

Magnetic Brain Response Mirrors Extraction of Phonological Features from Spoken Vowels

Jonas Obleser, Aditi Lahiri, and Carsten Eulitz

Abstract

■ This study further elucidates determinants of vowel perception in the human auditory cortex. The vowel inventory of a given language can be classified on the basis of phonological features which are closely linked to acoustic properties. A cortical representation of speech sounds based on these phonological features might explain the surprisingly inverse correlation between immense variance in the acoustic signal and high accuracy of speech recognition. We investigated timing and mapping of the N100m elicited by 42 tokens of seven natural German vowels varying along the phonological features tongue height (corresponding to the frequency of the first formant) and place of articulation (corresponding to the frequency of the second and third formants). Auditory-

evoked fields were recorded using a 148-channel whole-head magnetometer while subjects performed target vowel detection tasks. Source location differences appeared to be driven by place of articulation: Vowels with mutually exclusive place of articulation features, namely, coronal and dorsal elicited separate centers of activation along the posterior–anterior axis. Additionally, the time course of activation as reflected in the N100m peak latency distinguished between vowel categories especially when the spatial distinctiveness of cortical activation was low. In sum, results suggest that both N100m latency and source location as well as their interaction reflect properties of speech stimuli that correspond to abstract phonological features. ■

INTRODUCTION

When listening to speech, the human brain copes admirably with poor acoustic conditions and high variance across as well as within speakers. To this day, the mechanisms allowing such an effortless decoding of the speech signal are barely understood (Lahiri & Reetz, 2002; Sussman, 2000; Fitch, Miller, & Tallal, 1997).

Invasive and noninvasive research has revealed several principles of functional organization in auditory processing (Read, Winer, & Schreiner, 2002; Rauschecker & Tian, 2000; Langner, Sams, Heil, & Schulze, 1997; Pantev et al., 1995), such as cochleotopy, periodicity, or preferred spectral bandwidth. If the hierarchy and variety of discrete cortical fields in the human auditory cortex follows the organization found in macaques and chimpanzees, a complex interaction of such maps may enable appropriate representations of the diversity of speech sounds (Wessinger et al., 2001; Diesch, Eulitz, Hampson, & Ross, 1996). Recent studies with marmosets for species-specific vocalizations (Wang, Merzenich, Beitel, & Schreiner, 1995) and with trained gerbils for vowels (Ohl & Scheich, 1997) have implied such mappings.

In humans, magnetoencephalographic and electroencephalographic research has focused on the N100/N100m component as a possible signature of auditory pattern recognition and integration (Näätänen & Win-

kler, 1999). Especially in speech sound processing, the N100m is a valuable indicator for it is elicited reliably by vowels, syllables, and word onsets (Obleser, Elbert, Lahiri, & Eulitz, 2003; Sanders & Neville, 2003; Diesch & Luce, 2000; Poeppel et al., 1997; Diesch et al., 1996; Eulitz, Diesch, Pantev, Hampson, & Elbert, 1995). In combination with source imaging techniques, the N100m provides at least two parameters that can give further insight into auditory processing: N100m source configuration as a possible index of topographical coding and N100m peak latency as an index of temporal coding can be analyzed.

A robust finding is that the latency of the N100m is affected by acoustic correlates of speech segments such as first formant's frequency, with N100m peaking earliest for formant frequencies around 1 kHz (Roberts, Ferrari, Stufflebeam, & Poeppel, 2000). Topographical differences, as expected from animal research results, have not been clearly shown: Diesch and colleagues (1996) detected indications for separable vowel representations in the human auditory cortex, which were confirmed in a more recent MEG study (Obleser et al., 2003). Here, reliable statistical results demonstrated the preservation of acoustic dissimilarities of vowels in cortical maps. However, not all studies intended to demonstrate spatially distinct processing systems for vowels or were able to demonstrate them (Alku, Sivonen, Palomaki, & Tiitinen, 2001; Vihla, Lounasmaa, & Salmelin, 2000; Poeppel et al., 1997).

It remains unclear how the various mechanisms that have been proposed deal with poor acoustic conditions and high acoustic variance in the speech signal. An ubiquitous system of speech sound representation has not yet emerged from these results. Lahiri and Reetz (2002) proposed a linguistic model of word recognition based on the extraction of phonological features (FUL, “featurally underspecified lexicon”). Combinations of universal phonological features classify the vowels and consonants actually utilized in a given language (Kenstowicz, 1994). Extracting a minimally sufficient set of features instead of all the acoustic information available, that is, representing “underspecified” phonemes (Lahiri & Reetz, 2002), could account for variance and ambiguity in the speech signal.

Here, subjects listen to 42 tokens of seven natural German vowels while performing a target vowel detection task. We test the assumption that abstract phonological features indicating place of articulation and tongue height are utilized in the cortical processing of vowels as early as in the N100m response. We generally ask to what extent phonological features can account for the time course of activation and location of vowel sources in the auditory cortex.

We analyze peak latency as well as source location of the N100m—modeled as a equivalent current dipole (ECD) in three-dimensional source space—in response to German vowels that vary with respect to phonological features place of articulation and tongue height. The quasi-orthogonal arrangement of these features in German vowels (Figure 1) allows us to separate their influences on auditory processing.

If the phonological features also determine speech sound mapping in the auditory cortex, we expect to find relative changes in timing and topography of the N100m brain response due to changes in the features place of articulation and tongue height.

