Available online at www.sciencedirect.com
SCIENCE DIRECT®

Cognitive Brain Research 19 (2004) 82–91

COGNITIVE BRAIN RESEARCH

www.elsevier.com/locate/cogbrainres

Research report

Intra-subject replication of brain magnetic activity during the processing of speech sounds

Carsten Eulitz[a,b,*], Jonas Obleser[a], Aditi Lahiri[b]

[a] *Department of Clinical Psychology, University of Konstanz, Giessberg 10, P.O. Box D25, 78457 Konstanz, Germany*
[b] *Department of Linguistics, University of Konstanz, Giessberg 10, P.O. Box D25, 78457 Konstanz, Germany*

Accepted 18 November 2003

Abstract

The present study examined the cortical activity during processing of vocalic segments by means of whole-head magnetoencephalography (MEG) to see whether respective cortical maps are stable across repeated measurements. We investigated the spatial configuration and temporal characteristics of the N100m generators of the auditory-evoked field during the processing of the synthetic German vowels [a], [e] and [i] across 10 repeated measurements in a single subject. Between vowels, N100m latency as well as source location differences were found with the latency differences being in accordance with tonochronic principles. The spatial configuration of the different vowel sources was related to differences in acoustic/phonological features. Vowels differing maximally in those features, i.e., [a] and [i], showed larger Euclidean distances between N100m vowel sources than [e] and [i]. This pattern was repeatable across sessions and independent of the source modeling strategy for left-hemispheric data. Compared to a pure tone control condition, the N100m generators of vowels were localized in more anterior, superior and lateral parts of the temporal lobe and showed longer latencies. Being aware of the limited significance of conclusions drawn from a single case study, the study yielded a repeatable spatial and temporal pattern of vowel source activity in the auditory cortex which was determined by the distinctiveness of the formant frequencies corresponding to abstract phonological features.

Theme: Sensory systems
Topic: Auditory systems: Central Physiology

Keywords: Magnetoencephalography; N100m; Auditory system; Vowels; Speech recognition; Language processing

1. Introduction

Magnetoencephalography (MEG) is a useful method to describe cortical activation with good spatial and an excellent temporal resolution [3,12,19]. A considerable amount of MEG studies is therefore aimed at identifying cortical neural networks supporting circumscribed cognitive processes and describing the spatio-temporal pattern of activation. In the case of speech perception, auditory-evoked magnetic fields (AEF) have been studied during all processing levels down to the acoustic–phonetic level. The latter, in particular the processing of vowels has been studied in numerous experiments [1,2,13,15,18]. An interesting question in ongoing research is which neurophysiological mechanisms and parameters are utilized by the brain to extract linguistically relevant information which is then used for the lexical access. As suggested by Näätänen and Winkler [6], the AEF component N100m, occurring around 100 ms after stimulus onset, might be the earliest component of interest to study such questions. A strong argument is that this component is the earliest to show habituation across several seconds, thus being able to index information about previously perceived acoustic features for a longer period.

Earlier studies have shown that timing (tonochrony) and topography (phonemotopy) of the N100m may reflect differences in the quality of vowels. The tonochrony hypothesis (for a review, see Ref. [15]) suggests the peak latency of the N100m to be a relevant parameter reflecting the encoding of acoustic features of the incoming auditory signal. Although this piece of information is achieved at a

---

* Corresponding author. Tel.: +49-7531-884260; fax: +49-7531-884662.
E-mail address: Carsten.Eulitz@uni-konstanz.de (C. Eulitz).

0926-6410/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cogbrainres.2003.11.004

pre-linguistic level of processing, it might among others index the quality of vowels. As a further mechanism for encoding the quality of vowels, the activation of specialized neural networks in the vicinity of the auditory cortex (phonemotopy-hypothesis) has been discussed [1,2,9]. Earlier MEG studies have shown that at least some of the N100m source parameters, such as orientation of the equivalent current dipole (ECD) source are different during the processing of vowels as compared to a sine tone stimulus [2,18]. Support for the phonemotopy-hypothesis (but not for the N100m as the relevant ERP-component) comes from animal studies showing differential firing patterns for different behaviorally relevant utterances [17,20] and even the processing of learned vowels in gerbils [10]. With respect to the N100m, comparisons between different vowels have demonstrated that the number of distinctive acoustic and phonological features is correlated with the Euclidean distance between ECD sources [1,9]. Obleser et al. [8] have also demonstrated the reliability of such distance pattern in a repeated measurement. It has been often argued that large interindividual differences are caused by variability of vowel-sensitive neural substrates because the perception of phonetic stimuli is optimized in the course of language learning and the specialized neural networks supporting these functions are shaped individually. Therefore, it is reasonable to measure spatial distinctiveness of sources independently of interindividual differences in the spatial configuration of sources by means of Euclidean distances. Nevertheless, one has to be aware of the fact that the Euclidean distance is a measure of variability as well, i.e., a less-confident source location in one condition can mimic a larger distance relative to other conditions. Consequently, there are two ways to interpret a larger Euclidean distance, (i) spatially more distinct sources with an interindividually different neuroanatomy and neurophysiology underneath or (ii) just a less-confident source location in a particular experimental condition for whatever reasons. One way to demonstrate and quantify the repeatability of vowel source configurations is a single case study. Factors like interindividual differences in the neuroanatomy or the location of neural structures for speech sound processing will not affect the results. The present study was designed to test which of the abovementioned interpretations for differential Euclidean distances are more plausible.

To compare the results for the processing of vowels with earlier replicability studies with tonal stimuli [12,19], and to have additional information about activational differences during the processing of pure tones and acoustically complex stimuli (thus, extending the scope of conditions compared to Ref. [8]), a tone condition was introduced. The present study examined a variety of signal- and source-space parameters to test the repeatability of the N100m peak-latency and the ECD source location differences across 10 MEG measurements during the processing of the German vowels [a], [e], [i] as well as a 1-kHz sine tone for reference.

