

Research report

# Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds

D. Poeppel<sup>a,\*</sup>, E. Yellin<sup>b</sup>, C. Phillips<sup>c</sup>, T.P.L. Roberts<sup>a</sup>, H.A. Rowley<sup>a</sup>, K. Wexler<sup>c,d</sup>,  
A. Marantz<sup>c</sup>

<sup>a</sup> *Biomagnetic Imaging Laboratory, Department of Radiology, University of California San Francisco, 513 Parnassus Avenue, S-362, San Francisco, CA 94143-0628, USA*

<sup>b</sup> *Department of Electrical Engineering and Computer Science, MIT, Boston, MA, USA*

<sup>c</sup> *Department of Linguistics and Philosophy, MIT, Boston, MA, USA*

<sup>d</sup> *Department of Brain and Cognitive Sciences, MIT, Boston, MA, USA*

Accepted 28 May 1996

## Abstract

The auditory evoked neuromagnetic fields elicited by synthesized speech sounds (consonant–vowel syllables) were recorded in six subjects over the left and right temporal cortices using a 37-channel SQUID-based magnetometer. The latencies and amplitudes of the peaks of the M100 evoked responses were bilaterally symmetric for passively presented stimuli. In contrast, when subjects were asked to discriminate among the same syllabic stimuli, the amplitude of the M100 increased in the left and decreased in the right temporal cortices. Single equivalent current dipole modeling of the activity elicited by all stimulus-types localized to a well-circumscribed area in supratemporal auditory cortex. The results suggest that attentional modulation affects the two supratemporal cortices in a differential manner. Task-conditioned attention to speech sounds is reflected in lateralized supratemporal cortical responses possibly concordant with hemispheric language dominance.

*Keywords:* Auditory; Speech; Lateralization; Attention; Cognitive neuroscience; Functional neuroimaging

## 1. Introduction

A central challenge for cognitive neuroscience is to understand the physiological basis of speech and language processing and relate it to the anatomical asymmetry that is characteristic of cerebral language representation. One approach is to focus on the neural basis of the acoustic and speech-perceptual input processes and relate these to hemispheric asymmetry and language dominance. An important advantage gained by concentrating on speech perception as a partially independent subroutine of language comprehension is that one can use simple stimuli the physical makeup of which is well understood. The benefit of combining data from the brain imaging techniques magnetoencephalography (MEG) and magnetic resonance imaging (MRI) is that

one obtains both high temporal resolution and high spatial resolution information.

A number of investigators have recorded auditory evoked neuromagnetic fields elicited by speech sounds. The speech stimuli have included vowels, syllables, and words [11,16–19,35]. Other magnetoencephalographic studies have assessed pure tone processing to elucidate the functional organization of the auditory cortex [1,4,10,14,15,20,24,26–28,34,39]. The latter studies have been able to characterize in detail some components of the auditory evoked response and localize these components to supratemporal auditory cortex. Several distinct auditory evoked response components have been described. Following a suprathreshold auditory stimulus, one typically detects event-related fields (ERFs) 40–50 ms after stimulus onset, a prominent response around 100 ms (the M100, or N1m), and a response of opposite polarity to the M100 at 150–200 ms. In addition, some studies have documented a sustained field (SF) response that remains visible through-

\* Corresponding author. Fax: (1) (415) 502 4302; e-mail: poeppel@itsa.ucsf.edu

out the duration of a stimulus [11,26] and a response to stimulus offset comparable to the M100 [15].

There is consensus that for auditory stimuli the M100 evoked response is particularly important [23]. The M100 is practically always elicited by auditory stimuli and its sensitivity to acoustic stimulus features is demonstrated in that the amplitude and latency vary systematically with certain physical and temporal aspects of stimuli [13,23]. 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 cochleotopically organized fields [4,10,27,34,39]. Attempts to correlate response components with specific acoustic or linguistic–phonetic features of stimuli have generated more complex results [18,20].

To investigate the issue of the relative symmetry or asymmetry of speech sound processing, we recorded neuromagnetic data from the left and right superior temporal lobes independently while the subjects listened to syllable stimuli. Because of the importance of the M100 for the assessment of auditory stimuli, we restricted the present analysis to the M100 latency and amplitude and used a single equivalent current dipole model of the neuromagnetic field.

Several MEG studies have explored the evoked neuromagnetic responses to consonant–vowel syllables [16,18,19]. Only Kuriki and Murase [19] recorded from both hemispheres and compared the results with respect to the lateralization of responses. They reported that, 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 did not discuss whether their localizations placed the dipoles in primary or association auditory cortex. Whether there

are differential hemispheric responses to syllables and their intrinsic properties remains an open question.

Independent observations coming from the hemodynamic functional neuroimaging literature motivated the exploration of the interaction of speech stimuli with attentional requirements, such as those processes engaged by having to execute particular linguistic tasks. There is a large literature devoted to the study of attention both with event-related potentials and, more recently, with MEG. Two types of experimental protocols are commonly employed, ‘alternating ear’ and dichotic listening paradigms, in which one is asked to selectively attend to the input delivered to the left or right ears [1,33,38]. 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 and colleagues [38] studied the dichotic presentation of pure tones and reported that the same tone elicited larger M100 amplitudes when attended than unattended.