RESULTS

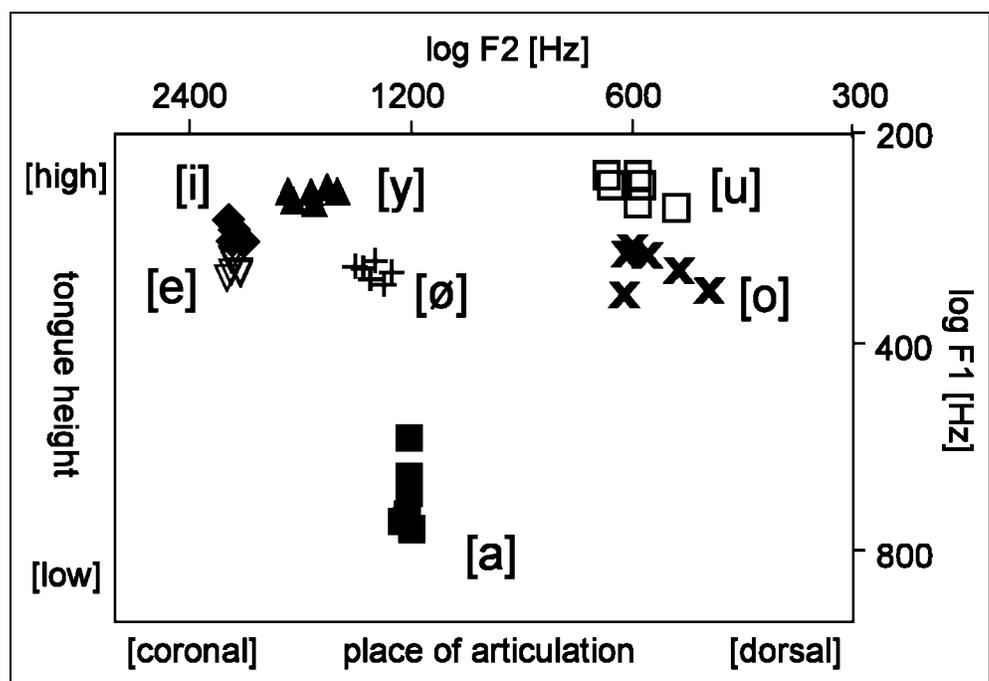
N100m RMS Peak Latency

Analyses on RMS peak latency yielded a main effect of vowel category, $F(6,114) = 25.81$, $\epsilon = .67$, $p < .0001$ (Figure 2). We therefore tested the parametric model with factors hemisphere, tongue height, and place of articulation. While dorsal vowels [o] and [u] elicited the latest N100m peaks and high vowels peaked later than non-high vowels, it also revealed a significant interaction of tongue height and place of articulation, $F(2,38) = 3.77$, $\epsilon = .87$, $p < .05$. A subsequent analysis for front vowels only showed this interaction to be driven by unrounded coronal vowels [e] and [i], $F(1,19) = 5.56$, $p < .03$: It was only for these vowels that tongue height led to a highly significant latency difference of 6 msec ($p < .0001$) (Figure 3).

N100m ECD Source Location

Vowel category did have a significant influence on ECD source location along the posterior–anterior axis, $F(6,78) = 5.41$, $\epsilon = .59$, $p < .01$. Subsequent analysis with factors hemisphere, tongue height, and place of articulation revealed a main effect of place of articulation, $F(2,26) = 9.75$, $\epsilon = .88$, $p < .01$: Irrespective of hemi-

Figure 1. A vowel space plotting the first formant frequency (*y*-axis, logarithmic display) against the second formant frequency (*x*-axis, logarithmic display) is shown for all vowel tokens used. Note the considerable acoustic variance within vowel categories as well as the close correspondence of formant frequencies to phonological features place of articulation and tongue height.



