ACADEMIC
PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

NeuroImage

NeuroImage 20 (2003) 1839–1847

www.elsevier.com/locate/ynimg

Auditory-evoked magnetic field codes place of articulation in timing and topography around 100 milliseconds post syllable onset

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

[a] *Department of Clinical Psychology, University of Konstanz, 78457 Konstanz, Germany*
[b] *Department of Linguistics, University of Konstanz, 78457 Konstanz, Germany*

Received 29 April 2003; revised 22 July 2003; accepted 23 July 2003

Abstract

This study demonstrates by means of magnetic source imaging how consonants and vowels that constitute a syllable differently affect the neural processing within the auditory cortex. We recently identified a topographically separate processing for mutually exclusive place features in isolated vowels (Obleser et al., in press). Does this mapping principle also hold for stop consonants with differing places of articulation? How is the N100m response to consonant–vowel (CV) syllables affected by the congruency of place information in the consonant and the vowel? Moreover, how is the N100m affected by coarticulation, i.e., the spreading of place features to adjacent phonemes? By systematically varying phonological information in the consonant as well as in the vowel of CV syllables, we were able to reveal a difference in N100m syllable source location along the anterior–posterior axis due to mutually exclusive places of articulation in the vowel of the syllable. We also found a change in source orientation rather than source location due to the same mutually exclusive features in the onset of the syllable. Furthermore, the N100m time course of the brain response delivered important complementary information to identify the phonological features present in the speech signal. Responses to all syllable categories originated in the perisylvian region anterior to the source of a band-passed noise stimulus. The systematic variation of both consonantal and vocalic place features and the study of their interaction on auditory processing proves to be a valuable method to gain more insight into the elusive phenomenon of human speech recognition.

Introduction

In human speech, single phonemes are articulated in a cascade rather than in serial order and overlap considerably. This overlapping articulation or coarticulation may be anticipatory (affected by upcoming sounds) as well as persevering (continued effect of preceding sound; Farnetani, 1997). As coarticulation is the standard rather than the exception in natural speech, it is a phenomenon worth considering more extensively in speech perception research.

This study is designed to examine the influence of coarticulation on consonant and vowel processing. Most brain imaging studies have been using tokens of /ba/, /da/, and /ga/ to examine stop consonant processing (e.g., Steinschneider et al., 1995; Gage et al., 2002). However, the influence of the coarticulated vowel cannot even be quantified in such a design because the vowel /a/ is held constant. Since it is well known how formant transitions depend on coarticulation (Fitch et al., 1997), this study employs two vowels with two different places of articulation additionally to different stop consonants to scrutinize the role of coarticulation in temporal and spatial mapping of speech sounds in the human auditory cortex.

Acoustically, a vowel exhibits a relatively stable frequency spectrum, whereas a consonant is characterized by a certain onset characteristic (e.g., a noise burst in stop consonants) and a formant transition, i.e., a sweep-like change in time–frequency pattern (Fig. 1). Phonetically, vowels constitute articulatory target states, and formant transitions bear information about these target states of the vocalic tract (Eggermont, 2001). They allow the recipient to anticipate upcoming phonemes, as formant transitions vary consider-

---

* Corresponding author. University of Konstanz, Universitätstrasse 10, P.O. Box D25, 78457 Konstanz, Germany. Fax: +49-0-7531-884601.
E-mail address: jonas.obleser@uni-konstanz.de (J. Obleser).

doi:10.1016/j.neuroimage.2003.07.019

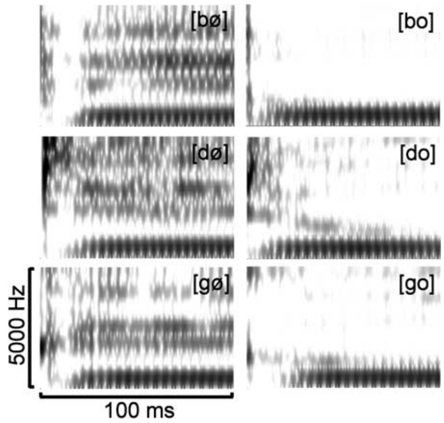

Fig. 1. Spectrograms (first 100 ms) of typical stimuli from all six syllable categories are presented. Note the obvious differences in formant frequencies between [ø]-containing (left column) and [o]-containing CV syllables (right column), as well as the differences between [d] (middle panel) and [g] (bottom panel) in shape of second formant transition.

ably depending on the quality of the ensuing vowels (Fitch et al., 1997; Sussman et al., 1998). Thus, in the speech time course, a consonant–vowel (CV) syllable can be recognized and used for further processing such as lexical access before the entire syllable is completely perceived.

Despite the known acoustic outcome of coarticulation, the crucial steps in speech sound decoding are still an unsolved problem (Scott and Johnsrude, 2003). Since the human brain recognizes a phoneme correctly even when it is produced by different speakers and in varying acoustic conditions, some kind of abstract pattern must be acquired and stored (Sussman et al., 1998; Kuhl, 2000; Lahiri and Reetz, 2002; Stevens, 2002; for analogue implications in animal vocalizations cf. Wang, 2000).