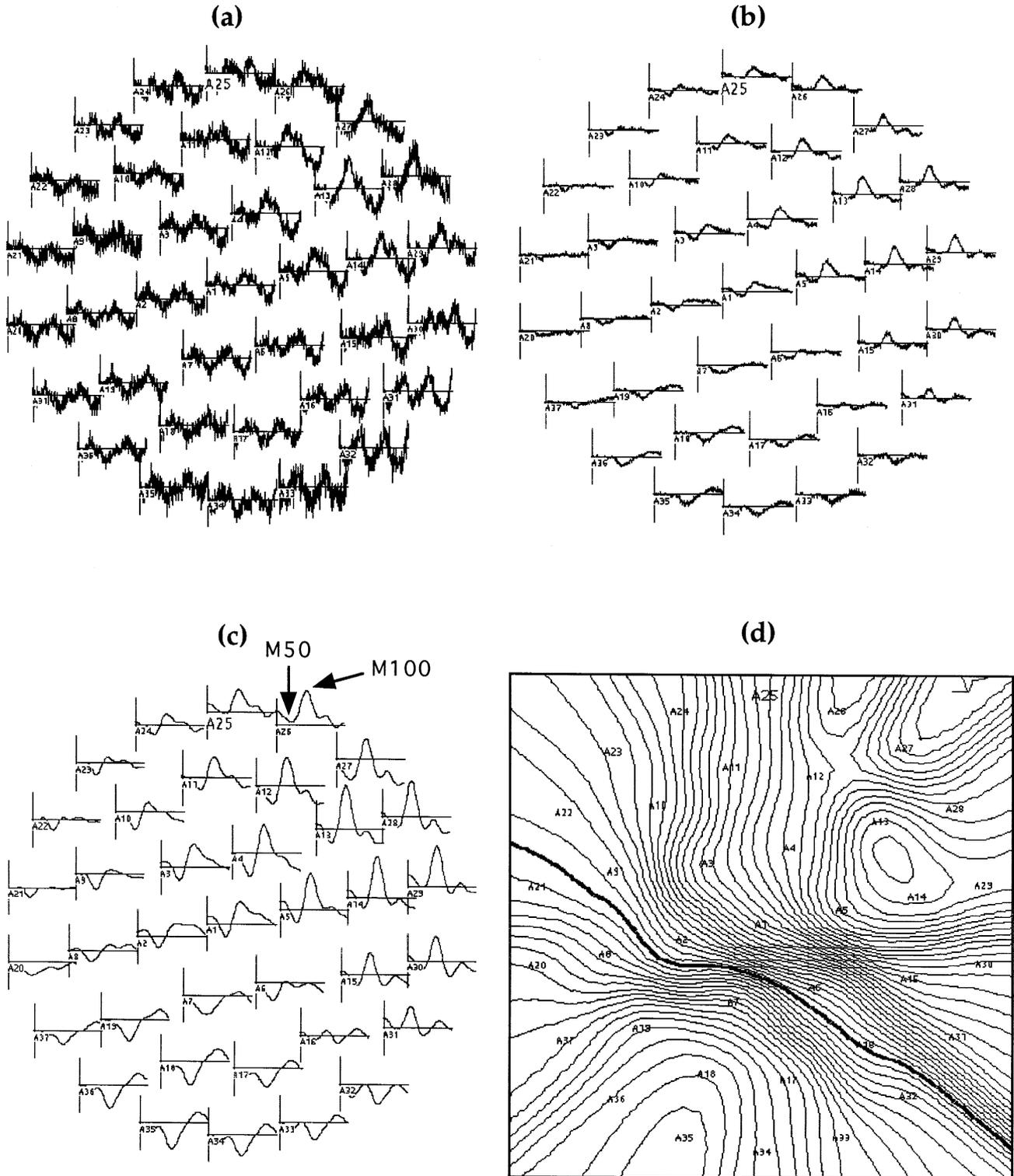
This study addresses an even simpler question. Subjects hear the same stimulus set in two conditions. In one condition they passively listen to syllables. In a second, attended condition, they listen to the same stimulus set and are required to make an overt discrimination on the stimuli. It is hypothesized that task-conditioned selective attention will modulate the response properties. Specifically, the response modulation is anticipated to be particularly evident in amplitude increases in the left temporal responses.

There is a discrepancy between much previous research that highlights *asymmetries* in speech and language processing and contemporary findings coming from neuroimaging modalities such as positron emission tomography (PET) and functional magnetic resonance imaging (fMRI), where *symmetrical* processing is often observed [2,29,31]. Briefly, the passive presentation of speech stimuli (whether as syllables, pseudowords, words, or connected speech) is very often associated with the bilateral

Fig. 1. Sensor layout displays and contour map. The auditory evoked magnetic field from one subject, elicited by the presentation of a vowel, is illustrated here 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. The stimulus was the vowel /a/, with a fundamental frequency of 100 Hz (male pitch) and a duration of 300 ms. The horizontal scale is approximately 240 ms/cm. The vertical display scale was set to 1000 fT/division. 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. (c) Sensor layout display of the data shown in (b), *bandpass filtered* from 1–20 Hz. The evoked response to 100 presentations of /a/ was averaged (b), and then digitally filtered (20 dB attenuation at 25 Hz). The vertical scale is 200 fT/division. The sensors on the upper right and lower left illustrate the prominent evoked response components. Large field deflections (of opposite polarities) at about 50 ms, 100 ms, and 200 ms following stimulus can be seen. These deflections represent the M50 (P50), M100 (N1m), and M200 (P2) auditory evoked response components. Although the evoked response components are visible in a single epoch, as seen in (a), signal averaging and filtering optimizes the signal-to-noise ratio. (d) Isofield *contour map* for the data shown in (c) is plotted at a latency of 108 ms (M100 peak). The contour interval is 10 fT and the scale is 500 fT. 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 vs. out) as visualized by a contour map.

activation of the superior temporal cortices. In contrast, task-modulated speech perception is associated with asymmetric activation results, often concordant with language dominance [30]. By explicitly separating *speech percep-*

*tual* processes from input-independent *language* processing, this work aims to resolve aspects of the existing disagreements. This study focuses on the contribution of left and right auditory cortices to the processing of speech



stimuli and the interaction with attentional mechanisms as reflected by properties of the M100 auditory evoked neuromagnetic response.

## 2. Materials and methods

### 2.1. Subjects

Six normal volunteers (three women, mean age = 29.6 years) participated in the experiments. The participants gave informed consent, separately for the MEG and the MRI components of the study. None of the subjects had any hearing loss or neurological abnormalities. Subjects S5 and S6 are strong left-handers, with no left-handed first-degree relatives. Handedness interacts in particular ways with hemispheric asymmetry as assessed by different measures for language and other cognitive domains, and the data for the two left-handed subjects are explicitly marked in the relevant graphs, and analysis was done both including and excluding these subjects.

### 2.2. Stimuli

In both the passive listening and the discrimination conditions, which were counterbalanced, subjects heard the syllables /bæ/, /pæ/, /dæ/, and /tæ/. For each subject, the stimulus amplitude was adjusted to a comfortable level (50–70 dB SPL). The 300 ms duration stimuli were constructed using a Klatt formant synthesizer running on a Macintosh (SenSyn™, Sensimetrics, Cambridge MA). The two voiced CVs (/bæ/ and /dæ/) were synthesized with a VOT of 20 ms and the two voiceless CVs were synthesized with a VOT of 80 ms. These values were chosen to approximate relatively typical tokens of each syllable. Each stimulus was presented 100 times, with the four syllables appearing in pseudorandom order at a variable ISI (1000–1500 ms).