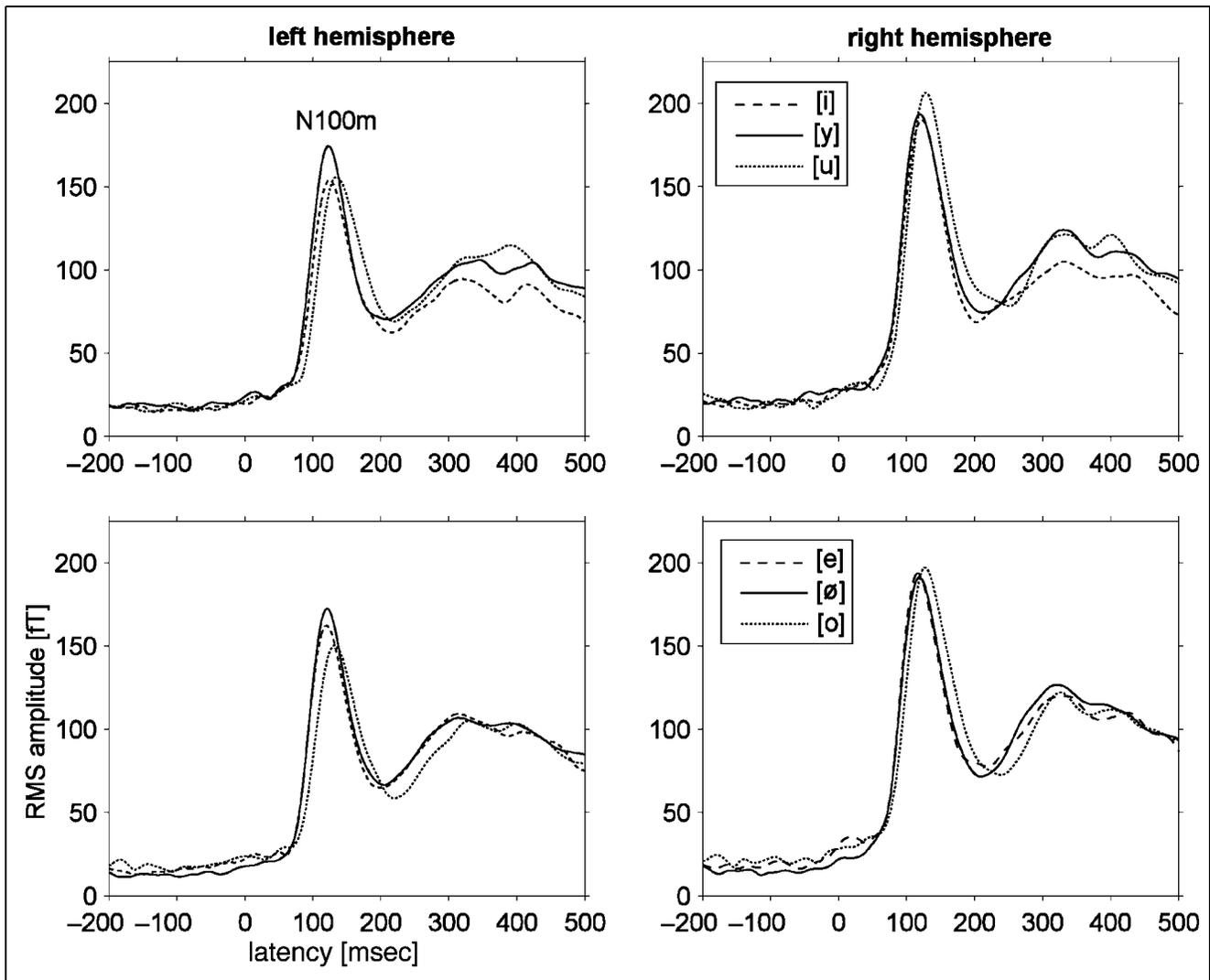


Figure 2. Grand-averaged root mean-squared amplitude of the elicited magnetic brain responses is plotted against latency, separately for high vowels (upper panels) and non-high vowels (lower panels) as well as left and right hemispheres (left and right columns, respectively). Analyses reported here are confined to the rising slope and peak of the most prominent component N100m.

sphere and tongue height, dorsal vowels were located more posterior than coronal vowels ($p < .01$) as well as coronal-labial vowels ($p < .01$) (Figure 3). The latter did not differ significantly in source location. Figure 4 illustrates the place of articulation effect by projecting mean ECD locations onto an axial slice of the Montreal Standard Brain, separately shown for high and non-high vowel sources in the left hemisphere. The more posterior position of dorsal vowel sources is evident even when the mean confidence ellipsoids are plotted, indicating 95% confidence for least squares estimates of the vowel source localizations (Kuriki & Murase, 1989). For the non-high vowel sources, which tended to be less spatially extended, this held true for [o] vs. [ø].

To distinguish between the influence of phonological features and a pure acoustic impact on the location differences along the posterior-anterior axis, the vowel [a] was considered. German [a] is not assigned to a certain place of articulation, but its ($F_2 - F_1$) difference,

which is often discussed as a relevant acoustic feature to separate vowel categories (Diesch & Luce, 2000; Ohl & Scheich, 1997), is more similar to dorsal than to coronal vowels. Hence, ECD location of vowel [a] was compared to mean dorsal, coronal, and coronal-labial places of articulation, $F(3,39) = 7.12$, $\epsilon = .86$, $p < .01$. Post hoc contrast analyses proved the [a] to be significantly more anterior than dorsal vowels ($p < .05$), but it did not differ from coronal and coronal-labial vowels.

Analyzing hemispheric differences, right-hemispheric ECDs were found to be located more anterior, $F(1,13) = 5.39$, $p < .0001$, and more inferior, $F(1,13) = 10.70$, $p < .01$, than left-hemispheric ones. The medial-lateral axis revealed no effects whatsoever. N100m source orientation in the sagittal plane corroborated the hemispheric differences found, that is, right-hemispheric ECDs were oriented more vertically, $F(1,13) = 12.39$, $p < .01$. Dipole orientation did not differ for vowel categories.