2. Materials and methods

2.1. Subjects

One 26 years old, female subject participated in the experiment. She was right-handed, as ascertained by the Edinburgh Handedness Questionnaire [11] and had no history of neurological, psychiatric, or otological disorders. Individual hearing thresholds for both ears were determined individually and for each stimulus prior to the initial testing. The subject gave written informed consent after the nature of the experimental procedures was explained and was paid US$250 for participation. The study was conducted in compliance with the Declaration of Helsinki and the standards established by the local ethics committee.

2.2. Stimulation

The critical stimuli were the three German vowels [a] (as in "f*a*ther"), [e] (similar to "b*ai*t" or "b*a*y" but not diphthongized) and [i] (as in "b*ea*t"). A fourth vowel, a long schwa-like vowel, was used as a target stimulus to assure that the subject attended constantly to the stimuli and maintained a phonetic processing mode. All phonetic stimuli were synthetic [5]. When listening to the critical vowels for the first time, the subject easily identified them as the corresponding German vowels. During the experimental study, the subject listened to a pseudo-random sequences of the four vowels, where the critical vowels occurred with a probability of 31% each and the schwa with 7%, and was asked to press a button with her right index finger whenever the target occurred. The vowels were delivered in three sequences consisting of 520 vowel stimuli presented with a randomized stimulus onset asynchrony of 2.0 ± 0.2 s. All vowels had a fundamental frequency F_0 of 129 Hz, falling linearly to 119 Hz and equal frequencies for the F_4 (3900 Hz) and F_5 (4700 Hz). They differed in the formant frequencies F_1 to F_3, as depicted in Table 1. After each sequence of vowels, 50 1-kHz sine tones (same ISI), i.e., 150 for the whole experimental session, were presented to get an additional landmark for spatially relating vowel evoked brain responses to a functional reference known to be localized in the auditory cortex. All stimuli had a sampling frequency of 10 kHz, a duration of 600 ms, a 50-ms gaussian onset ramp and a 150-ms gaussian offset ramp.

They were presented binaurally at 50 dB SPL above respective hearing threshold via a non-magnetic and echo-

Table 1
Format frequencies of vowel stimuli used

	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)
[a]	780	1250	2600
[e]	370	2250	2800
[i]	250	2700	3400
Schwa	350	1400	2500

free stimulus delivery system (with almost linear frequency characteristics in the critical range of 200–4000 Hz). Recordings were made in a sound-attenuated and magnetically shielded room (Vaccumschmelze). During the experimental session, the subject watched silent videos in order to maintain constant vigilance and to reduce excessive eye movements.

Exactly the same experimental session was repeated 10 times, one session every 2 days. Great care was taken to keep experimental conditions as constant as possible. All measurements were conducted by the same research assistant and were obtained in a supine position in order to minimize movement artifacts. The positioning of the sensor relative to the subject's head and the resting position were held as equal across sessions as possible using a mechanical reference system mounted to the dewar. The recordings were always made between 10 AM and noon.

2.3. Data acquisition and data analyses

Auditory-evoked magnetic fields (AEFs) evoked by the three different vowel stimuli and the sine tone were recorded using a whole head neuromagnetometer (MAGNES 2500, 4D Neuroimaging). Epochs of 1200-ms duration (including a 200-ms pre-trigger baseline) were recorded with a bandwidth from 0.1 to 200 Hz and a 678.17-Hz sampling rate. If the peak-to-peak amplitude exceeded 3.5 pT or the co-registered EOG-signal was larger than 100 μV in one of the channels, epochs were rejected. Epochs containing button presses were also excluded. About 250 to 480 artifact-free epochs for every session and critical vowel were averaged after off-line noise correction. A 20-Hz low-pass filter (Butterworth 12 dB/oct, zero-phase shift) was subsequently applied to the average.

Data analyses focused on the N100m component. N100m peak amplitude was calculated as the maximum Root Mean Square (RMS) over 34 magnetometer channels selected to include the field extrema over the left and the right hemisphere, respectively. N100m peak latency was defined as the sampling point where the corresponding RMS reached the maximum. The source space results were derived in two different ways using a one-dipole model as well as a spatio-temporal four-dipole model. First, an equivalent current dipole (ECD) in a spherical volume conductor (fitted to the shape of the regional head surface) was modeled at every sampling point separately for the left and the right hemisphere. The channel selections over the left and the right hemispheres were the same as used for the definition of the signal space parameters. Source space parameters, such as the location and orientation of the ECD as well as the dipole moment were defined as the median of five successive sampling points in the rising flank of the N100m, i.e., the median for the 7.4 ms before the peak. To include the values for calculating the median, a number of criteria had to be fulfilled: (1) Goodness of fit greater than 0.90, (2) ECD location larger than 1.5 cm in medial–lateral direction from the center of the brain and 3–8 cm in superior direction, measured from the connecting line of the pre-auricular points.

As the signal power of the N100m over the right hemisphere was considerably smaller than over the left hemisphere (especially for the vowels), the source modeling with a single ECD model did not deliver satisfactory results for the right hemisphere. As a consequence, we used a more general source model, i.e., a spatio-temporal four-dipole model. The four ECDs were used to simultaneously explain the activity from 100- to 240-ms post stimulus onset in both hemispheres. First, two dipoles were introduced to account for the activity from 100 to 140 ms, i.e., the N100m onset. These two dipoles were then fixed in location and orientation and the two other dipoles were introduced to model the residual activity for the whole latency range. Finally, all four dipoles got all degrees of freedom and had to account for the activity in the latency range from 100 to 240 ms. The outcome of this final analysis, in particular that of the two N100m sources, was used for further analyses.

To test the repeatability of temporal and spatial characteristics of the N100m, peak amplitude and latency results from all 10 observations were submitted to repeated measures 2 × 4 ANOVA with factors hemisphere and stimulus (1-kHz tone, [a], [e], [i]). Single ECD source strength and locations of the left hemisphere only were submitted to a univariate analysis with repeated factor stimulus (1-kHz tone, [a], [e], [i]).