### 2.3. 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, San Diego, CA). The sensor-array was placed over the left or right temporal lobe and its position was adjusted to record effectively from auditory cortex by evaluating whether a suprathreshold pure tone (1 kHz) elicited a strong and reliably localizable response. The position of a participant's head with respect to the sensors was recorded by a transceiver-based system which localized fiducial landmarks, thereby also setting up a spatial reference frame for the MEG data.

Epochs of 600 ms that included a 100 ms pre-stimulus interval and a 200 ms post-stimulus interval were acquired. Neuromagnetic data were recorded using a 1.0 Hz high-pass

cutoff. The sampling rate was 1041.7 Hz with a bandwidth of 400 Hz. Fig. 1a–d illustrates typical data and the standard sequence of data acquisition and processing. The steps include the simultaneous acquisition of the time-varying evoked neuromagnetic field in 37 channels for each epoch (Fig. 1a), signal averaging (Fig. 1b), filtering of the averaged evoked field data (Fig. 1c), and reconstruction of contour maps of the evoked field at specific time-points (Fig. 1d). These data may be used to compute the root mean square (RMS) of responses across channels and best-fitting single equivalent current dipoles at selected points.

### 2.4. Procedure

During the recordings, the subjects lay on their left or right side, depending on the hemisphere of recording. The stimuli were presented to the contralateral ear. Contralateral auditory presentation typically yields larger response amplitudes [21,22,24]. A small but tightly fitting earphone was inserted into the stimulated ear and the other ear was plugged to eliminate possible distracting noise from ambient sound.

In the discrimination condition, subjects had to execute a 2-alternative forced-choice decision, grouping stimuli by VOT (/bæ/ and /dæ/ vs. /pæ/ and /tæ/) in a voiced/voiceless classification. Subjects pressed one response button when perceiving the short VOT /bæ/ or /dæ/ and another when perceiving the long VOT /pæ/ or /tæ/. Subjects pressed the response button with the hand ipsilateral to the recording side to minimize the interference from the motor component of the task. Stimuli were presented and subjects' responses and response latencies were collected using the PsyScope software [6].

### 2.5. Magnetic resonance images

For each subject, high resolution volumetric magnetic resonance images (SPGR sequence,  $128 \times 128 \times 124$  matrix, resolution  $\sim 1 \times 1 \times 1.5$  mm, TR = 36 ms, TE = 8 ms, flip =  $70^\circ$ ) were acquired using a 1.5 Tesla SIGNA magnetic resonance scanner (GE Medical Systems, Milwaukee, WI). The MR images were judged to be normal by a neuroradiologist (H.A.R.).

### 2.6. Data analysis

For each stimulus, the 100 epochs (or more) were averaged for further processing. The averaged waveforms were digitally filtered with a simulated 4th-order Butterworth bandpass filter with a low cutoff frequency of 1.0 Hz and a high cutoff frequency of 20 Hz, and all further analyses were performed on the filtered data. The root mean square (RMS) of the field strength across the 37 channels was calculated for each sampled point in time.

To estimate the source parameters of the M100, a single equivalent current dipole model was used. The localization

algorithm inversely (iterative least-squares minimization) computes a single dipole in a spherical volume of uniform conductivity. Representative *single* dipoles were chosen in the following way: the latency of the M100 peak was determined, and then all the dipoles in a 10 ms window around the M100 peak were computed (i.e. peak latency  $\pm$  5 ms). Of all the dipoles in that interval, the dipole with the smallest confidence volume was chosen as representative for that condition.

The left and right preauricular points and nasion were marked during the MEG recordings and served as the basis for a common coordinate system for the MEG and MRI data sets. 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).

The M100 RMS peak and latency data were analyzed using standard analysis of variance (ANOVA) methods using the UNIXSTAT package.

### 3. Results

#### 3.1. Behavioral results

Of the total of 4800 trials acquired in the discrimination condition (6 subjects  $\times$  2 hemispheres  $\times$  4 syllables  $\times$  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. Five of the 6 subjects performed with an accuracy rate ranging from 98.25 to 100%, corresponding to an absolute number of errors ranging from 0 to 14. One subject (S6) performed worse, making a total of 47/799 categorization errors. The analysis of variance of the reaction time data for the behavioral discrimination condition showed no significant main effect of ear,  $F_{1,5} = 3.707$ ,  $P = 0.112$ , or syllable (CV),  $F_{3,15} = 1.012$ ,  $P = 0.415$ , and no significant hemisphere  $\times$  syllable interaction,  $F_{3,15} = 2.233$ ,  $P = 0.126$ .

#### 3.2. Evoked neuromagnetic field results

##### 3.2.1. Passive listening

Table 1 lists the M100 RMS peaks and the latencies of the M100 RMS peaks for each subject and each syllable. For the passive condition, analyses of variance of the latency data reveal no significant or marginally significant main effects or interactions. Importantly, there was no effect of hemisphere.

Although the amplitude data revealed a complex patterns of results, the dominant finding is that the M100 largely reveals bilateral symmetry; however, the interaction hemisphere  $\times$  voicing was significant,  $F_{1,5} = 9.039$ ,

Table 1  
M100 RMS peak and latency for passive syllable and discriminate syllable conditions

Subject	Stim	a: Stops – passive				b: Stops – discriminate			
		M100 RMS peak value (fT)		Lat. of M100 RMS peak (ms)		M100 RMS peak value (fT)		Lat. of M100 RMS peak (ms)	
		Left	Right	Left	Right	Left	Right	Left	Right
S1	bæ	96	100	96	101	105	89	122	108
	pæ	94	133	101	110	91	94	100	101
	dæ	84	83	111	116	112	64	119	107
	tæ	74	98	108	102	114	64	120	112
S2	bæ	90	103	104	112	105	97	107	102
	pæ	80	130	103	117	116	101	113	107
	dæ	78	105	105	108	113	102	103	101
	tæ	88	105	101	106	134	96	112	96
S3	bæ	145	125	102	85	146	121	105	89
	pæ	128	151	98	88	131	126	100	92
	dæ	136	128	101	82	157	110	104	92
	tæ	126	138	97	88	121	122	104	89
S4	bæ	178	160	101	91	231	140	100	85
	pæ	133	135	104	90	152	125	102	83
	dæ	201	173	102	93	255	165	100	88
	tæ	150	119	98	93	185	130	96	82
S5	bæ	164	132	93	83	184	125	95	83
	pæ	175	152	97	89	176	120	98	88
	dæ	189	145	95	84	214	116	98	85
	tæ	187	149	95	87	136	121	98	86
S6	bæ	143	121	96	96	149	184	98	118
	pæ	107	65	96	95	98	102	100	122
	dæ	167	106	96	94	145	150	98	129
	tæ	80	101	109	88	81	63	94	113

$P = 0.030$ , with the left hemisphere response being larger with voiced stimuli.