Several linguistic approaches assume that this abstract template is a set of phonological features, and that speech sounds are stored as bundles of features (cf. Chomsky and Halle, 1968; Lahiri and Reetz, 2002; Stevens, 2002). Under this assumption, the auditory system might scan incoming speech for discriminative features rather than for exact phoneme identity, as the feature information may be more robust and sufficient for accurate lexical access (Lahiri and Reetz, 2002). In previous electro- and magnetoencephalographic studies, the heuristic value of such an abstract feature representation has been probed for the perception of vowels in preattentive processing (Eulitz and Lahiri, 2004; Vihla and Eulitz, 2003) as well as in an active phonological processing mode (Obleser et al., 2003, 2004). In both experimental approaches, changes in place of articulation yielded differences in topography and time course of brain responses. German front and back vowels (hereafter referred to as "coronal" and "dorsal" vowels) activated distinct areas in the auditory cortex, irrespective of other features such as tongue height or lip rounding (Obleser et al., 2004). The fact that these differences covaried with more general feature categories across phoneme boundaries (i.e., the location difference was found for different dorsal vowels, such as [u] and [o], compared to all coronal vowels, such as [i], [y], [e], and [ø]) suggests that the processing of speech around 100 ms poststimulus onset is influenced by abstract phonological features rather than by phoneme category.

In the present study, our interest is focused on the mechanisms of place of articulation mapping across vowels and consonants in CV syllables. What drives the auditory cortical processing in the time range of the N100m component when coarticulated CV syllables with congruent or differing places of articulation are presented?

In humans, magnetoencephalographic (MEG) and electroencephalographic research has focused on the N100/N100m component as a possible signature of auditory pattern recognition and integration (Naatanen and Winkler, 1999). The N100m is a valuable indicator for decoding processes, especially in speech sound processing, for it is elicited reliably by vowels, syllables, and word onsets (Eulitz et al., 1995; Poeppel et al., 1997; Diesch and Luce, 2000; Obleser et al., 2003; Sanders and Neville, 2003). 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. Temporally high-resolution sampling is an advantage of MEG over brain imaging techniques such as PET and fMRI, as these techniques summate over longer periods of time and therefore over different stages of acoustic, phonological, and task-relevant processing.

Gage and colleagues were able to demonstrate that the N100m response also integrates information over time and has a small integration time window of at least up to 40 ms (Gage et al., 1998; Gage and Roberts, 2000). Consequently, one can expect that differential formant transitions within the first 40 ms of a syllable may have a profound impact on the N100m responses and varying places of articulation will differentially modulate the N100m. This implies also that not only differences due to different stop consonants, e.g., [gø] vs [dø], might be observed but also differences between syllables that share the same stop consonant but not the consecutive vowel, e.g., [gø] vs [go].

Here, we analyze peak latency, generator location, and orientation of the auditory N100m response to syllables that vary orthogonally in stop consonant's and vowel's place of articulation in order to examine the role of phonological place features in speech recognition.

Materials and methods

Subjects

Twenty-two subjects (11 female) with a mean age of 24.8 ± 3.3 years (M ± SD) took part in the experiment.

Table 1
Phonological features conventionally assigned to the vowels and stop consonants used, whereby Coronal and Dorsal are more general but synonymous labels for the vowel-specific place features Front and Back, and the consonantal place features Alveolar and Velar, respectively

	Vowel place of articulation	
	Coronal	Dorsal
Stop place of articulation		
Labial	[bø]	[bo]
Coronal	[dø]	[do]
Dorsal	[gø]	[go]

None reported a history of neurological, psychiatric, or otological illness. All subjects were monolingual native speakers of German. Only right-handers as ascertained by the Edinburgh Handedness Questionnaire (Oldfield, 1971) were included [(R − L)/(R + L) = 0.9]. Subjects gave written informed consent and were paid €20 for their participation.

Stimulus material

Naturally spoken German syllables were used to investigate place of articulation in stop consonants and vowels. [b], [d], and [g] (labial, coronal, dorsal place of articulation, respectively) were combined either with a coronal rounded vowel [ø] or a dorsal rounded vowel [o], resulting in six consonant–vowel syllables [bø], [dø], [gø], [bo], [do], and [go] (Table 1). The features coronal and dorsal refer respectively to front and back vowels as well as consonants with alveolar and velar places of articulation.

Syllables were edited from recordings of spoken words of a female speaker. The speaker was instructed to pronounce the words with extra long vowels. Speech signals were 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 20-kHz digitized speech signal, 350-ms portions containing the plosive, the formant transition, and the steady-state vowel signal were cut out. All 36 stimulus audio files were faded out with 50-ms Gaussian ramps and normalized for peak amplitude. Natural syllables were preferred over synthetic stimuli, since the naturally occurring variance in exact pitch, formant frequency, and time course of speech sounds extends the external validity of our measurements: Pitch frequency (279 ± 27 Hz, M ± SD) and formant frequencies varied since we used six tokens of each syllable category. F_1 in the steady-state vowel phase ranged for both [o] and [ø] from 450 to 550 Hz, F_2 ranged for [o] from 800 to 1050 Hz and for [ø] from 1250 to 1550 Hz. Voice onset time (VOT) was influenced by the stop consonant (Kruskal–Wallis $P > 0.01$) and longest for [g] syllables whereas the vowel did not affect VOT consistently (Mann–Whitney $P > 0.30$). No prevoicing (negative VOT) occurred.

Experimental design

Prior to the measurement, individual hearing thresholds were determined for both ears using 35-ms onset fragments of all syllable categories. The same onset fragments were then used in a categorization task where subjects had to determine syllable identity from onset fragments in 60 trials. This behavioral task was employed to ensure that subjects were able to identify syllables from a stimulus time window which is reflected in the N100m brain response (Gage and Roberts, 2000).

In this task as well as in the actual MEG measurements, stimuli 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.