Furthermore, the relative source configurations were evaluated by calculating the Euclidean distances between the three vowel ECD sources separately for each hemisphere and each experimental session. The comparison of Euclidean distances was additionally employed to elucidate invariant spatial relationships between vowel sources irrespective of absolute localization errors [1].

Significant differences were further explored by means of contrast analyses: For all dependent variables, the tone condition was contrasted with the mean of all three vowel conditions, and vowel conditions were compared to each other in pairwise contrasts. Where appropriate, p-values were Greenhouse–Geisser corrected.

3. Results

3.1. Signal space results

A summary of the averaged waveforms for all experimental conditions and all sessions is shown in Fig. 1. The two channels with maximum incoming and outgoing magnetic field flux over the left and the right hemisphere were chosen. Across sessions and vowel conditions, the same channels were selected. For the tone condition, the maximum field strength was obtained over neighboring channels which were also used to present the waveforms across

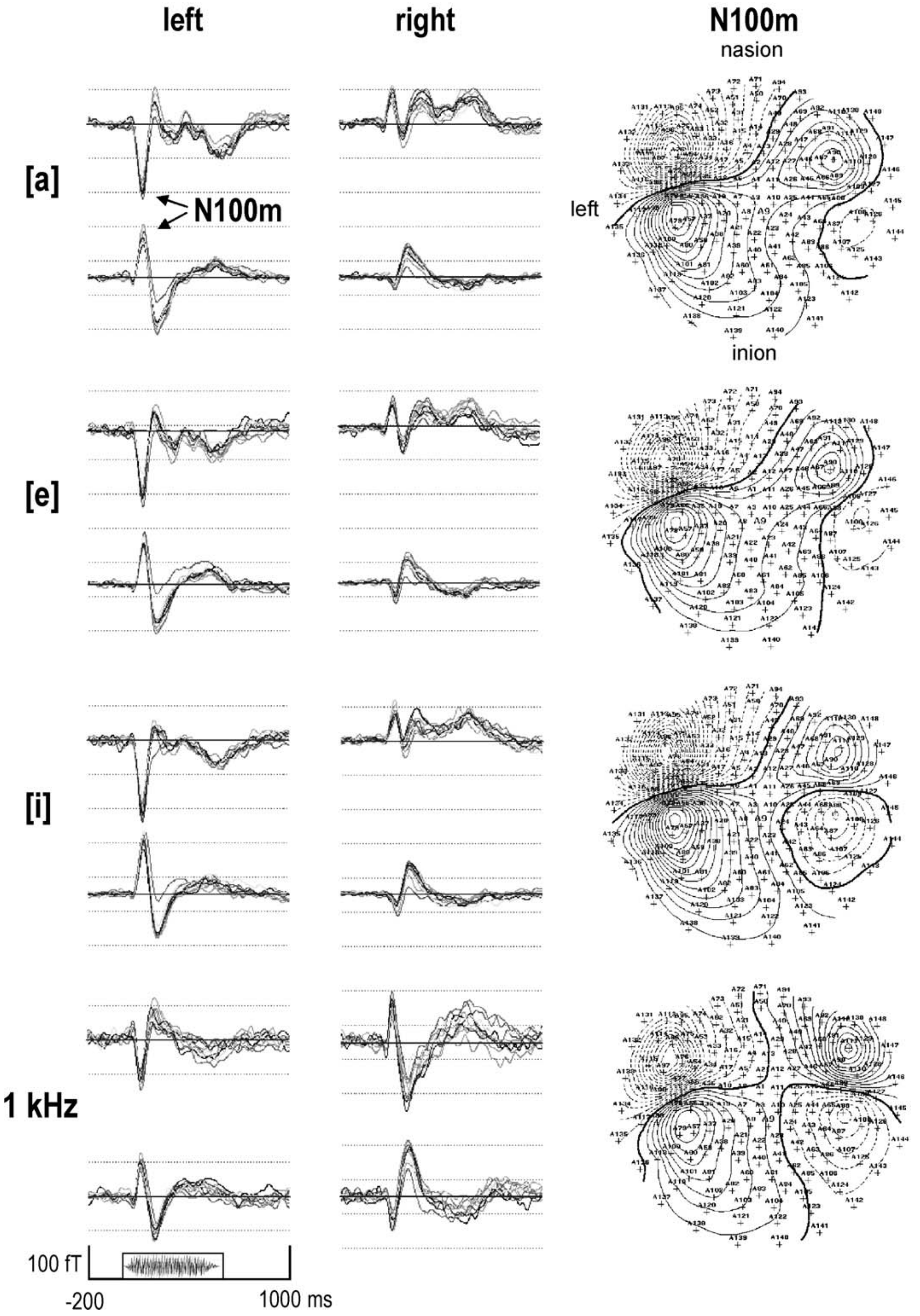

Fig. 1. Waveforms from channels measuring maximum incoming and maximum outgoing field are shown. Results of all 10 sessions are plotted separately for each condition (panels) and each hemisphere. The right column shows isofield contour plots for the N100m peak latency for each condition using a contour step of 10 fT.

sessions (Fig. 1). The slight shift in N100m topography is also seen in the isocontour map for that particular component which is shown in the left column. Peak latency and amplitude of the N100m was similar across sessions. There was one session differing slightly from all others but the difference was more marked for the P200 m. As the same magnetic sensor was chosen for all sessions, marginal differences in positioning of the subject’s head may have led to such differences. Analyses in the source space are more robust against the variability in the positioning.