### 3.2.2. Discrimination task

Table 1 (b) lists the M100 RMS peaks and the latencies for each subject for the syllable discrimination condition. As in the passive condition, the analyses of variance of the latency data revealed no significant effects, particularly no effects of hemisphere.

To focus on the effect of having an attentionally explicit task it is necessary to better characterize the effect the task. To quantify the effect, Fig. 2 shows the difference, for each subject and each syllable, between the RMS peak in the syllable discrimination and the RMS peak in the passive condition. The data reveal a significant difference with respect to how the left and right temporal cortex responses are modulated by the presence of the discrimination task.

Consider first the left hemisphere data from subjects S1–S4 (right-handed subjects). In 14 of 16 measurements, the task increased the mean RMS value in the left, with the increase ranging from 1.4 fT to 54.8 fT. In the right hemisphere measurements from the same subjects, a decrease was observed in 15 of 16 measurements, with values ranging from 3.1 fT to 39.5 fT. Subjects S5 and S6 (left-handers) patterned in two directions. The response pattern for S5 was comparable to the patterns for the right-handed subjects, with the exception of one single value (/tæ/) which significantly altered the overall data

for this subject. The comparison of responses for S6 are in the opposite direction: the values in the left hemisphere, on average, decreased, whereas the right hemisphere values showed large increases. Important is that the changes, with few exceptions, were extremely large. Typical RMS values of evoked M100 responses lie between 100 and 200 fT. The task-dependent modulations were thus on the order of tens of percent of change.

In the initial analysis of variance for all subjects, there was no main effect of task,  $F_{1,5} = 0.349$ ,  $P = 0.580$ , and no hemisphere  $\times$  task interaction,  $F_{1,5} = 3.933$ ,  $P = 0.104$ . However, the data were analyzed in two additional more meaningful ways: once by recoding by ‘presumed dominant hemisphere’ (which meant recoding the values for S6) and once by omitting from the analyses the two left-handed subjects (S5, S6). Particularly noteworthy about these re-analyses was the hemisphere  $\times$  task interaction. (When the data were analyzed as a group of  $n = 6$ , that effect was not significant.) When the two left-handers were omitted, that interaction was highly significant,  $F_{1,3} = 39.957$ ,  $P = 0.008$ , and when the one left-handed subject was recoded (S6), the effect was also extremely large,  $F_{1,5} = 55.307$ ,  $P = 0.001$ . The latency data showed a significant main effect of hemisphere after the recoding,  $F_{1,5} = 12.214$ ,  $P = 0.017$ . The latencies were longer in the left temporal recordings. That effect was no longer significant when the two left-handed participants were omitted from the analysis. When the data were recoded by presumed dominant hemisphere, there was no main effect of hemisphere,  $F_{1,5}$

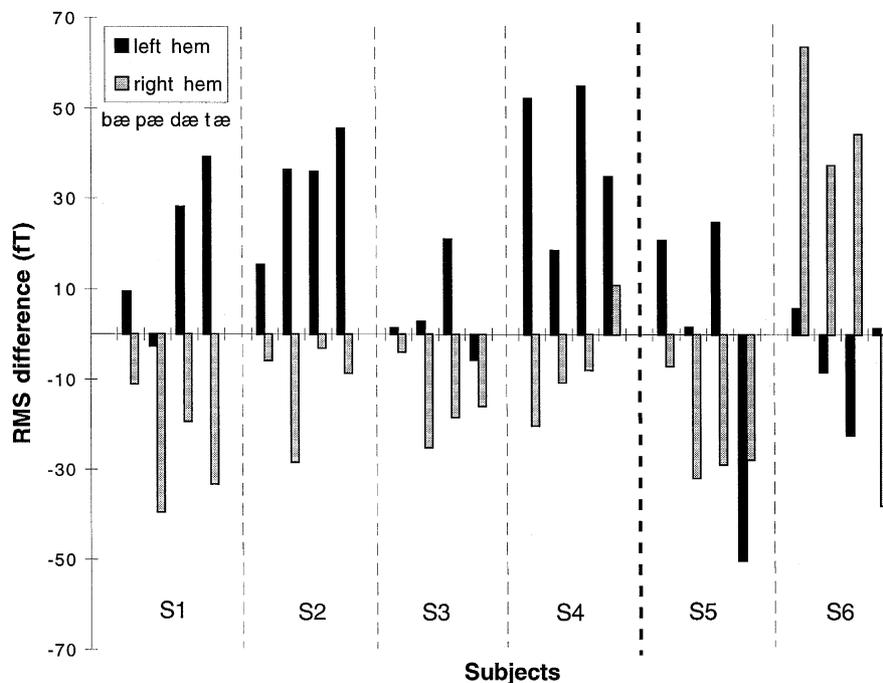


Fig. 2. M100 RMS difference values (task – passive) for each subject and each syllable. Each bar represents the M100 amplitude difference for each syllable and each hemisphere. Positive values (above abscissa) indicate that the values in the task condition were larger than the corresponding measurements in the passive condition. Negative values 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, grey bars represent the comparison for the right hemisphere.