For MEG measurements, syllables were aligned in pseudorandomized sequences of 572 stimuli with a variable stimulus onset asynchrony of 1.8 ± 0.2 s. Every subject listened to three of such sequences. To sustain attention to the stimuli, subjects performed target detection tasks (Obleser et al., 2004): In every sequence, the tokens of two given syllables had a low cumulated probability of 10% and served as targets, e.g., [bo] and [dø]. Subjects had to press a button with their right index finger each time they detected such a target. Each subject encountered all six syllable categories as targets across the whole experiment. Since all syllable categories exhibited acoustic diversity (cf. Zielinski and Rauschecker, 2000; Obleser et al., 2004), subjects had to map stimuli on syllable 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 syllables aloud and recognized all stimuli correctly. Binaural loudness was slightly readjusted where necessary to ensure perception in the head midline. Subjects watched silent videos (Obleser et al., 2003, 2004) in order to maintain constant alertness and to reduce excessive eye movements.

Data acquisition and reduction

Auditory magnetic fields (AEFs) evoked by syllable stimuli were recorded using a whole head neuromagnetometer (MAGNES 2500, 4D Neuroimaging, San Diego, CA) in a magnetically shielded room (Vaccumschmelze, Hanau, Germany). Epochs of 800 ms duration (including a 200-ms pretrigger baseline) were recorded with a bandwidth from 0.1 to 100 Hz and a 687.17-Hz sampling rate. If the peak-to-peak amplitude exceeded 3.5 pT in one of the channels or the coregistered EOG signal was larger than 100 μV, epochs were rejected. Epochs of target tokens as well as epochs containing false-positive button presses were also

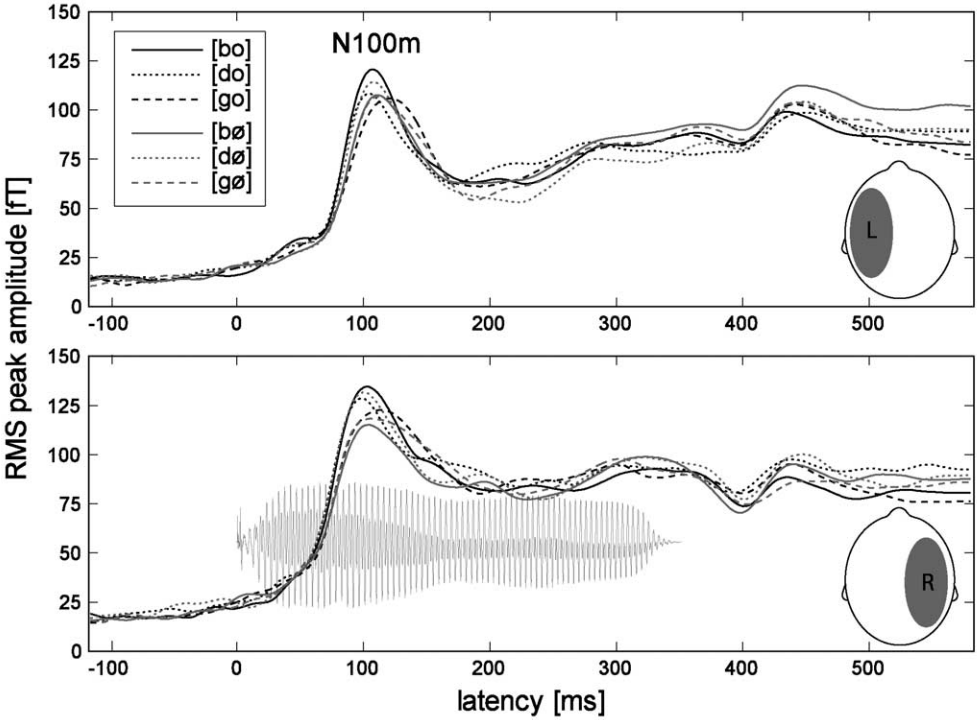

Fig. 2. Grand average waveforms from $N = 22$ subjects are shown for all syllable categories separately for left and right hemispheres (upper and lower panel). Analyses reported here were confined to the most prominent waveform deflection N100m.

excluded. Up to 250 artifact-free epochs that remained for every subject and vowel were averaged after off-line noise correction. A 20-Hz lowpass filter (Butterworth 12 dB/oct, zero phase shift) was subsequently applied to the average.

The N100m component was evident in all subjects and all conditions and was defined as the prominent waveform deflection in the time range between 90 and 160 ms (Fig. 2). Isofield contour plots of the magnetic field distribution were visually inspected to ensure that N100m and not P50m or P200m were analyzed.

N100m peak latency was defined as the sampling point in this latency range by which the first derivative of the root mean square (RMS) amplitude reached its minimum and second derivative was smaller than zero. RMS amplitude was calculated across 34 magnetometer channels selected to include the field extrema over the left and the right hemisphere, respectively.

Prior to statistical analyses, all brain response latencies were corrected for a constant sound conductance delay of 19 ms in the delivery system. Using the same sets of channels, 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 (Sarvas, 1987). The N100m source parameters were determined as the median of the three best successive ECD solutions from the rising slope of the N100m. Great care was taken not to select solutions after the peak, as underlying source configuration may change after the N100m peak (Scherg et al., 1990). The resulting ECD solution represents the center of gravity for the massed and synchronized neuronal N100m activity. To be included in this calculation, single ECD solutions had to meet the following plausibility criteria: (i) goodness of fit greater than 0.90, (ii) 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, and (iii) ECD orientation pointing inferior and posterior.