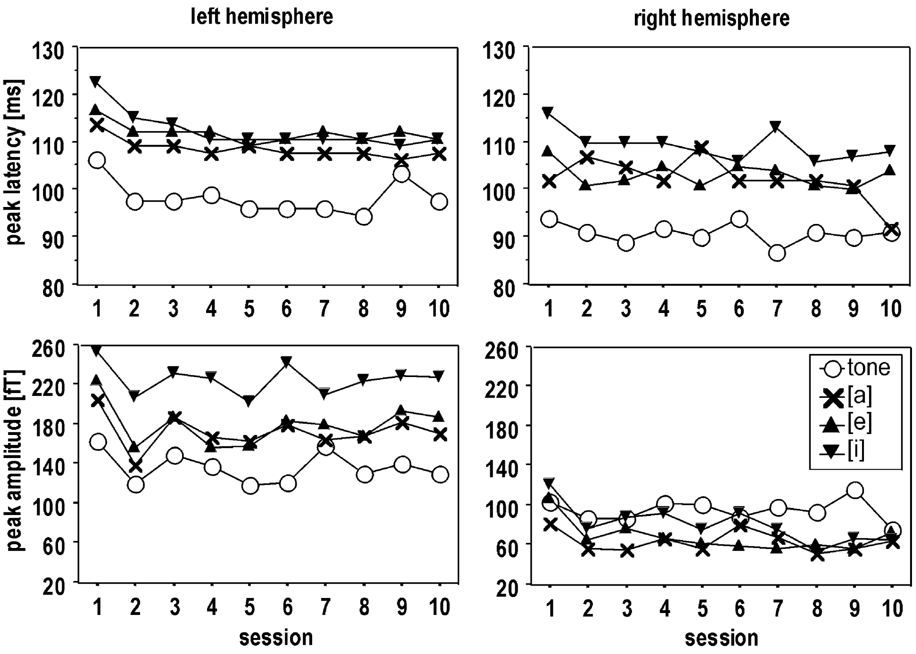

Fig. 2. N100m RMS peak latency (upper panel) and peak amplitude (lower panel) for all 10 sessions are plotted separately for both hemispheres (left and right column). Note the earlier N100m latency in both hemispheres and the less asymmetric RMS amplitude in the tone condition (open circles).

The N100m amplitude in the vowel conditions was larger over the left as compared to the right hemisphere. This interhemispheric difference was not as strong for the tone condition: An interaction of hemisphere and stimulus type emerged [$F(3,27) = 108.86$, $\varepsilon = 0.86$, $p < 0.0001$]. If referring to the N100m/P200m amplitude difference, this interhemispheric difference was even more pronounced for the vowel conditions and the tone condition showed a right-hemispheric preponderance. Looking at the left hemisphere only, the vowel conditions showed a larger N100m amplitude than the tone condition [$F(3,27) = 146.56$, $\varepsilon = 0.75$, $p < 0.0001$]—contrast analyses proved the tone response to be significantly weaker ($p < 0.0001$). The reverse was true when analyzing the right hemisphere only: the ANOVA also revealed a main effect of stimulus type [$F(3,27) = 15.51$, $\varepsilon = 0.75$, $p < 0.0001$], but tone responses were stronger than mean vowel responses ($p < 0.0001$; see Fig. 2).

As shown in the right column of Fig. 1, the isocontour maps of the N100m component showed dipolar field patterns over both hemispheres, thus source analyses using a single ECD model is plausible.

The N100m latency was markedly shorter for the tone condition [main effect of condition: $F(3,27) = 124.4$, $\varepsilon = 0.72$, $p < 0.0001$; tone vs. mean vowel contrast $p < 0.0001$] in both hemispheres. Furthermore, the latency of N100m in the right hemisphere was shorter than that in the left hemisphere for all conditions. Among the vowel conditions, the N100m latency was markedly longer for the [i] ([i] vs. mean other conditions $p < 0.01$) in both hemispheres (see Fig. 2). However, the latency differences between vowel conditions were overall not symmetric across hemispheres. In the left hemisphere, the latency of the [e] was close to the latency of [i] (contrasts: [e] vs. [i] ns; [e] vs. [a] $p < 0.001$) whereas in the right hemisphere it was close to the latency of [a] (contrasts: [e] vs. [a] ns; [e] vs. [i] $p < 0.001$; see also Table 2).

3.2. Source space results

Source analyses with a single ECD model applied separately for the left- and right-hemispheric data revealed reasonable results only for the left-hemispheric data (obviously, the signal power for right-hemispheric data was too small). Therefore, the presentation of results will be confined to left-hemispheric results. The mean ECD locations and orientations overlaid onto the individual Magnetic

Table 2
Descriptive statistics for N100m latency and left-hemispheric ECD source parameters (mean ± S.D.)

Parameter	[a]	[e]	[i]	1-kHz tone
Latency left (ms)	108.8 ± 2.00	112.0 ± 1.93	112.6 ± 3.91	98.5 ± 3.69
Latency right (ms)	102.4 ± 4.50	103.1 ± 2.51	109.4 ± 3.17	90.9 ± 2.13
Posterior–anterior (cm)	1.37 ± 0.19	1.23 ± 0.15	1.40 ± 0.18	0.95 ± 0.31
Medial–lateral (cm)	5.48 ± 0.33	5.32 ± 0.17	5.18 ± 0.28	4.99 ± 0.54
Inferior–superior (cm)	6.87 ± 0.24	6.57 ± 0.23	6.44 ± 0.27	6.18 ± 0.35
|Q| (nAm)	27.38 ± 3.46	30.23 ± 3.97	41.21 ± 5.08	28.49 ± 7.02

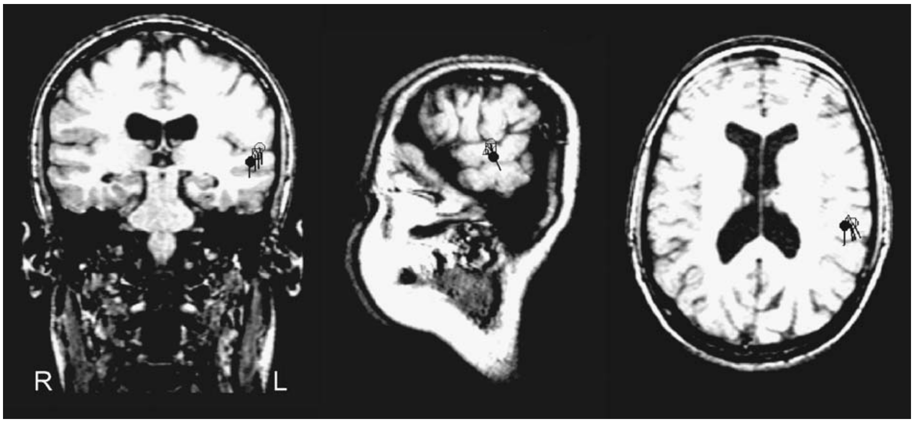

Fig. 3. Mean left-hemispheric ECD solutions across sessions for the tone (filled circle) and vowel (open symbols; circle for [a], square for [e], triangle for [i]) conditions are projected onto the subject's magnetic resonance image. Note the spatial separation and differential orientation of the vowel sources compared to the tone source.