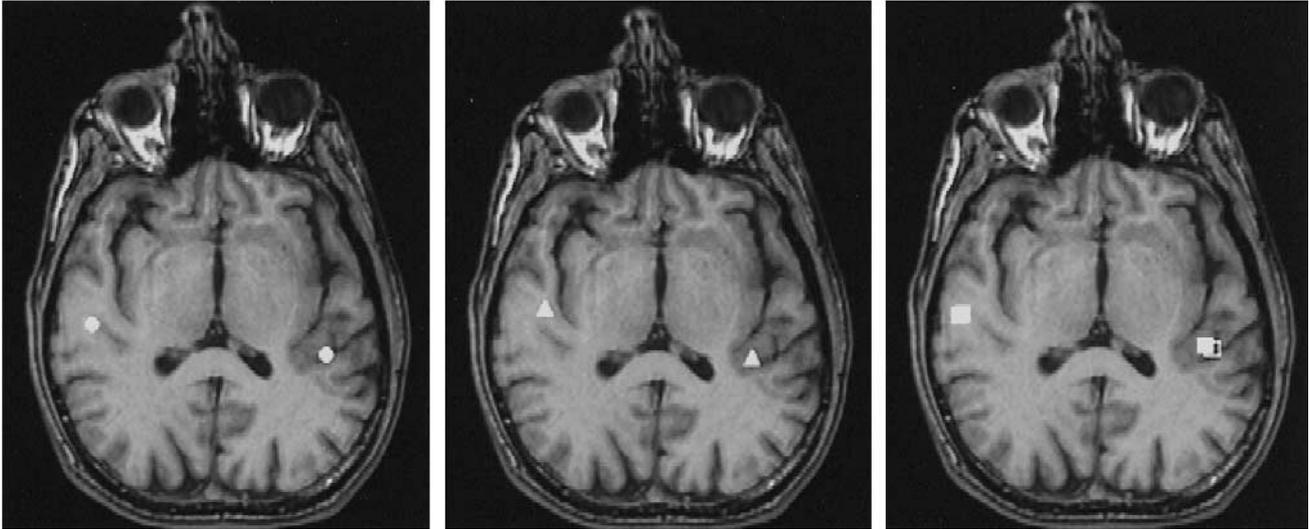
**Tone (500Hz)****Vowel (/a/)****Syllables**

Fig. 3. Single dipoles at the M100 RMS peak coregistered with MRI. Each plotted point represents the single dipole chosen for a given stimulus condition for the subject S3. The right side of each MR image corresponds to the left hemisphere. Single dipoles were derived using a iterative least-squares minimization algorithm. All the dipoles appear to be on the supratemporal plane. (a) Single equivalent current dipole for 500 Hz pure tone, monaural passive presentation. (b) Single dipole for passively presented vowel /a/ with a 100 Hz fundamental frequency. (c) Dipoles for syllable stimuli. Filled squares correspond to passive presentation, open squares to the task condition. The dipoles colocalize to an extent that they are individually barely visible in the overlay.

= 0.089,  $P = 0.777$  in the passive listening condition. In contrast, for the discrimination condition, there was a main effect of hemisphere,  $F_{1,5} = 11.282$ ,  $P = 0.020$ . Similarly, when the two left-handed subjects were omitted, there was no main effect of hemisphere in the passive listening

condition,  $F_{1,3} = 0.460$ ,  $P = 0.546$ , but a main effect of hemisphere approaching significance in the task-modulated condition,  $F_{1,3} = 8.432$ ,  $P = 0.062$  despite the low power ( $n = 4$ ). In summary, the presence of the task appeared to ‘induce’ the observed hemispheric asymmetry.

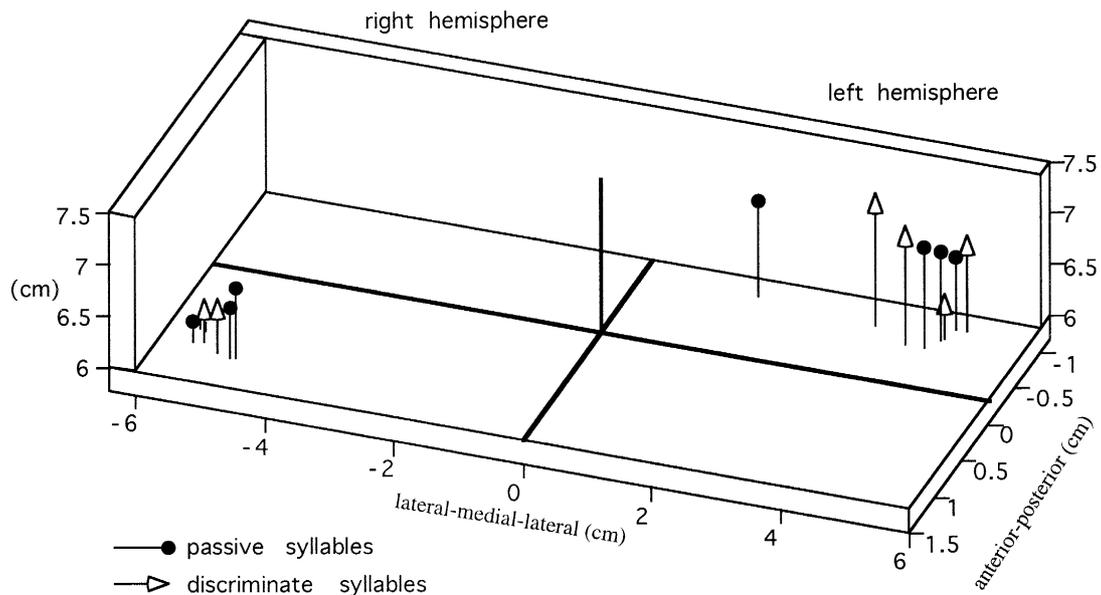


Fig. 4. Single dipole locations for syllables (passive and discriminate) for one subject. Filled circles represent the dipoles derived for the passive syllable condition, open triangles represent the dipoles for the syllable discrimination condition. As ascertained by MR (cf. Fig. 3), all dipoles localize to supratemporal auditory cortex.

### 3.3. Localization results

For all subjects, the best single equivalent current dipole was computed for each of the four syllables in both conditions. The dipole source localization results are presented in Figs. 3 and 4. Fig. 3 illustrates the coregistration of MEG-derived dipoles with the same subject's MRI. Fig. 4 displays the syllable-elicited dipoles for one representative subject. The dipoles generated in both conditions clustered in very tight regions in the left and right auditory cortices (in the same area of supratemporal auditory association cortex that was activated by vowels and tones in independent recordings).