Additional nonspeech condition

The recording session was followed by an additional passive listening session with 150 trials of a band-passed noise stimulus serving as a nonspeech functional landmark of belt area activation in primary auditory cortex (Kaas et al., 1999; Wessinger et al., 2001) A 350-ms white noise probe (5-ms onset, 50-ms Gaussian offset ramps) was band-pass-filtered (center frequency 1 kHz, width 0.3 oct) and was presented binaurally at 50 dB SL with a randomized SOA of 1.8 ± 0.2 s. In further data processing, noise condition trials were treated like syllable condition trials.

Statistical analyses

All analyses were done with the SAS 8.0 software package. Parameters analyzed were the RMS peak latency and amplitude of the N100m (for $N = 22$ subjects) as well as N100m source strength, orientation, and location in poste-

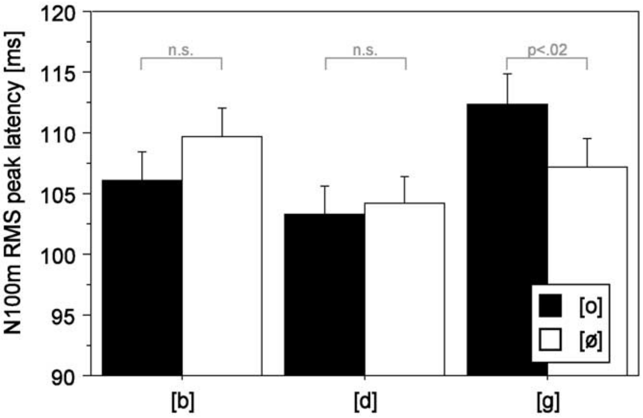

Fig. 3. Differences in N100m peak latency depending on place of articulation combinations in syllables show that the combination of a dorsal [g] with a dorsal [o] leads to a significant prolongation of the N100m peak latency.

rior–anterior, medial–lateral, and inferior–superior dimension (for $N = 16$ subjects). In a repeated-measures analysis of variance (ANOVA), a 2 × 2 × 3 design with the repeated-measures factors hemisphere (left, right), stop consonant place of articulation (labial [b], coronal [d], dorsal [g]), and vowel place of articulation (coronal [ø], dorsal [o]) was tested in order to quantify the influence of phonological features.

As the sphericity assumption is frequently violated in psychophysiological measurements (Picton et al., 2000), sphericity tests using Mauchly's criterion were performed. Unless otherwise stated, univariate tests proved to be appropriate and are reported here. However, when violation of sphericity occurred, more appropriate multivariate tests using Wilks λ-approximated F values were performed.

For post hoc comparisons of significant effects, contrast analyses were used, with significance levels adjusted by Holm's ranked Bonferroni procedure (Holland and Copenhaver, 1988).

Results

Behavioral data

Every subject accomplished the categorization task of 35-ms syllable onsets without difficulty and well above chance level. Across all 60 test items, subjects responded in 76 ± 3.6% (M ± SEM) of trials correctly, while chance level was down to 16.67% with six response options. Since only 35 ms of syllable onset were presented, vowel identification from only the formant transition, i.e., from coarticulatory cues, was important. Errors in vowel identification were not equally distributed across syllable categories, as revealed by a 2 × 3 repeated-measures ANOVA with factors vowel and stop consonant (Wilks λ-approximated $F(2,19) = 32.05$, $P < .0001$): In the category [do], significantly more perception errors than in any other category occurred; that is, initial [d] consonant strongly misled subjects to perceive [dø] instead of [do] (4.7 ± 0.5 errors, M ± SEM, all post hoc tests $P < 0.0001$). As expected, very few errors occurred in stop consonant identification (0.9 ± 0.2 errors, M ± SEM). The behavioral task during MEG measurements, i.e., target detection, was accomplished well by all subjects. Despite the ambiguity of the target syllables, 89.7 ± 1.6% of the button presses were correct.

MEG data

In all 22 subjects, stimulation elicited bihemispherical N100m responses (Fig. 2) showing the typical dipolar field patterns, although most of the subjects showed either slight left- or right-hemispheric preponderance. Influence of subjects' gender (Obleser et al., 2001) was analyzed separately and revealed that female subjects showed a more pronounced left-hemispheric N100m asymmetry than male subjects (Obleser et al., submitted). However, gender did not affect any syllable-specific effects reported here whatsoever.

For 16 subjects (eight female), data quality allowed the selection of adequate and physiologically plausible single dipole models to account for N100m activity in both hemispheres (mean N100m SNR 1:5.9). An average goodness of fit of 0.97 ± 0.16 (M ± SEM) was achieved, and average