Resonance images are shown in Fig. 3. The center of activity is located in the vicinity of the auditory cortex but is rather close to the superior temporal sulcus than to the Sylvian fissure. Fig. 4 summarizes the source space results for the N100m (see also Table 2). As demonstrated by the goodness of fit and the confidence volume, the source localizations were reasonably good across all sessions. The slightly worse data for the tone conditions illustrates

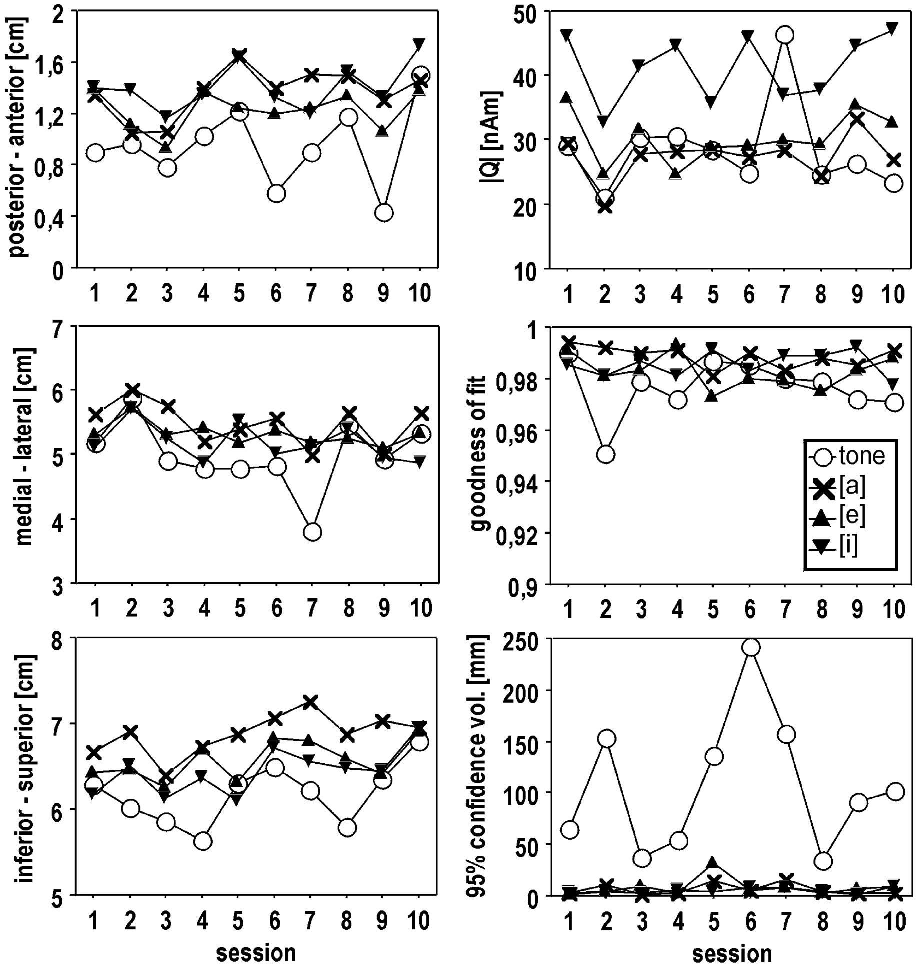

Fig. 4. For the left hemisphere, ECD solutions for all 10 sessions are displayed. The left column depicts source location parameters in posterior–anterior (upper panel), medial–lateral (middle panel) and inferior–superior direction (lower panel). The right column shows the ECD Source Strength |Q| (upper panel), the goodness of fit (middle panel) and the confidence volume (lower panel).

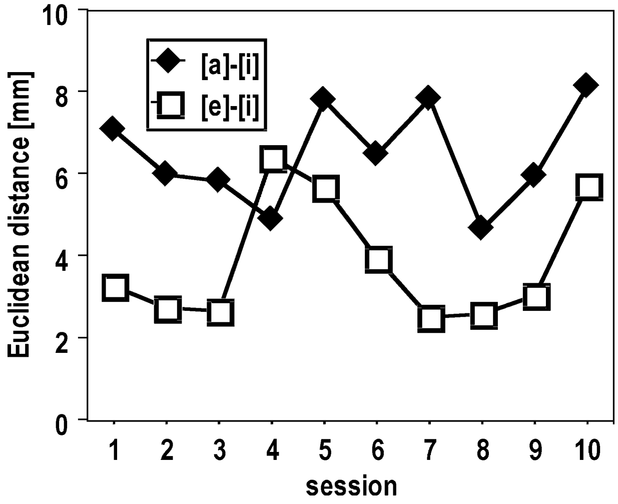

Fig. 5. Euclidean distances between N100m sources (single ECD model) of the acoustically most dissimilar vowels [a] and [i] (filled diamonds) and of the more similar vowels [e] and [i] (open squares) is shown for all 10 sessions. The distance [a]–[i] is systematically larger than the distance [e]–[i].

the sensitivity of the applied source localization technique to the signal power of input data.