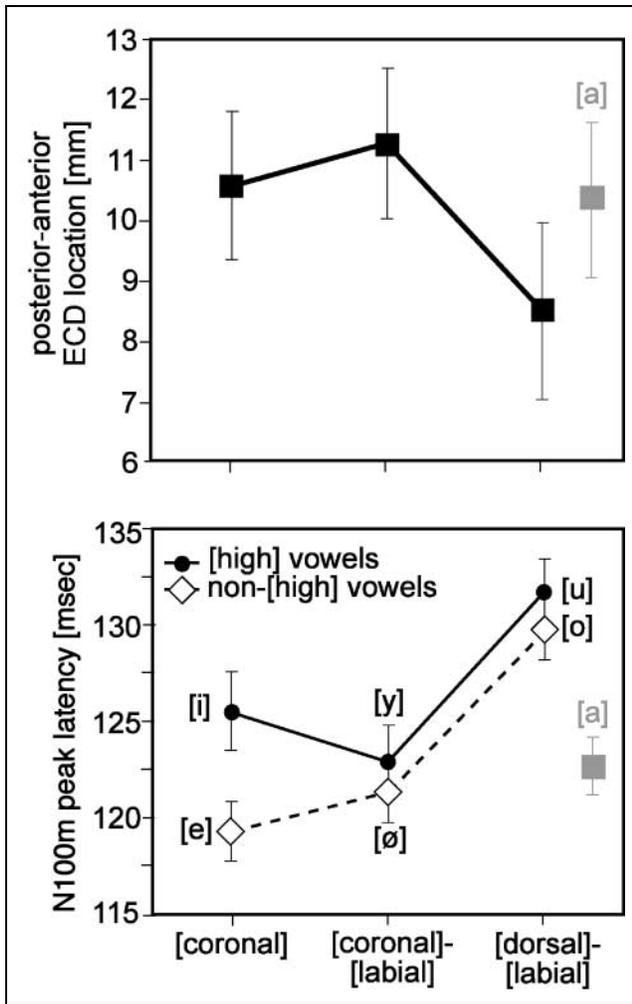


Figure 3. The upper panel displays the mean ECD locations (± 1 SEM) in the posterior–anterior direction for the three places of articulation. The lower panel shows the mean N100m peak latencies (± 1 SEM) for the same places of articulation in synopsis. High vowels are indicated by filled circles and non-high vowels by open diamonds. Note the more pronounced latency difference between [e] and [i] as compared to coronal–labial and dorsal vowels regardless to almost identical differences in the first formant frequency. The vowel [a] is shown as a reference only and was not tested in this parametric model.

When transforming mean ECD locations in Talairach coordinates (Lancaster et al., 2000), locations in both hemispheres matched the coordinates for the superior temporal or medial temporal gyrus (left-hemispheric coordinates $-57 > x > -60$, $-25 > y > -29$, $0 > z > -1$; right-hemispheric coordinates $54 > x > 51$, $-17 > y > -20$, $-1 > z > -3$). Whenever the applied automatic procedure did return more specific locations, they all allocated to Brodmann’s area 22.

Euclidean Distances between ECD Locations

We further examined intraindividual Euclidean distances between ECD sources (Obleser et al., 2003; Diesch et al.,

1996), as across-subject effects in ECD location may be blurred by interindividual variance in gross anatomical structure of the temporal lobes (Shapleske, Rossell, Woodruff, & David, 1999) as well as the more fine-grained differences in formation of experience-dependent speech sound representation (Buonomano & Merzenich, 1998).

An omnibus repeated-measures analysis of variance with 21 distances and two hemispheres revealed overall differences in distances, $F(20,260) = 3.46$, $\epsilon = .23$, $p < .01$, but no hemispheric differences. We then tested more specifically whether vowels differing in only one place feature (i.e., coronal vowels [e]/[i] vs. coronal–labial vowels [ø]/[y]) were more closely collocated than vowels differing in two features (i.e., coronal vowels [e]/[i] vs. dorsal–labial vowels [o]/[u]). A $2 \times 2 \times 2$ repeated-measures analysis of variance with factors hemisphere, tongue height, and feature difference yielded only a main effect of feature difference in the expected direction, $F(1,13) = 25.87$, $p < .001$: smaller distances between ECDs of vowels differing only in labiality amounted to 5.9 ± 1.0 mm ($M \pm SEM$), whereas vowels differing in both dorsality and labiality yielded ECD distances of 8.8 ± 1.23 mm.

Regarding features of tongue height (cf. Table 1), the mutually exclusive features low and high led to comparably large ECD distances between vowels [a] and [i] (7.0 ± 0.9 mm), while vowels [e] and [i], which are only differentiated by the absence or presence of feature high, elicited smaller distances (5.5 ± 1.0 mm), $F(1,13) = 6.75$, $p < .02$.

DISCUSSION

The aim of our study was to scrutinize the role of phonological features in the processing and identification of speech sounds. When we analyzed latency and relative spatial distribution of N100m brain responses, the phonological features turned out to be a valuable heuristic concept:

The N100m dipole solutions of all vowels were located in the auditory cortex (cf. Figure 4). The Talairach-transformed mean location coordinates suggested a cortical area outside of AI in the vicinity of the planum temporale.

At a more fine-grained spatial scale, we found an effect of the place of articulation onto the location of the center of electromagnetic brain activation. On the posterior–anterior axis, the ECD sources of vowels with mutually exclusive, never co-occurring phonological features coronal and dorsal were separable. Because dorsal and coronal places of articulation never co-occur in the vowel systems of natural languages, their spatial separation in the auditory cortex is very salient and follows predictions made from neural plasticity research (Buonomano & Merzenich, 1998) and developmental research on language acquisition (Kuhl, 2000; Jusczyk, 1999). The latter