For subject S3, the dipole correlations ranged from 0.96 to 0.99, high correlations that correspond well to those obtained with pure tone stimuli. Fig. 4 shows that the dipoles for the two conditions were extremely closely spaced; there was a single outlier, and the remaining 15 dipoles (2 hemispheres  $\times$  2 conditions  $\times$  4 syllable types = 16 total dipoles) were tightly spaced in all three dimensions. (Some dipoles in the right were on top of one another and are consequently not seen in the display.) The left and right regions in which the syllable-elicited dipoles were clustered were the same regions of auditory cortex in which the subject showed the vowel and tone dipole clusters (cf. Fig. 3). Similar correlation and localization results were observed for the other subjects. Particularly consistent were 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 S3 the cluster of dipoles was more anterior in the right, this was true for other auditory (vowel and tone) dipole localizations as well. This more anterior positioning in the right supratemporal plane is reported in many auditory MEG studies and is commonly attributed to the large morphological asymmetries between the left and right supratemporal planes [12].

## 4. Discussion

This study focused on whether the execution of a speech sound categorization task modulates the responses measured from auditory cortex in a selective way. There are two aspects to the issue. The first is whether the neuromagnetic M100 response is in principle influenced by attentional demands or whether the M100 reflects processing that is independent of any task requirements. A number of investigators using both EEG and MEG technologies have been able to show that late auditory evoked response components can be modulated by certain task conditions (for review, see [21]). For example, Woldorff et al. [38] presented subjects dichotically with tones of different frequencies and showed that the M100 generated by a

stimulus in the attended ear was larger than the M100 generated by the same stimulus when presented to the unattended ear. This study demonstrates that an even simpler manipulation, the mere absence or presence of a simple task, can modulate the amplitude of the M100 evoked field.

The second issue investigated is motivated by data on neuroimaging of speech sounds with PET and fMRI and from previous selective attention studies using MEG (e.g. [1]). Many observations in the hemodynamic imaging literature are consistent with the model 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 linguistic tasks. Several studies suggest that when subjects execute '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 [8,9,40]. The experiment presented here tested the hypothesis that there is an attention-dependent gating of activity that differentially affects the two hemispheres [30]. The same stimulus material was presented in two different experimental conditions: passive listening or performing a discrimination task. 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.

The main finding of the study was that a simple attentional modulation affected the two supratemporal cortices differently. Passive listening to syllables showed complex but – with respect to most analyses – bilaterally symmetric responses in left and right superior temporal cortices. In contrast, the execution of a simple categorization task on the same stimuli revealed that a region in left supratemporal cortex increased its response amplitude (large M100 amplitude increases) while the (presumably homologous) region in right temporal cortex revealed significantly decreased responses. Attention to speech sounds was thus reflected in lateralized supratemporal cortical responses possibly concordant with hemispheric language dominance.

The interpretation of these results requires considerable caution. In particular, the sample size was small and the large-array biomagnetometer recording technique is relatively new, so we are not yet aware of all the subtleties of the data obtained with extracranial recording. Acknowledging these caveats and limitations, the pattern of results is consistent with a model that posits that early aspects of speech sound processing (e.g. up to the level of accessing a lexical buffer) are mediated bilaterally in the posterior aspect of the superior temporal gyrus in supratemporal auditory cortex. Moreover, the data support the hypothesis that there is an attention-dependent modulation of activity affecting left superior temporal cortex. Unexpectedly, the right temporal cortex was also affected by attentional

demands, although in the opposite direction to the left superior temporal cortex.

#### 4.1. Bilateral symmetry of speech sound processing

The apparent *symmetry* of early evoked responses to speech is surprising if one subscribes to a model in which the analysis of speech sounds is performed exclusively by left (or right) temporal cortical mechanisms. In particular, because cortical responses elicited by auditory stimuli can be measured as early as 15–20 ms after stimulus onset in direct recordings [5], sound recognition may be a fast process that occurs before 100 ms post-stimulus (although not necessarily as early as 20 ms). One might therefore reason that responses as late as 100 ms should reflect higher-order linguistic processing. Since there is ample evidence for left-hemisphere specialization of many language processes one might expect these type of speech stimuli to preferentially engage the left temporal cortex [32,36,37]. The data described here, however, are more compatible with a model in which speech sounds are processed in both superior temporal cortices.

Other electrophysiological research has yielded similar findings of bilateral processing. For example, Creutzfeldt et al. [7] and Ojemann [25] report on intracranial recordings performed in 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. Similarly, Boatman et al. [3] 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 posterior superior temporal gyri (STG). Recordings revealed a significant activation increase in posterior STG during speech perception while stimulation of one specific posterior STG electrode pair dramatically impaired phoneme discrimination across subjects.

Importantly, the deficit-lesion literature and the functional imaging literature both motivate the hypothesis that the left and right superior temporal gyri both construct speech sound representations, although no claim about the precise nature the computations has emerged. The neuropsychological evidence from *pure word deafness*, a rare deficit in which patients have selectively impaired speech perception (although speaking, reading, and writing remain intact) is most consistent with the position that both the left and right superior temporal gyri play an essential role in speech sound processing [30]. Similarly, a review of the older (<sup>18</sup>FDG PET and 133-Xe) and more recent (PET and fMRI) functional neuroimaging literature reveals that speech perceptual processes often activate both superior temporal gyri, regardless of whether the stimuli are single

phonemes, syllables, words, or connected speech [30], particularly in passive listening conditions. 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 strongly supports the hypothesis that speech perception proper is subserved by the posterior superior temporal gyrus in both hemispheres.

Interestingly, there are nevertheless aspects of the results that point to possible hemispheric differences. For instance, the M100 latency is virtually always slightly longer in the left hemisphere. In both syllable experiments, the left hemisphere latency was uniformly longer. Although none of the latency effects were statistically significant, this may be due to the small sample size. If the latency difference effect is real, it is quite interesting because it goes in the opposite direction than one might expect. A currently popular idea is that the left hemisphere is both specialized for speech and language and specialized for fast temporal processing [37]. If that is true, it is somewhat counterintuitive that the left auditory cortex appears to be associated with slower latencies for the major auditory evoked response component (M100) than the right.

If the decoding and integration of speech sounds is in fact mediated by both supratemporal cortices, a number of new issues are raised. Mainly, one now owes an explanation addressing at what processing level the extensively documented asymmetries for language processing arise. The data suggest that while auditory speech perception is mediated bilaterally, the further processing of speech stimuli, i.e. the input-modality independent language representations (word and sentence level) are mediated by the left hemisphere.

#### 4.2. Lateralization: hemispheric asymmetry and task dependence

The experimental hypothesis was that executing a task would selectively change processing in the left hemisphere, thus potentially obscuring the role of the right temporal cortex in the processing of speech sounds. 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. 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, and it will have to be further elucidated in future experiments.