The dipole moment was systematically higher for [i] compared to all other conditions. The dipole moment in the tone condition was not different from the [a] and [e] conditions, however, the location was different—thereby explaining the differences between these conditions seen for the (channel-selection dependent) RMS data. The ANOVA for the dipole moment confirmed overall stimulus differences [$F(3,27)=20.57$, $\varepsilon=0.50$, $p<0.001$] and proved the [i] source to be significantly stronger than all three other conditions ($p<0.01$).

The ECD location parameters all exhibited significant main effects of stimulus type. Along the posterior–anterior axis [$F(3,27)=19.04$, $\varepsilon=0.63$, $p<0.0001$], the tone source was more posterior than all three vowel sources ($p<0.001$). The most posterior vowel source [e] was significantly more posterior than [a] and [i] sources (both $p<0.02$). The [a] and [i] sources did not differ significantly.

Along the medial–lateral axis [$F(3,27)=6.46$, $\varepsilon=0.60$, $p=0.01$], the tone source was located deepest (Table 2), differing significantly from the mean across vowel sources ($p<0.05$). Furthermore, a significant difference between the depth of [a] and [i] sources was found ($p<0.01$).

The inferior–superior axis [$F(3,27)=23.76$, $\varepsilon=0.65$, $p<0.0001$] differentiated tone from vowel sources ($p<0.01$), but also revealed differences between all three vowel sources with [a] more superior than [e] as well as [e] more superior than [i] (all $p<0.01$).

The orientation of dipoles differed between conditions in the sagittal [$F(3,27)=29.63$, $\varepsilon=0.60$, $p<0.0001$], the axial [$F(3,27)=60.01$, $\varepsilon=0.69$, $p<0.0001$], as well as the coronal plane [$F(3,27)=37.20$, $\varepsilon=0.45$, $p<0.0001$]. The orientation differences were most pronounced between tone and the vowel conditions. Among vowel sources, the [a] source differed from both the [e] and [i] sources.

As illustrated in Fig. 4, despite the quite consistent differences between conditions, the source localization results showed a considerable variance across sessions. To describe the consistency of the source location pattern, Euclidean distances between vowels were calculated (Fig. 5). A significant distance difference was revealed [$F(3,18)=8.27$, $\varepsilon=0.77$, $p<0.01$], with [a]–[i] (0.67 cm) being significantly larger than [e]–[i] (0.39 cm) across sessions ($p<0.001$). The third, fully determined distance [a]–[e] amounted on average to 0.48 cm and was also significantly smaller than [a]–[i] ($p<0.01$) but did not differ from [e]–[i].

As an alternative way of source analysis, a spatio-temporal model was used to explain the magnetic field distributions using all recording channels for a selected latency range in both hemispheres simultaneously. Fig. 6 demonstrates that with this source localization technique the [a]–[i] distance was also larger than [e]–[i] in most of the sessions. This distance pattern was seen for both hemispheres. The larger variability of the Euclidean distances across sessions in the right hemisphere compared to the left

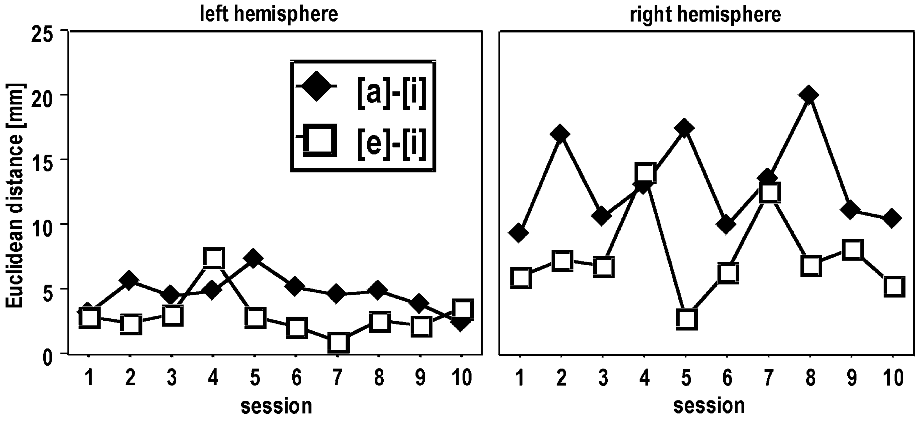

Fig. 6. Euclidean distances between N100m sources (spatio-temporal model) of [a] and [i] vs. [e] and [i] are shown for all 10 sessions (for symbols, see Fig. 5).

one illustrates again the sensitivity of this measure to the signal power of the AEF data.

Besides a main effect of hemisphere [$F(1,9) = 78.33$, $p < 0.001$] with right hemispheric distances being larger, a main effect of distance emerged [$F(2,18) = 6.17$, $\varepsilon = 0.63$, $p < 0.03$]. Contrast analyses proved the [a]–[i] distance (left hemisphere 0.46 cm, right hemisphere 1.33) again to be larger than that of [e]–[i] (left hemisphere 0.31 cm, right hemisphere 0.77 cm) ($p < 0.01$). The hemisphere by distance interaction was not significant.

4. Discussion

The repeatability of the N100m peak-latency and the ECD source location differences across 10 MEG measurements in one subject was analyzed to crosscheck the plausibility of the tonochrony and phonemotopy hypotheses about the functional organization of the auditory cortex. Both the latency and the source parameters of the N100m not only delivered support for these hypotheses, but also indicated the limitations of the MEG technique to study these questions. First, we discuss the differences between tones and vowels in general, showing that processing of speech sounds takes place at least partly in different brain regions than processing of pure tones—at least in German where pitch (corresponding to tone in a linguistic sense) is not a linguistic feature to distinguish between lexical items. Second, we discuss the differences between vowels to draw conclusions with respect to the hypotheses mentioned above.