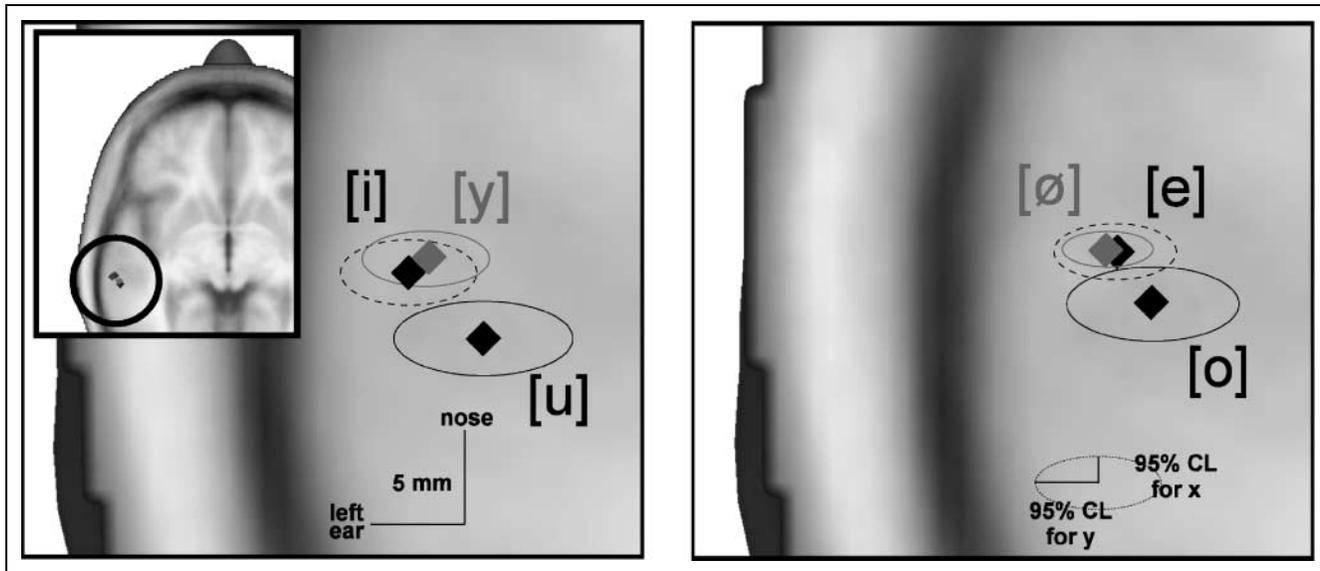


Figure 4. ECD source locations in the left hemisphere are projected onto an axial slice of the Standard Montreal brain, separately for high vowels (left column) and non-high vowels (right column). Gray-scale differences in source location symbols correspond to typed vowel symbol, respectively. Ellipsoids indicate the mean 95% confidence regions across individual source locations. Note that the spatial separation is most pronounced when just dorsal is the distinctive feature.

authors have shown that primarily unbiased auditory representations are “warped” in infancy by inherent rules and probabilities of the speech signal.

The place of articulation mapping we found might partly be congruent with animal research by Ohl and Scheich (1997) and human magnetoencephalographic research by Diesch and Luce (2000), since our place of articulation differences are mirrored in spectral (F_2-F_1) differences and the location differences in the auditory cortex are perpendicular to main tonotopic gradients (which are expected to run for the human N100m mainly from lateral to medial, cf. Pantev et al., 1995).

To further strengthen the (F_2-F_1) mapping assumption, the source location of the vowel [a] was included in the analyses. The physical parameter (F_2-F_1) of [a] is close to dorsal vowels [u] and [o] (cf. Figure 1), but

statistical analysis showed that [a] and dorsal locations differed significantly—even more so, when comparing [a] location only to the closest (F_2-F_1) neighbor vowel [u] ($p < .001$) and the most distant (F_2-F_1) vowel [i] ($p > .50$). Consequently, the ECD location results for vowel [a] rule out a linear (F_2-F_1) mapping of N100m in humans. Instead of pure acoustics, phonological features (i.e., more abstract and more pre-processed acoustic properties) may better account for the mapping along posterior–anterior axis we report here. In such a phonological feature space, distances between vowels may be warped towards representing perceptually relevant rather than linear (F_2-F_1) differences (Kuhl, 2000).

Prior studies failed to show group-statistic differences in N100m localization of vowels (Alku, Sivonen, Palomaki, & Tiitinen, 2001; Poeppel et al., 1997; Diesch et al.,

Table 1. Overview over the Assignment of Phonological Features as Well as the Pitch (F_0) and Formant Frequency (F_1 , F_2 , F_3) Variability in the Vowel Categories Used

Vowel Category	Place of Articulation	Tongue Height	F_0	F_1	F_2	F_3	(F_2-F_1)
			min–max	min–max	min–max	min–max	min–max
[a]	–	low	103–113	552–747	1188–1224	2663–3171	442–641
[i]	coronal	high	127–132	267–287	2048–2120	2838–3028	1764–1816
[e]		non-high	109–125	302–322	2055–2143	2711–2890	1741–1821
[y]	coronal–labial	high	115–144	238–248	1516–1769	1978–2097	1275–1528
[ø]		non-high	108–125	301–325	1293–1447	1945–2079	981–1142
[u]	dorsal–labial	high	112–118	231–256	522–645	2117–2292	266–415
[o]		non-high	109–125	293–346	471–609	2481–2688	131–303

The derived parameter (F_2-F_1) is also given.