There is prior evidence that suggests that there may be differences between auditory cortical computation condi-

tioned by the perceptual attribute attended to in an experiment. For example, Zatorre et al. [40], 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 the syllables were the same (phonetic judgment), and in the other subjects had to judge whether or not the pitch changes at the end of the syllables were the same (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 activated left hemisphere sites, notably Broca's area. In contrast, the pitch judgment condition showed activation foci more sharply in the right hemisphere, notably right inferior frontal areas. These results were suggestive, but did not resolve questions about processing in auditory cortex itself. It is still unknown whether or not primary auditory cortex and the immediately adjacent auditory association areas represent speech (and other auditory) input in the same manner, bilaterally.

Like other investigators, we found that the source localizations in the left were often more posterior than those in the right [11,19,24]. Unlike previous work, however, we did not find anterior posterior differences between stimulus types within a hemisphere: syllables (and tones and vowels) localized to the same area in supratemporal auditory cortex. Kuriki and Murase [19] recorded the evoked magnetic field generated in response to pure tones (170 Hz and 1 kHz) 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 [11] and is assumed to be associated with the substantial morphological asymmetry of the superior temporal plane in many human subjects [12] – although no one has argued that this anterior–posterior asymmetry has functional consequences. With respect to the speech stimuli, Kuriki and Murase [19] found no left–right asymmetries when looking at the waveforms generated by the vowel or the syllable. However, single 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.

Eulitz et al. [11] 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 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 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. Most analyses of the latencies and amplitudes of the major evoked response components have failed to detect any significant cortical asymmetries before the sustained field response (approximately 200 ms after stimulus onset), at least with respect to vowels.

The present 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. Insofar as cortical areas respond differentially to attentional requirements, one is in danger of systematically over- or underestimating the contribution of a given cerebral area to the processing of the domain in question. In any case, the interaction reported here predicts that in imaging studies that use subtractive methodology one could find such an effect as well. In particular, it is predicted that one will detect response decreases in right auditory association cortex and increases in left auditory association cortex on the supratemporal plane when passive listening results are subtracted from task-modulated results.

The effects of attention and arousal need to be clarified in further experiments to ascertain precisely what kind of attentional modulation is occurring. The issue one needs to pursue is precisely what kind of attentional manipulation leads to this response attenuation or enhancement. Several possibilities exist. (1) Only attending to the *particular stimulus set* changes the outcome. The effect is therefore due to subjects having to make overt linguistic discriminations on the presented material. (2) Attending to *any speech or language stimuli* 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 task will lead to such a modulation, including, say, the execution of a visual task.

Of particular importance for the interpretation of the task-induced asymmetry results is the finding of Arthur et al. [1] that an asymmetric attention effect was not found with tones. These authors recorded the auditory evoked neuromagnetic field elicited by tones. They presented target tones in a series of standards (in one ear, recording from the contralateral hemisphere) and analyzed the effect of attention on the major evoked field components. Although they found significant effects of selective attention after 200–250 ms in both hemispheres, selective attention to tone targets did not differentially engage the left and right auditory cortices at 100 ms.

One issue that requires additional consideration is the possible relation of these results to hemispheric dominance for language. It appears that the attentional requirements

shift activity to the dominant hemisphere. It is intriguing to find that one left-handed subject clearly showed a reverse pattern than the other five in the comparison between the passive and attended conditions. Since that subject's responses reversed on many other patterns, he was deemed at having a reasonable chance of being right-dominant. If task requirements such as the ones we used selectively engaged the dominant hemisphere, we may have incidentally tested hemispheric dominance, even though there were no obvious hemispheric differences in the passive condition. This is a provocative possibility.

## Acknowledgements

For their expert technical assistance we would like to thank Susanne Honma, Mary Mantle, and Dr. Leonardo Topa. For insightful critical discussion off this work we thank David Caplan, Albert Galaburda, Martha McGinnis, David Pesetsky, and Steven Pinker. Biomagnetic Technologies, Inc., provided technical assistance throughout the execution of this project. DP was supported by the McDonnell-Pew Center for Cognitive Neuroscience. Additional financial support for this project came from the NEC Corporation, NSF SGER grant #SBR-9423242, the NSF-funded RTG program 'Language: Acquisition and Computation' (#DIR9113607).