Before doing so, let us briefly comment on the consistently left-lateralized N100m amplitude and source strength in the vowel condition which was not as strong for the tone condition. Previous studies often failed to show hemisphere-specific effects for speech segments on early components such as the N100 [2,13]. However, the reason for this could have been the mixing (near to balance) of the gender of subjects. A recent study by Obleser et al. [7] indicated that the hemispheric asymmetry of the N100m is dependent on subject's gender, i.e., only female subjects showed a left-hemispheric-dominant N100m amplitude. The data of the present subject also fits into this dominance pattern. The fact that the dominance pattern was not paralleled by the tone condition (which was absent in Obleser's et al. study) supports the notion that the left-dominant pattern is stimulus specific, at least specific to spectrally complex sounds. It may even be specific to speech sounds which has to be tested in further experiments with spectrally complex non-speech sounds.

4.1. Tone–vowel differences

Contrary to earlier studies [1,2], pure tones of the control condition were presented in separate blocks to avoid interference with the processing of speech sounds which were task relevant as well. The purpose of the tone condition to serve as an anatomical and functional landmark was assumed to be best fulfilled using this block design. The N100m peak latency was earlier for the tone condition and the generator was located more posterior, inferior and medial relative to the vowel sources. Furthermore, the tone ECD was oriented differently compared to the vowel ECDs. These results replicate earlier reports showing that processing of acoustic information in complex stimuli takes longer and involves different neural substrates than the processing of simple tones [2,16][1]. As known from studies in animals [4] and imaging studies in human subjects [21], the processing of pure tones is mostly restricted to the auditory core areas whereas band-passed noise and species-specific vocalizations activate the surrounding belt areas as well. In the studies of Rauschecker and Tian [14], the anterolateral belt areas were most specific for monkey calls and interpreted to be the auditory "what" system. Under the assumption that the tone condition in our study indexes the auditory core area, the displacement of the vowel sources relative to the core area is anterolateral, i.e., might be interpreted as an activation of the auditory "what" system, at least in this particular subject, and matches results from a recent examination with syllables and band-passed noises (Obleser et al., submitted for publication).

A difference in source locations between tone and vowel conditions has not always been found when using the source locations in head frame coordinates [2,8,18]. It has been often argued that this is caused by the interindividual variability of vowel-sensitive neural substrates because the perception of phonetic stimuli is optimized in the course of language learning and the specialized neural networks supporting these functions are shaped individually. The present study provides empirical support for this interpretation, as the spatial relationship between tone and vowel sources in head frame coordinates was stable across repeated measurements in a single subject.

4.2. Tonochrony hypothesis

The N100m latency differences between vowels were stable and in line with previous results [15]. The vowels were partly different compared to those in previous studies. Here, we investigated the vowel [e] instead of [u]. It was the [e] which showed an interesting interhemispheric difference. In the left hemisphere, its latency was close to the [i] and in the right hemisphere closer to the [a], i.e., only in the left

---

[1] While interpreting source localization differences one has to be aware that dipole locations only represent the center of gravity of active cortical tissue. These location differences do not necessarily prove distinct, mutually exclusive phonemic centers, representing one and only one specific vowel. They may also result from overlapping and differently shaped activated cortical patches—similar to those in Ohl and Scheich's study [10] where the $F_1 - F_2$ difference was correlated with the extent of the activated stripe in the auditory cortex.

hemisphere the phonologically closer related stimuli showed a temporally more similar behavior of the N100m. As to whether or not this reflects a relevance of the tonochrony beyond a pre-linguistic processing level has to be demonstrated in future studies.

4.3. Phonemotopy hypothesis

In our subject, systematic differences in source locations between vowels were found. The [a] and [i] sources were more anterior than [e], the [i] was more medial than [a], and in the inferior–superior direction [a] was most superior, [i] most inferior and [e] in-between but significantly different from both. This means that along two axes—the medial–lateral and the inferior–superior—the spatial order of vowels is in accordance with their spectral dissimilarity in the F_1–F_2 dimension. This regularity cannot be obtained in the anterior–posterior direction. Here, it seems that the [e] which is phonologically not specified for tongue height, is localized distantly from the other vowels and closer to the tone. However, a generalization of that result is premature, as a group study with the same vowels (but without the tone) did not show this difference for the corresponding head frame coordinates [8]. Their study revealed source location differences just in the inferior–superior direction rather supporting the notion of a considerable interindividual variability of the location of neural substrates for vowel processing.

The spatial distinctiveness of sources independently on interindividual differences in the spatial configuration of sources can be analyzed by means of Euclidean distances [1,8]. These distances revealed another stable spatial relationship between vowel sources: The acoustically and phonologically most dissimilar vowels [a] and [i], differing most markedly in F_1 and F_2, led in 9 out of 10 sessions to source distances larger than those between the phonologically more similar vowels [e] and [i]. This relationship was also evident when using spatio-temporal source modeling. Furthermore, a group study with the same stimuli and a similar experimental setup led to a strikingly similar result [8]. This finding that the cortical map is determined by the distinctiveness of the formant frequencies corresponding to abstract phonological features is to our knowledge unparalleled in previous human electrophysiological studies. An earlier study [1] yielded a tendency towards a phonemotopic organization of N100m sources (the [u]–[i] distance was larger than the [e]–[ɛ] distance in 60% of all cases and conditions). A reason for failing statistical significance might have been the comparably poor signal to noise ratio achieved with 100–120 averages. The same argument may explain the missing phonemotopy of the N100m in the auditory cortex in other similarly designed studies [13,18]. Support for the notion of a phonemotopic organization of the auditory cortex comes also from a recent fMRI study: Using CV-syllables, Zielinski et al. [22] detected distinct but contiguous activational centers for the different syllables in the supratemporal plane.

As mentioned in Introduction, one has to be aware of the fact that the Euclidean distance is a measure of variability as well (see for instance right hemispheric results in Fig. 6). Larger Euclidean distances can therefore be interpreted as (i) spatially more distinct sources or (ii) just a less confident source location in a particular experimental condition. The present study demonstrates that the preferred interpretation, i.e., (i), can indeed be true for studies of vowel processing in the auditory cortex.