1996; Eulitz et al., 1995). In the present study, a number of points may have contributed to the significant effect regardless of the comparably small location difference between dorsal and coronal vowels, such as (i) a good signal-to-noise ratio (SNR) due to more than 125 averages in every instance (average SNR 13:1); (ii) small spatial smearing of topographies during data acquisition due to measurements in supine position; and (iii) the well-controlled alertness of subjects. Furthermore, the task was chosen such that subjects had to categorize and identify vowels in order to accomplish the task which is more demanding than a simple template matching as in our previous experiments with just one target vowel. Not just the mental effort for the task, but the task itself, may also have contributed to better ECD source separation in the present results. In the somatosensory domain, for example, using MEG within-subject variations of the homuncular hand representation within the primary somatosensory cortex has been shown to be modulated by attention (Braun et al., 2002): The detection of local versus global features of a well-controlled afferent input to the somatosensory system changed the homuncular hand representation. This indicated that the hand representation within the primary somatosensory cortex is not statically fixed but is dynamically modulated by top-down mechanisms to support task requirements. Although, similar mechanisms may have modulated the representations in the auditory cortex in the present study, a controlled study has to be made to answer this question satisfactorily.

To link up the present results to those of previous studies, Euclidean distances between vowel sources have been calculated. Vowels differing in only one phonological feature were collocated more closely than vowels differing in more features—thereby confirming and extending previous results (Obleser et al., 2003). Unfortunately, the analysis of relative distances abandons absolute spatial information and therefore cannot answer the question whether linear acoustic vowel mapping or a more phonologically motivated processing is reflected in the N100m.

As in previous studies (Obleser et al., 2003; Roberts et al., 2000; Poeppel et al., 1997; Diesch et al., 1996; Eulitz et al., 1995), the latency of N100m response was influenced by vowel category. We replicated the finding that vowels [a], [i], and [u] differ in latency with [u] eliciting latest N100m responses. Roberts et al. (2000) suggested that this latency difference is a function of the first formant frequency which is correlated with tongue height. However, when comparing vowels with respect to tongue height and place of articulation, the pattern of latency differences is more sophisticated: Place of articulation and tongue height interacted on N100m latency, as coronal vowels [i] and [e] exhibited a statistically significant peak difference of about 6 msec. Within coronal–labial and dorsal vowel pairs, no such differences were found although these vowels showed

very similar F_1 frequency differences (cf. Figure 1). A comparable latency difference for [e] and [i] has been observed before (Obleser et al., 2003), but could not be interpreted in terms of phonological features, as neither rounded nor unrounded vowels were used. To our knowledge, the interaction of both the place of articulation and tongue height on the N100m latency is found for the first time. Furthermore, the first formant frequencies alone do not account for this interaction. Consequently, the F_1 frequency hypothesis of N100m latency (Roberts et al., 2000) probably has to be extended, maybe towards a more complex parameter such as the interaction of various formant frequencies along earlier auditory processing stages (Diesch & Luce, 2000; Ohl & Scheich, 1997).

In sum, we found that the latency and source configuration of the N100m component to natural vowels is sensitive to the spectral dissimilarities associated with phonological features. In addition, mutually exclusive place of articulation features dorsal and coronal activated spatially distinct regions in auditory association cortices. Further studies shall scrutinize the ubiquity of this spatial dorsal versus coronal difference of the N100m brain response. As phonological features are also assigned in consonants (Lahiri & Reetz, 2002), our results shall be further elaborated using consonant–vowel–syllables.

Place of articulation and tongue height both influenced N100m peak latency, a finding not fully compatible with an F_1 frequency hypothesis of N100m latency (Roberts et al., 2000). The present data suggest that the time course of activation as reflected in the N100m peak latency delivers additional information to distinguish between vowel categories when the spatial distinctiveness of cortical activation is low. Further studies should therefore consider the interaction of N100m latency with N100m source locations in more detail. The concept of ubiquitous phonological features turned out to be a valuable guideline in exploring speech sound processing of the human brain.

METHODS

Subjects

Twenty subjects (11 women) with a mean age of 25.5 ± 4.2 years ($M \pm SD$) took part in the experiment. None reported a history of neurological, psychiatric, or otological illness. All subjects were monolingual native speakers of German. Only right-handers were included, as ascertained by the Edinburgh Handedness Questionnaire (Oldfield, 1971). Subjects gave written informed consent and were paid € 20 for their participation. Prior to the measurement, individual hearing thresholds were determined for both ears using a calibration vowel [i].¹

Stimulus Material

We investigated brain responses to seven naturally spoken German vowels [a], [e], [i], [o], [u] (as in “father,” “bay,” “bee,” “doe,” “do,” respectively), [ø] and [y]

(both do not occur in English and are the rounded counterparts of [e] and [i]). The classification of the vowels in terms of their phonological features, as well as their pitches and formant frequencies, is given in Table 1. For every vowel category, we selected six tokens, edited from recordings of spoken words of a male speaker. The speaker was instructed to pronounce the words with extra long vowels. The speech signal was recorded with a Sennheiser MD-421 microphone on a DAT recorder (TDA-1, TASCAM). Off-line editing was performed with the CSL Speechlab 4300b (Kay Elemetrics, Lincoln Park, NJ). From the 10-kHz digitized speech signal, 350-msec portions containing only the steady-state vowel signal were cut out. All 42 stimulus audio files were ramped with 50-msec Gaussian on- and offsets and normalized for peak amplitude. Pitch frequency (119 ± 10 Hz, $M \pm SD$) and formant frequencies varied within vowel categories (cf. Figure 1, Table 1), thus introducing considerable acoustic diversity.