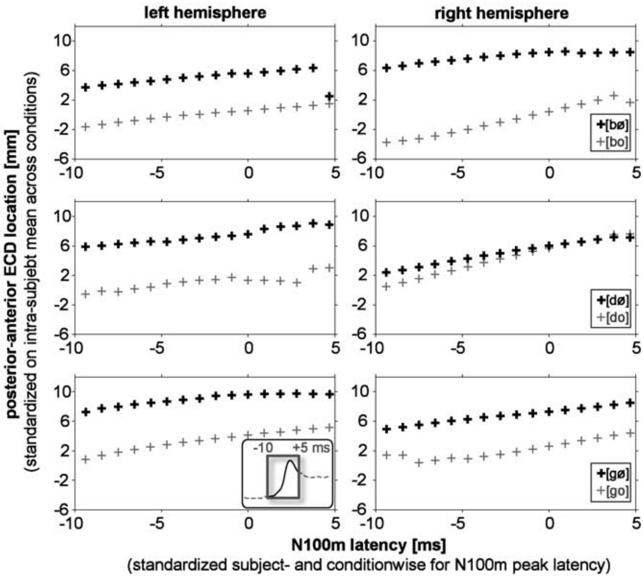

Fig. 4. Changes of source locations along the anterior–posterior dimension (y axis) are shown over 10 ms before and 5 ms after the N100m peak (x axis, cf. scheme in left bottom panel). Left- and right-hemispheric responses are shown separately in the left and the right column, and separate rows show sources of separate syllable onsets ([b], [d], [g] from top to bottom row). Source locations as well as latency have been standardized subject- and conditionwise, such as that $x = 0$ indicates individual and conditionwise N100m peaks and shifts on the y axis indicate intraindividual condition differences in posterior–anterior source locations. Note that [o]-containing syllables are located more posterior than [ø]-containing syllables, irrespective of syllable's onset consonant.

confidence volume indicating 95% certainty of estimated dipole location amounted to 151 ± 106 mm^3 (M ± SEM).

N100m peak latency

As can be observed in Figs. 2 and 3, N100m peak latency was affected by an interaction of vowel place of articulation and stop consonant place of articulation ($F(2,42) = 5.78$, $P < 0.006$). Especially the interaction between coronal/dorsal consonant and coronal/dorsal vowel place of articulation was significant (Fig. 3), $F(1,21) = 4.91$, $P < 0.05$: It was only for the fully dorsal syllable [go] that a prolonged N100m response latency appeared. For the coronal [d], prolongation by vowel change from [ø] to [o] was not evident ($F < 1$), and vowels following labial [b] also had no significant influence ($F(1,21) = 3.2$, $P < 0.10$).

Additional to this syllable-induced N100m latency effect, there was a strong tendency toward a faster-responding right hemisphere ($\Delta_t = 7$ ms, $F(1,21) = 3.78$, $P < 0.07$).

N100m source location and orientation

Along the posterior–anterior dimension, source generators of all syllables containing the dorsal vowel [o] were located significantly more posterior than generators of syllables containing the coronal [ø] ($F(1,15) = 10.97$, $P < 0.005$). This effect was of very comparable magnitude in both hemispheres. In [b] and [g] syllable onsets, vowel change from [o] to [ø] shifted the center of gravity on average 2.4 and 2.3 mm anterior. In coronal [d] syllable onsets, the anteriority shift was not as pronounced, although the interaction was not significant.

The spatial separation became more evident when the intraindividual location changes along the posterior–anterior axis across the time course of the N100m were observed (Fig. 4). It is remarkable how location differences of several millimeters were stable across the rising slope of the N100m ($t = 0$ indicates individual and conditionwise N100m peaks), and how they appeared to be driven by vowel's place of articulation rather than stop consonant's place of articulation: When comparing syllables with coronal and dorsal stop consonants irrespective of the vowel, no such location difference was observed ($F < 1$).

While the exact location of a dipolar source gives us an impression about the center of gravity of the active cortical patch, its orientation allows us to draw conclusions about the involvement of differentially oriented cortical patches, such as different parts of Heschl's gyrus or the planum temporale.

Interestingly, stop consonant place of articulation influenced the orientation of the source generator in the sagittal plane relative to the Sylvian fissure (Wilks λ-approximated $F(2, 14) = 6.49$, $P < 0.01$, Fig. 5): dorsal syllable onsets [gø] and [go] elicited cortical responses that were by 6° more vertically oriented than coronal onsets [dø] and [do] ($F(1,15) = 9.96$, $P < 0.01$), which were oriented more horizontally pointing posterior. Labial onsets [bø] and [bo] elicited activity that was not oriented differentially either from coronal or from dorsal syllable onsets.

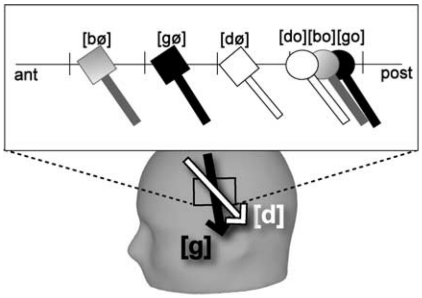

Fig. 5. Empirical grand mean data (upper panel) show the orientation difference between coronal (white ECD symbols) and dorsal (black ECD symbols) syllable onsets combined with the posterior–anterior location difference between coronal and dorsal vowel syllables (i.e., square symbols are more anterior than circle symbols in all colors). The source orientation difference we found for coronal (white arrow) vs dorsal (black) syllable onsets is also shown schematically in the lower panel. Labial onsets (gray symbols in upper panel) did not differ in orientation.

N100m peak amplitude and ECD source strength

In sensor space RMS amplitude, an interaction of vowel and stop consonant place of articulation occurred ($F(2,42) = 8.72$, $P < 0.0001$). However, when analyzing amplitudes in source space where exact generator distance from the sensors and exact head position are taken into account, all syllable-dependent amplitude differences vanished (vowel × stop consonant interaction $F < 1$).

