, J. C., & Vaughan, H. G. J. (1993). Temporal Encoding of Phonetic Features in Auditory Cortex. *Ann N Y Acad Sci* (682), 415-417.

Stockert, T. (1982). On the structure of word deafness and mechanisms underlying the fluctuation of disturbances of higher cortical functions. *Brain and Language*, *16*, 133-146.

Stromswold, K., Alpert, N., Rausch, S., & Caplan, D. (in press). Localization of sentence processing by positron emission tomography. *Brain & Language*.

Studdert-Kennedy, M., & Shankweiler, D. (1970). Hemispheric specialization for speech perception. *Journal of the Acoustic Society of America*, *48*(2), 579-594.

Swinney, K., & Wikswo, J. (1980). A calculation of the magnetic field of a nerve action potential. *Biophys. J.*, *32*, 719-732.

Tailarach, J., & Tournoux, P. (1988). *Co-planar stereotactic atlas of the human brain* (2nd ed.). Stuttgart: Thieme Verlag.

Takahashi, N., Kawamura, M., Shinotou, H., Hirayama, K., Kaga, K., & Shindo, M. (1992). Pure word deafness due to left hemisphere damage. *Cortex*, *28*, 295-303.

Tanaka, Y., Taguchi, K., Sakabe, N., & Igarashi, E. (1964). Pure auditory agnosia associated with alteration of voice and personal character. *Folia Phoniat*, *17*, 185-194.

Tanaka, Y., Yamadori, A., & Mori, E. (1987). Pure word deafness following bilateral lesions: a psychophysical analysis. *Brain*, *110*, 381-403.

Teuber, H. (1978). The brain and human behavior. In R. Held, H. Leibowitz, & H. Teuber (Eds.), *Handbook of Sensory Physiology* New York: Springer-Verlag.

Tinbergen, N. (1951). *The study of instinct*. Oxford: Oxford University Press.

Tramo, M., Bharucha, J., & Musiek, F. (1990). Music perception and cognition following bilateral lesions of auditory cortex. *Journal of Cognitive Neuroscience*, *2*(3), 195-212.

Ulrich, G. (1978). Interhemispheric functional relationships in auditory agnosia: an analysis of the preconditions and a conceptual model. *Brain and Language*, *5*, 286-300.

Vignolo, L. (1982). Auditory agnosia. *Phil. Transactions R. Soc. London, B*, *298*, 49-57.

Werker, J., & Tees, R. (1992). The organization and reorganization of speech perception. *Annual Review of Neuropsychology*, *15*, 377-402.

Wernicke, C. (1874). *Der aphasische Symptomenkomplex*. Breslau: Cohn and Weigart.

Wernicke, C., & Friedlaender, C. (1883). Ein Fall von Taubheit in Folge von doppelseitiger Läsion des Schläfenlappens. *Fortschritte der Medizin*, *1*, 177-185.

Wise, R., Chollet, F., Hadar, U., Friston, K., Hoffner, E., & Frackowiak, R. (1991). Distribution of cortical neural networks involved in word comprehension and word retrieval. Brain, 114, 1803-1817.

Wohlfahrt, G., Lindgren, A., & Jernelins, B. (1952). Clinical picture and morbid anatomy in a case of 'pure word deafness'. Journal of Nervous and Mental Disease, 116, 818-827.

Wolberg, S., Temlett, J., & Fritz, V. (1990). Pure word deafness. S Afr Med J, 78, 668-670.

Woldorff, M., Gallen, C., Hampson, S., Hillyard, S., Pantev, C., Sobel, D., & Bloom, F. (1993). Modulation of early sensory processing in human auditory cortex during auditory selective attention. Proceedings of the National Academy of Sciences of the USA, 90, 8722-8726.

Yamamoto, T., Williamson, S., Kaufman, L., Nicholson, C., & Llinas, R. (1988). Magnetic localization of neuronal activity in the human brain. The Proceedings of the National Academy of Sciences of the USA, 85, 8732-8736.

Yaqub, B., Gascon, G., Al-Nosha, M., & Whitaker, H. (1988). Pure word deafness (acquired verbal auditory agnosia) in an Arabic speaking patient. Brain, 111, 457-466.

Zatorre, R., Evans, A., Meyer, E., & Gjedde, A. (1992). Lateralization of phonetic and pitch discrimination in speech processing. Science, 256, 846-849.

Ziegler, D. (1952). Word deafness and Wernicke's aphasia: report of cases and discussion of the syndrome. Archives of Neurology and Psychiatry, Chicago, 67, 323-331.

Ziehl, F. (1896). Ueber einen Fall von Wortaubheit und das Lichtheimsche Krankheitsbild der subcorticalen sensorischen Aphasie. Deutsche Zeitschrift für Nervenheilkunde, 8, 259.

Zurif, E. (in press). Brain regions of relevance to syntactic processing. In L. Gleitman (Eds.), An Invitation to Cognitive Science Cambridge, MA: MIT Press.