Experimental Design

Vowels were presented binaurally with 50 dB SL via a nonmagnetic echo-free stimulus delivery system with almost linear frequency characteristic in the critical range of 200–4000 Hz. Vowels were aligned in pseudo-randomized sequences of 572 stimuli with a variable stimulus onset asynchrony of 1.8 ± 0.2 sec. Every subject listened to three of such sequences. To sustain subjects' attention to the stimuli, a target detection task was employed: In every sequence, the realizations of two given vowels had a low cumulated probability of 10% and served as targets. Subjects had to press a button with their right index finger each time they detected such a target. For reasons of symmetry in target assignment, the central vowel [ø] never functioned as a target, but each subject encountered all six target vowel categories across the whole experiment. As all vowel categories exhibited acoustic variance (Figure 1), subjects had to map stimuli on vowel category prototypes to decide whether a given stimulus is a target or not (i.e., subjects had to maintain a phonological processing mode throughout the experiment).

In a test sequence, subjects repeated vowel stimuli aloud and recognized all stimuli as typical German vowels. Binaural loudness was slightly readjusted to ensure perception in the head midline. Subjects watched silent videos in order to maintain constant alertness and to reduce excessive eye movements.

Data Acquisition and Reduction

Auditory magnetic fields (AEFs) evoked by vowel stimuli were recorded using a whole-head neuromagnetometer (MAGNES 2500, 4D Neuroimaging, San Diego) in a magnetically shielded room (Vaccumschmelze, Hanau, Germany). Epochs of 800-msec duration (including a 200-msec pretrigger baseline) were recorded

with a bandwidth from 0.1 to 100 Hz and a 508.63-Hz sampling rate. If the peak-to-peak amplitude exceeded 3.5 pT in one of the channels or the co-registered EOG-signal was larger than 100 μ V, epochs were rejected. Epochs of target tokens as well as epochs containing false-positive button presses were also excluded. Up to 250 artifact-free epochs that remained for every subject and 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.

Further analysis closely followed our previous study (Obleser et al., 2003) and was confined to the rising slope and peak of the N100m component defined as the prominent waveform deflection in the time range between 90 and 160 msec (Figure 2). Isofield contour plots of the magnetic field distribution were visually inspected. N100m peak latency was defined as the sampling point by which the root mean square (RMS) amplitude reached its maximum. RMS was calculated across 34 magnetometer channels selected to include the field extrema. This was done separately over the left and the right hemisphere. Prior to statistical analyses, all brain response latencies were corrected for a constant sound conductance delay of 19 msec in the delivery system. Using the same sets of channels, an 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 hemispheres (Sarvas, 1987). The N100m source parameters were determined as the median of five adjacent ECD solutions in the rising slope of the N100m. The resulting ECD solution represents the center of gravity for the massed and synchronized neuronal activity. To be included in this calculation, single ECD solutions had to meet the following criteria: (i) goodness of fit greater than .90; (ii) confidence volume (containing the source with 95% certainty) smaller than 1 cm³; (iii) 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 preauricular points.

Statistical Analyses

All analyses were done with the SAS 8.0 software package. Parameters of interest were the RMS peak latency of the N100m (for $n = 20$ subjects) as well as N100m source orientation and location in posterior–anterior, medial–lateral, and inferior–superior dimension (for $n = 14$ subjects). In a repeated-measures analysis of variance, two designs were employed: Initially, a 2×7 design with repeated-measures factors hemisphere (left, right) and vowel category ([a], [e], [i], [ø], [y], [o], [u]) was carried out as an omnibus test. Given any significant effects of vowel category, a $2 \times 2 \times 3$ repeated-measures design with factors hemisphere (left, right), tongue height

(high vowels [i], [y], [u] vs. non-high vowels [e], [ø], [o]), and place of articulation (coronal vowels [i], [e] vs. coronal-labial vowels [y], [ø] vs. dorsal-labial vowels [o], [u]) followed for testing the influence of phonological features (Table 1). Features coronal and dorsal correspond to the front versus back vowel distinction, respectively, and labial corresponds to the feature roundedness (Lahiri & Reetz, 2002).

Exploring representation of vowels relative to each other in three-dimensional source space, Euclidean distances between vowel ECD sources were calculated on an intrahemispheric and intraindividual basis. For $k = 7$ vowels, there are $(k^2 - k)/2 = 21$ possible Euclidean distances. In an omnibus repeated-measures analysis of variance, all distances were submitted to a 2×21 design with factors hemisphere (left, right) and distance (21 possible distances). In case of significant differences, further comparisons of interest were explored.

As the sphericity assumption is frequently violated in psychophysiological measurements (Picton et al., 2000), Greenhouse-Geisser corrected p values are reported where applicable. For further exploration of significant differences, contrast analyses were used.

Acknowledgments

Research was supported by the Deutsche Forschungsgemeinschaft (Leibniz funds and FOR 348/3-3) and the Volkswagen foundation. We gratefully acknowledge the help of Barbara Awiszus in collecting the data. Many thanks also to Henning Reetz for advice in stimulus preparation, to Michaela Schlichting for recording and editing of the stimulus material, and to Nathaniel Pihama for correcting the manuscript.

Reprint requests should be sent to Jonas Obleser, University of Konstanz, Universitätsstrasse 10, P.O. Box D25, 78457 Konstanz, Germany, or via e-mail: jonas.obleser@uni-konstanz.de.