Comparison with nonspeech condition

Where applicable, a 2 × 2 repeated-measures ANOVA with factors hemisphere (left, right) and condition (noise condition vs the mean of all syllable conditions) was tested for all dependent variables reported above to quantify differences between syllable and nonspeech noise processing. N100m peak latency and amplitude could be tested in 21 of 22 subjects; one subject was excluded due to data loss. However, magnetic source imaging is highly dependent on the dipolarity of the evoked magnetic field and demands an excellent signal-to-noise ratio. In the noise condition, a poorer signal-to-noise ratio (smaller N100m amplitudes due to strong habituation in the blocked presentation) and stronger lateralized activity (i.e., many subjects showed strongly dipolar fields patterns only over the right hemisphere) allowed satisfactory source space analysis for both hemispheres in only 13 subjects.

No differential N100m peak latency was apparent for the noise condition ($F < 1$); however, N100m peak amplitude was significantly smaller for this condition in signal space (N100m RMS peak amplitude, $F(1,20) = 55.6$, $P < 0.0001$) as well as in source space (ECD source strength, $F(1,12) =$

6.59, $P < 0.03$). The ECD sources fitted in the noise condition were located in an adjacent, but more posterior, cortical patch than the syllable ECDs ($\Delta_{\text{posterior-anterior}} = 6.4$ mm, $F(1,12) = 9.05$, $P < 0.01$). Correspondingly, the noise condition ECDs were located 3.6 mm more superior and were tilted by 10° more perpendicular to the Sylvian fissure, as one would infer from the more posterior position in the perisylvian region. But both effects failed to attain significance, presumably because of considerable variance in noise condition ECD modeling and reduced statistical power.

Discussion

To the best of our knowledge, this is the first brain imaging study that scrutinized the processing of CV syllables by systematically varying the place of articulation in both the consonant and the vowel. The sets of CV syllables allowed minute examination of the auditory processing of coarticulated place features.

Based on previous studies with isolated vowels (Obleser et al., 2004) and nonwords (Vihla and Eulitz, 2003), we expected topographic and temporal differences in the N100m component due to place of articulation changes within CV syllables (e.g., [gø] or [do]) as well as between CV syllables (e.g., [go]–[do] or [go]–[gø]). If mutually exclusive places coronal and dorsal are coarticulated, e.g., in the syllable [gø], auditory N100m topography may be determined by the place information of the stop consonant or that of the vowel or by an interaction of both.

Surprisingly, the present results suggest that the spatial mapping of CV syllables is largely dominated by place information from the vowels (cf. Fig. 4): we found significantly different locations of N100m sources along the posterior–anterior axis which were dependent on the place of articulation of the vowel but independent of different syllable onsets. In all syllables, the coronal vowel [ø] elicited activity with a more anterior center of gravity than the dorsal vowel [o], which becomes particularly obvious when intraindividual differences between source generators of these place features are considered. The effect matches in magnitude and directionality the results of Obleser et al. (2004) where the same place features were investigated in a set of German vowels spoken by a male.

As it is well known that formant transitions in coarticulated CV syllables bear information about the adjacent vowel (Fitch et al., 1997; Sussman et al., 1998; Eggermont, 2001), this finding is not surprising. Phonetic studies of articulatory processes and their resulting acoustic spectra indicated that stop consonants are prone to the assimilatory influence of the succeeding vowel's place of articulation (Keating and Lahiri, 1993; Sussman et al., 1998), which is what we saw in the present main effect of the place of articulation of the vowel on N100m source topography.

Our data, in combination with preceding studies of isolated vowels, suggest that a mapping of place features is implemented in the human perisylvian cortex, anterior to primary auditory areas (as shown by the source location of band-passed noise known to emerge from core and belt areas in primary auditory cortex, cf. Kaas et al., 1999; Wessinger et al., 2001) and orthogonal to the main tonotopic gradient (Langner et al., 1997; Ohl and Scheich, 1997; Diesch and Luce, 2000; Read et al., 2002; Obleser et al., 2004).

The N100m topography shifts due to different places of articulation suggest a basic mechanism in human speech recognition which analyzes the speech signal for invariant spectro-temporal cues, that is, phonological features which in turn allow immediate access to further processing stages involving the mental lexicon (Lahiri et al., 1984; Lahiri and Reetz, 2002).

Place of articulation, particularly coronal and dorsal, are phonological features that have their robust spectral counterparts in second formant frequency or the difference between second and first formant frequency (Peterson and Barney, 1952), which may explain why we succeeded in tracking topographic processing differences with magnetic source imaging, a macroscopic, noninvasive technique with limited spatial resolution. Another important reason might be that we utilized methods that can enhance the signal-to-noise ratio; e.g., subjects are tested in resting supine position and under constant vigilance, their focus of attention is on the stimulus material as ensured by a task rather than passive listening, and the task is phonetic in nature and forces subjects into the desired processing mode.

Stop consonants affected the N100m sources in a different manner: the orientation of the ECD source in the posterior–anterior plane was by 6° more horizontal for coronal [d] syllable onsets than for dorsal [g] onsets. The rotation of an equivalent current dipole can indicate that the configuration of the underlying brain activity has changed without a displacement of the centroid of activity, e.g., that the activity now incorporates additional rather than entirely different areas in the auditory cortex, or that the underlying activity is slightly shifted along the curvature of the auditory cortical areas (Pantev et al., 1995). Interestingly enough, it was again the difference between mutually exclusive place features, coronal vs dorsal, that showed this significant change in source configuration. The lip rounding, i.e., the feature labial, which in contrast is likely to cooccur with a coronal as well as a dorsal place of articulation, did not elicit temporally or spatially separable brain responses.