## References

- [1] Arthur, D., Lewis, P., Medvick, P. and Flynn, E., A neuromagnetic study of selective auditory attention, *Electroenceph. Clin. Neurophysiol.*, 78 (1991) 348–360.
- [2] Binder, J., Rao, S., Hammeke, T., Yetkin, F., Jesmanowicz, A., Bandettini, P., Wong, E., Estkowski, L., Goldstein, M., Haughton, V. and Hyde, J., Functional magnetic resonance imaging of human auditory cortex, *Ann. Neurol.*, 35 (1994) 662–672.
- [3] Boatman, D., Lesser, R. and Gordon, B., Auditory speech perception in the left temporal lobe: an electrical interference study, *Brain Language*, 51 (1995) 269–290.
- [4] Cansino, S., Williamson, S. and Karron, D., Tonotopic organization of human auditory cortex, *Brain Res.*, 663 (1994) 38–50.
- [5] Celesia, G., Organization of auditory cortical areas in man, *Brain*, 99 (1976) 403–414.
- [6] Cohen, J., MacWhinney, B., Flatt, M. and Provost, J., PsyScope: a new graphic interactive environment for designing psychology experiments, *Behav. Res. Meth.*, 25 (1993) 257–271.
- [7] Creutzfeldt, O., Ojemann, G. and Lettich, E., Neuronal activity in the human lateral temporal lobe, *Exp. Brain Res.*, 77 (1989) 451–475.
- [8] Démonet, J., Price, C., Wise, R. and Frackowiak, R., Differential activation of right and left posterior sylvian regions by semantic and phonological tasks: a positron-emission tomography study in normal human subjects, *Neurosci. Lett.*, 182 (1994) 25–28.
- [9] Démonet, J.-F., Chollet, F., Ramsay, S., Cardebat, D., Nespoulous, J.-L., Wise, R., Rascol, A. and Frackowiak, R., The anatomy of phonological and semantic processing in normal subjects, *Brain*, 115 (1992) 1753–1768.
- [10] Elberling, C., Bak, C., Kofoed, B., Lebech, J. and Saermark, K., Auditory magnetic fields: source location and tonotopic organization in the right hemisphere of the human brain, *Scand. Audiol.*, 11 (1982) 61–65.
- [11] Eulitz, C., Diesch, E., Pantev, C., Hampson, S. and Elbert, T., Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli, *J. Neurosci.*, 15 (1995) 2748–2755.
- [12] Geschwind, N. and Levitsky, W., Human brain: left–right asymmetries in temporal speech region, *Science*, 161 (1968) 186–187.
- [13] Hari, R., Activation of the human auditory cortex by speech sounds, *Acta Otolaryngol. (Stockh.)*, Suppl. 491 (1991) 132–138.
- [14] Hari, R., Hämäläinen, M., Ilmoniemi, R., Kaukoranta, E., Reinikainen, K., Salminen, J., Alho, K., Näätänen, R. and Sams, M., Responses of the primary auditory cortex to pitch changes in a sequence of tone pips: neuromagnetic recordings in man, *Neurosci. Lett.*, 50 (1984) 127–132.
- [15] Hari, R., Pelizzzone, M., Mäkelä, J., Hällström, J., Leinonen, L. and Lounasmaa, O., Neuromagnetic responses of human auditory cortex to on- and offsets of noise bursts, *Audiology*, 26 (1987) 31–43.
- [16] Hayashi, M., Auditory neuromagnetic fields evoked by spectral transition of syllables, *J. Robot. Mechatr.*, 5 (1993) 409–412.
- [17] Hayashi, M., Mashiko, T., Imada, T. and Odaka, K., Brain magnetic fields evoked by five Japanese vowels, *JCA*, 14 (1992) 2–3.
- [18] Kaukoranta, E., Hari, R. and Lounasmaa, O., Responses of the human auditory cortex to vowel onset after fricative consonants, *Exp. Brain Res.*, 69 (1987) 19–23.
- [19] Kuriki, S. and Murase, M., Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds, *Exp. Brain Res.*, 77 (1989) 127–134.
- [20] Mäkelä, J., Hari, R. and Leinonen, L., Magnetic responses of the human auditory cortex to noise/square wave transitions, *Electroenceph. Clin. Neurophysiol.*, 69 (1988) 423–430.
- [21] Näätänen, R., *Attention and Brain Function*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1992.
- [22] Näätänen, R., The role of attention in auditory information processing as revealed by event-related potentials and other brain measures of cognitive function, *Behav. Brain Sci.*, 13 (1990) 201–288.
- [23] Näätänen, R. and Picton, T., The N1 wave of the human electric and magnetic response to sound: a review and analysis of the component structure, *Psychophysiology*, 24 (1987) 375–425.
- [24] Nakasato, N., Fujita, S., Seki, K., Kawamura, T., Matani, A., Tamura, I., Fujiwara, S. and Yoshimoto, T., Functional localization of bilateral auditory cortices using an MRI-linked whole head magnetoencephalography (MEG) system, *Electroenceph. Clin. Neurophysiol.*, 94 (1995) 183–190.
- [25] Ojemann, G., The cortical organization of language, *J. Neurosci.*, 11 (1991) 2281–2287.
- [26] Pantev, C., Eulitz, C., Elbert, T. and Hoke, M., The auditory evoked sustained field: origin and frequency dependence, *Electroenceph. Clin. Neurophysiol.*, 90 (1994) 82–90.
- [27] Pantev, C., Hoke, M., Lehnertz, K., Lütkenhöner, B., Anogianakis, G. and Wittkowski, W., Tonotopic organization of the human auditory cortex revealed by transient auditory evoked magnetic fields, *Electroenceph. Clin. Neurophysiol.*, 69 (1988) 160–170.
- [28] Pantev, C., Hoke, M., Lütkenhöner, B. and Lehnertz, K., Tonotopic organization of the auditory cortex: pitch versus frequency representation, *Science*, 246 (1989) 246–248.
- [29] Petersen, S., Fox, P., Posner, M., Mintun, M. and Raichle, M., Positron emission tomographic studies of the processing of single words, *J. Cognit. Neurosci.*, 1 (1989) 153–170.
- [30] Poeppel, D., *The Neural Basis of Speech Perception*, MIT Dissertation, 1995, Cambridge MA.
- [31] Price, C., Wise, R., Ramsay, S., Friston, K., Howard, D., Patterson, K. and Frackowiak, R., Regional response differences within the human auditory cortex when listening to words, *Neurosci. Lett.*, 146 (1992) 179–182.
- [32] Rauschecker, J., Tian, B. and Hauser, M., Processing of complex

- sounds in the macaque nonprimary auditory cortex, *Science*, 268 (1995) 111–114.
- [33] Rif, J., Hari, R., Hämäläinen, M. and Sams, M., Auditory attention affects two different areas in the human supratemporal cortex, *Electroenceph. Clin. Neurophysiol.*, 79 (1991) 464–472.
- [34] Romani, G., Williamson, S. and Kaufman, L., Tonotopic organization of the human auditory cortex, *Science*, 216 (1982) 1330–1340.
- [35] Sams, M., Reijo, A., Hämäläinen, M., Hari, R., Lounasmaa, O., Lu, S. and Simola, J., Seeing speech: visual information from lip movements modifies activity in the human auditory cortex, *Neurosci. Lett.*, 127 (1991) 141–145.
- [36] Studdert-Kennedy, M. and Shankweiler, D., Hemispheric specialization for speech perception, *J. Acoust. Soc. Am.*, 48 (1970) 579–594.
- [37] Tallal, P., Miller, S. and Fitch, R., Neurobiological basis of speech: a case for the preeminence of temporal processing, *Ann. N.Y. Acad. Sci.*, 682 (1993) 27–47.
- [38] Woldorff, M., Gallen, C., Hampson, S., Hillyard, S., Pantev, C., Sobel, D. and Bloom, F., Modulation of early sensory processing in human auditory cortex during auditory selective attention, *Proc. Natl. Acad. Sci. USA*, 90 (1993) 8722–8726.
- [39] Yamamoto, T., Williamson, S., Kaufman, L., Nicholson, C. and Llinas, R., Magnetic localization of neuronal activity in the human brain, *Proc. Natl. Acad. Sci. USA*, 85 (1988) 8732–8736.
- [40] Zatorre, R., Evans, A., Meyer, E. and Gjedde, A., Lateralization of phonetic and pitch discrimination in speech processing, *Science*, 256 (1992) 846–849.