Note

1. In a pilot experiment, hearing thresholds were determined for all seven vowels on four healthy subjects. No consistent differences appeared. Therefore, a vowel with average hearing threshold was selected for calibration.

REFERENCES

- Alku, P., Sivonen, P., Palomaki, K., & Tiitinen, H. (2001). The periodic structure of vowel sounds is reflected in human electromagnetic brain responses. *Neuroscience Letters*, *298*, 25–28.
- Braun, C., Haug, M., Wiech, K., Birbaumer, N., Elbert, T., & Roberts, L. E. (2002). Functional organization of primary somatosensory cortex depends on the focus of attention. *Neuroimage*, *17*, 1451–1458.
- Buonomano, D. V., & Merzenich, M. M. (1998). Cortical plasticity: From synapses to maps. *Annual Review of Neuroscience*, *21*, 149–186.
- Diesch, E., Eulitz, C., Hampson, S., & Ross, B. (1996). The neurotopography of vowels as mirrored by evoked magnetic field measurements. *Brain and Language*, *53*, 143–168.
- Diesch, E., & Luce, T. (2000). Topographic and temporal indices of vowel spectral envelope extraction in the human auditory cortex. *Journal of Cognitive Neuroscience*, *12*, 878–893.
- Eulitz, C., Diesch, E., Pantev, C., Hampson, S., & Elbert, T. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *Journal of Neuroscience*, *15*, 2748–2755.
- Fitch, R. H., Miller, S., & Tallal, P. (1997). Neurobiology of speech perception. *Annual Review of Neuroscience*, *20*, 331–353.
- Jusczyk, P. W. (1999). How infants begin to extract words from speech. *Trends in Cognitive Science*, *3*, 323–328.
- Kenstowicz, M. (1994). *Phonology in generative grammar*. Cambridge: Blackwell.
- Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science, U.S.A.*, *97*, 11850–11857.
- Kuriki, S., & Murase, M. (1989). Neuromagnetic study of the auditory responses in right and left hemispheres of the human brain evoked by pure tones and speech sounds. *Experimental Brain Research*, *77*, 127–134.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven, N. Warner, & T. Rietveld (Eds.), *Labphon VII* (pp. 637–675). Berlin: Mouton.
- Lancaster, J. L., Woldorff, M. G., Parsons, L. M., Liotti, M., Freitas, C. S., Rainey, L., Kochunov, P. V., Nickerson, D., Mikiten, S. A., & Fox, P. T. (2000). Automated Talairach atlas labels for functional brain mapping. *Human Brain Mapping*, *10*, 120–131.
- Langner, G., Sams, M., Heil, P., & Schulze, H. (1997). Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: Evidence from magnetoencephalography. *Journal of Comparative Physiology, A: Sensory, Neural, and Behavioral Physiology*, *181*, 665–676.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, *125*, 826–859.
- Obleser, J., Elbert, T., Lahiri, A., & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies. *Cognitive Brain Research*, *15*, 207–213.
- Ohl, F. W., & Scheich, H. (1997). Orderly cortical representation of vowels based on formant interaction. *Proceedings of the National Academy of Science, U.S.A.*, *94*, 9440–9444.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.
- Pantev, C., Bertrand, O., Eulitz, C., Verkindt, C., Hampson, S., Schuierer, G., & Elbert, T. (1995). Specific tonotopic organizations of different areas of the human auditory cortex revealed by simultaneous magnetic and electric recordings. *Electroencephalography and Clinical Neurophysiology*, *94*, 26–40.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., Jr., Miller, G. A., Ritter, W., Ruchkin, D. S., Rugg, M. D., & Taylor, M. J. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology*, *37*, 127–152.
- Poeppel, D., Phillips, C., Yellin, E., Rowley, H. A., Roberts, T. P., & Marantz, A. (1997). Processing of vowels in supratemporal auditory cortex. *Neuroscience Letters*, *221*, 145–148.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex.

- Proceedings of the National Academy of Science, U.S.A.*, 97, 11800–11806.
- Read, H. L., Winer, J. A., & Schreiner, C. E. (2002). Functional architecture of auditory cortex. *Current Opinion in Neurobiology*, 12, 433–440.
- Roberts, T. P., Ferrari, P., Stufflebeam, S. M., & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology*, 17, 114–129.
- Sanders, L. D., & Neville, H. J. (2003). An ERP study of continuous speech processing: I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, 15, 228–240.
- Sarvas, J. (1987). Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. *Physics in Medicine and Biology*, 32, 11–22.
- Shapleske, J., Rossell, S. L., Woodruff, P. W., & David, A. S. (1999). The planum temporale: A systematic, quantitative review of its structural, functional and clinical significance. *Brain Research Reviews*, 29, 26–49.
- Sussman, H. (2000). Phonemic representation: A twenty-first century challenge. *Brain and Language*, 71, 37–240.
- Vihla, M., Lounasmaa, O. V., & Salmelin, R. (2000). Cortical processing of change detection: Dissociation between natural vowels and two-frequency complex tones. *Proceedings of the National Academy of Science, U.S.A.*, 97, 10590–10594.
- Wang, X., Merzenich, M. M., Beitel, R., & Schreiner, C. E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. *Journal of Neurophysiology*, 74, 2685–2706.
- Wessinger, C. M., VanMeter, J., Tian, B., Van Lare, J., Pekar, J., & Rauschecker, J. P. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, 13, 1–7.