The signatures of place features on the N100m peak latency promote a functional role for temporal coding in human speech recognition as well (Roberts et al., 2000; Gage et al., 2002; Obleser et al., 2004): While previous studies demonstrated a prolonged N100m latency for dorsal vowels, i.e., [u] and [o] (Poeppel et al., 1997; Roberts et al., 2000; Obleser et al., 2004), the combination of dorsal stop and dorsal vowel features also seemed to delay N100m peak latency, and more so than could be expected from dorsal vowel place or longer VOTs in all [g] syllables alone (Fig. 3). This suggests that the assimilatory effect of a dorsal vowel is very influential on a dorsal consonant like [g]. The low formant frequencies resulting from the presence of the place feature dorsal in both the consonant and the vowel as in [go] may prolong the temporal integration process reflected in the N100m (Gage and Roberts, 2000) substantially.

Peak latency of the N100m reflects stimulus attributes such as spectral composition and onset characteristic—attributes that themselves serve in speech recognition and may allow speech sound discrimination (Roberts et al., 2000). For German vowels, we found repeatedly highly significant latency differences in response to closely related vowels [i] and [e] where topographical separation was not evident (Obleser et al., 2003, 2004). The utilization of temporal information in addition to topographic changes in firing patterns and the interaction of both mechanisms might be a valuable tool for hierarchically higher processing stages such as lexical access that deserves further exploration.

Our findings might be unprecedented in magnetoencephalographic or brain imaging studies of consonant–vowel syllables. However, previous neuroscientific research on speech did not allow the testing of such effects. Studies tended to consider either vowels in isolation (Eulitz et al., 1995; Poeppel et al., 1997; Roberts et al., 2000; Obleser et al., 2003) or CV syllables (Poeppel et al., 1996; Jancke et al., 1999; Zielinski and Rauschecker, 2000), or both (Jancke et al., 2002), but there was no careful examination of the information-bearing elements (Fitch et al., 1997; Eggermont, 2001) provided by coarticulation, since none of these studies systematically varied place information in the vowels they used. Brain imaging studies using metabolic or blood flow measures (PET, fMRI) also reflect brain activity that is summated over different stages of acoustic and phonological processing. This can account for the more distributed activity (Jancke et al., 2002; Scott and Johnsrude, 2003) found by these studies, compared to the focal activity seen in the present study. One should also keep in mind that an equivalent current dipolar source reflects the center of gravity of the brain activity at a given time point, not the exact spatial extent of activation. Our results are nevertheless consonant with brain imaging studies of auditory processing, as we also found the main focus of speech-evoked activity to originate anterior of the primary auditory cortex (cf. Scott and Johnsrude, 2003): All syllables were processed anterior of nonspeech noise, which is known to activate primary auditory core and belt areas (Kaas et al., 1999; Wessinger et al., 2001).

The exceptions to the rule that stop consonants in general are prone to the assimilatory influence of the succeeding vowel's place of articulation (Keating and Lahiri, 1993; Sussman et al., 1998) are the coronal stop consonants which are less influenced by the following vowels. Instead, coronal consonants have a considerable effect on the following vowels, such that dorsal vowels become more coronal if preceded by a coronal consonant (Stevens and House, 1963). This articulatory (and consequently acoustic) asymmetry can account for two otherwise opaque phenomena in our present results. First, the place of articulation difference found along the anterior–posterior axis was not as clear-cut for the coronal syllable onset as it was for labial and dorsal onsets (cf. Fig. 4). This somewhat diminished N100m topography difference between [dø] and [do] could be predicted from the result of Stevens and House—the modifying context of a coronal stop consonant did not allow a dorsal vowel to be produced with a clear dorsal-like low second formant (Fig. 1).

Second, subjects frequently mistook syllables with [o] for their [ø]-containing counterparts when preceded by the coronal stop, when they had to identify syllables from 35-ms onset fragments. This was a highly significant effect and could not otherwise be explained than by the fact that the coronal consonant influenced the vowel rather than the other way around (cf. the spectrogram of [do] compared to [bo] and [go] in Fig. 1).

Conclusion

The present data demonstrate that for a comparably early processing stage (the earliest that reflects access to important coarticulatory information), speech sound mapping in the auditory cortex is sensitive to phonological features such as place of articulation and the coarticulatory processes in which these features are embedded. (i) Temporal coding repeatedly appeared as a parameter that can deliver additional and complementary insight into the dynamics of speech processing. (ii) The topographic mapping of place features that we identified in isolated vowels emerged again in coarticulated syllables and thereby confirmed what has been suggested by phonetic as well as psycholinguistic research: the acoustic correlates of specific phonemes are subject to dramatic change in different coarticulatory contexts, and the concomitant perceptual invariances may be better understood when turning to more abstract underlying phonological features. The systematic variation of both consonantal and vocalic place features and the study of their interaction on auditory processing proved to be a valuable method to gain more insight into the elusive phenomenon of human speech recognition.

Acknowledgments

Research was supported by the Deutsche Forschungsgemeinschaft (Leibniz Funds, FOR 348) and the Volkswagen Foundation. Barbara Awiszus and Sonja Schumacher helped collect data. Special thanks go to Melanie Hochstätter and Henning Reetz, who helped record stimulus material. Henning Reetz also contributed to the design of the study. Thomas

Elbert, Patrick Berg, and two anonymous reviewers helped improve data presentation and interpretation.

References

Chomsky, N., Halle, M., 1968. The Sound Pattern of English. Harper and Row, New York.

Diesch, E., Luce, T., 2000. Topographic and temporal indices of vowel spectral envelope extraction in the human auditory cortex. J. Cogn. Neurosci. 12, 878–893.