The location of vowel sources in head frame coordinates changes by several millimeters across sessions. One of the main questions of the present study was weather this reflects rather biological and technical noise or has it a biological meaning? It might for instance reflect a dynamic recruitment of changing neural substrates for the processing of speech signals. There are several arguments for the first point of view. (1) The standard deviations of source locations across sessions are in the range of the 95% confidence intervals for the source location in a single session. (2) The variability of source locations in the pure tone condition is similar to those for the vowels. Furthermore, the variability is in a similar range as the one for the pure tone localizations in the study of Pantev et al. [12]. As pure tones are known to have stable tonotopically organized cortical representations mainly restricted to the auditory core area [4,21], we can conclude that the variability of source location parameters across sessions results rather from biological and technical noise and is not an index of the dynamic formation of always slightly differing vowel representations. Hence, the cortical areas for processing of complex acoustic stimuli, such as vowels as indexed by the N100m component are stable, at least within one subject. The problem to find stable source location differences between N100m vowel sources when using head frame coordinates in previous studies [1,2,8] is probably indeed an index of the interindividual variability of the neural network structure supporting an acquired function, such as speech perception.

5. Conclusion

In sum, in this single case study we found support for both tonochrony as well as phonemotopy as functional principles to encode incoming speech signals. Stable spatial relations between vowels as well as between vowels and a pure tone were found for the N100m sources, independent of the source modeling strategy. However, to fully understand the functional meaning of N100m latency and location differences, more vowels varying in other phonological features have to be studied. We are aware of the limited significance of the conclusions one can draw from a single case study. Nevertheless, the study yielded a repeatable spatial and temporal pattern of vowel source activity in the auditory cortex which was determined by the distinctiveness of the formant frequencies corresponding to abstract phonological features.

Acknowledgements

Research was supported by the Deutsche Forschungsgemeinschaft. The authors wish to thank Michaela Schlichtling, Valerie Zitzlsperger and Ursula Lommen for their help during data acquisition. We especially thank Eugen Diesch for supplying the stimulus material.

References

[1] E. Diesch, C. Eulitz, S. Hampson, B. Ross, The neurotopography of vowels as mirrored by evoked magnetic field measurements, Brain Lang. 53 (1996) 143–168.

[2] C. Eulitz, E. Diesch, C. Pantev, S. Hampson, T. Elbert, Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli, J. Neurosci. 15 (1995) 2748–2755.

[3] R. Hari, S. Levanen, T. Raij, Timing of human cortical functions during cognition: role of MEG, Trends Cogn. Sci. 4 (2000) 455–462.

[4] J.H. Kaas, T.A. Hackett, M.J. Tramo, Auditory processing in primate cerebral cortex, Curr. Opin. Neurobiol. 9 (1999) 164–170.

[5] D.H. Klatt, Software for a cascade/parallel formant synthesizer, J. Acoust. Soc. Am. 67 (1980) 971–995.

[6] R. Naatanen, I. Winkler, The concept of auditory stimulus representation in cognitive neuroscience, Psychol. Bull. 125 (1999) 826–859.

[7] J. Obleser, C. Eulitz, A. Lahiri, T. Elbert, Gender differences in functional hemispheric asymmetry during processing of vowels as reflected by the human brain magnetic response, Neurosci. Lett. 314 (2001) 131–134.

[8] J. Obleser, T. Elbert, A. Lahiri, C. Eulitz, Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies, Cogn. Brain Res. 15 (2003) 207–213.

[9] J. Obleser, A. Lahiri, C. Eulitz, Magnetic brain response mirrors extraction of phonological features from spoken vowels, J. Cogn. Neurosci. 16 (2004) 31–40.

[10] F.W. Ohl, H. Scheich, Orderly cortical representation of vowels based on formant interaction, Proc. Natl. Acad. Sci. U. S. A. 94 (1997) 9440–9444.

[11] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia 9 (1971) 97–113.

[12] C. Pantev, C. Gallen, S. Hampson, S. Buchanan, D. Sobel, Reproducibility and validity of neuromagnetic source localization using a large array biomagnetometer, Am. J. E.E.G. Technol. 31 (1991) 81–101.

[13] D. Poeppel, C. Phillips, E. Yellin, H.A. Rowley, T.P. Roberts, A. Marantz, Processing of vowels in supratemporal auditory cortex, Neurosci. Lett. 221 (1997) 145–148.

[14] J.P. Rauschecker, B. Tian, Mechanisms and streams for processing of "what" and "where" in auditory cortex, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 11800–11806.

[15] T.P. Roberts, P. Ferrari, S.M. Stufflebeam, D. Poeppel, Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception, J. Clin. Neurophysiol. 17 (2000) 114–129.

[16] B. Rockstroh, J. Kissler, B. Mohr, C. Eulitz, U. Lommen, C. Wienbruch, R. Cohen, T. Elbert, Altered hemispheric asymmetry of auditory magnetic fields to tones and syllables in schizophrenia, Biol. Psychiatry 49 (2001) 694–703.

[17] B. Tian, D. Reser, A. Durham, A. Kustov, J.P. Rauschecker, Functional specialization in rhesus monkey auditory cortex, Science 292 (2001) 290–293.

[18] M. Vihla, O.V. Lounasmaa, R. Salmelin, Cortical processing of change detection: dissociation between natural vowels and two-frequency complex tones, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 10590–10594.

[19] J. Virtanen, J. Ahveninen, R.J. Ilmoniemi, R. Naatanen, E. Pekkonen, Replicability of MEG and EEG measures of the auditory N1/N1m-response, Electroencephalogr. Clin. Neurophysiol. 108 (1998) 291–298.

[20] X. Wang, M.M. Merzenich, R. Beitel, C.E. Schreiner, Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: temporal and spectral characteristics, J. Neurophysiol. 74 (1995) 2685–2706.

[21] C.M. Wessinger, J. VanMeter, B. Tian, J. Van Lare, J. Pekar, J.P. Rauschecker, Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging, J. Cogn. Neurosci. 13 (2001) 1–7.

[22] B.A. Zielinski, J.P. Rauschecker, Phoneme-specific functional maps in the human superior temporal cortex, Soc. Neurosci. Abstr. 26 (2000) 1969.