Eggermont, J.J., 2001. Between sound and perception: reviewing the search for a neural code. Hear. Res. 157, 1–42.

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. J. Neurosci. 15, 2748–2755.

Eulitz, C., Lahiri, A., 2004. Neurobiological evidence for abstract phonological representations in the mental lexicon during speech recognition. J. Cogn. Neurosci., in press.

Farnetani, E., 1997. Coarticulation and connected speech processes, in: Hardcastle, W.J. (Ed.), The Handbook of Phonetic Sciences, Blackwell, Oxford, pp. 371–404.

Fitch, R.H., Miller, S., Tallal, P., 1997. Neurobiology of speech perception. Annu. Rev. Neurosci. 20, 331–353.

Gage, N., Poeppel, D., Roberts, T.P., Hickok, G., 1998. Auditory evoked M100 reflects onset acoustics of speech sounds. Brain Res. 814, 236–239.

Gage, N.M., Roberts, T.P., 2000. Temporal integration: reflections in the M100 of the auditory evoked field. Neuroreport 11, 2723–2726.

Gage, N.M., Roberts, T.P., Hickok, G., 2002. Hemispheric asymmetries in auditory evoked neuromagnetic fields in response to place of articulation contrasts. Brain Res. Cogn Brain Res. 14, 303–306.

Holland, B.S., Copenhaver, M., 1988. Improved Bonferroni-type multiple testing procedures. Psychologic. Bull. 104, 145–149.

Jancke, L., Mirzazade, S., Shah, N.J., 1999. Attention modulates activity in the primary and the secondary auditory cortex: a functional magnetic resonance imaging study in human subjects. Neurosci. Lett. 266, 125–128.

Jancke, L., Wustenberg, T., Scheich, H., Heinze, H.J., 2002. Phonetic perception and the temporal cortex. NeuroImage 15, 733–746.

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

Keating, P., Lahiri, A., 1993. Fronted velars, palatalized velars, and palatals. Phonetica 50, 73–101.

Kuhl, P.K., 2000. A new view of language acquisition. Proc. Natl. Acad. Sci. USA 97, 11850–11857.

Lahiri, A., Gewirth, L., Blumstein, S.E., 1984. A reconsideration of acoustic invariance for place of articulation in diffuse stop consonants: evidence from a cross-language study. J. Acoust. Soc. Am. 76, 391–404.

Lahiri, A., Reetz, H., 2002. Underspecified recognition. in: Gussenhoven, C., et al. (Eds.), Laboratory Phononology VII. Mouton, Berlin, pp. 637–675.

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. J. Comp. Physiol. A 181, 665–676.

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

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

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

Obleser, J., Lahiri, A., Eulitz, C., 2004. Magnetic brain response mirrors extraction of phonological features from spoken vowels. J. Cogn. Neurosci., in press.

Ohl, F.W., Scheich, H., 1997. Orderly cortical representation of vowels based on formant interaction. Proc. Natl. Acad. Sci. USA 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. Electroencephalogr. Clin. Neurophysiol. 94, 26–40.

Peterson, G., Barney, H., 1952. Control methods used in a study of the vowels. J. Acoust. Soc. Am. 24, 175–184.

Picton, T.W., Bentin, S., Berg, P., Donchin, E., Hillyard, S.A., Johnson Jr., R., 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. Neurosci. Lett. 221, 145–148.

Poeppel, D., Yellin, E., Phillips, C., Roberts, T.P., Rowley, H.A., Wexler, K., Marantz, A., 1996. Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. Brain Res. Cogn Brain Res. 4, 231–242.

Read, H.L., Winer, J.A., Schreiner, C.E., 2002. Functional architecture of auditory cortex. Curr. Opin. Neurobiol. 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. J. Clin. Neurophysiol. 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. Brain Res. Cogn. Brain Res. 15, 228–240.

Sarvas, J., 1987. Basic mathematical and electromagnetic concepts of the biomagnetic inverse problem. Phys. Med. Biol. 32, 11–22.

Scherg, M., Vajsar, J., Picton, T.W., 1990. A source analysis of the late human auditory evoked field. J. Cogn. Neurosci. 1, 336–355.

Scott, S.K., Johnsrude, I.S., 2003. The neuroanatomical and functional organization of speech perception. Trends Neurosci. 26, 100–107.

Steinschneider, M., Reser, D., Schroeder, C.E., Arezzo, J.C., 1995. Tonotopic organization of responses reflecting stop consonant place of articulation in primary auditory cortex (A1) of the monkey. Brain Res. 674, 147–152.

Stevens, K.N., 2002. Toward a model for lexical access based on acoustic landmarks and distinctive features. J. Acoust. Soc. Am. 111, 1872–1891.

Stevens, K.N., House, A.S., 1963. Perturbation of vowel articulations by consonantal context: an acoustical study. J. Speech Hearing Res. 6, 111–128.

Sussman, H.M., Fruchter, D., Hilbert, J., Sirosh, J., 1998. Linear correlates in the speech signal: the orderly output constraint. Behav. Brain Sci. 21, 241–259.

Vihla, M., Eulitz, C., 2003. Topography of the auditory evoked potential in humans reflects differences between vowels embedded in pseudowords. Neurosci. Lett. 338, 189–192.

Wang, X., 2000. On cortical coding of vocal communication sounds in primates. Proc. Natl. Acad. Sci USA 97, 11843–11849.

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. J. Cogn Neurosci. 13, 1–7.

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