Phonological features of speech segments are reflected in the Auditory Evoked Brain Response around 100 ms post stimulus onset

Jonas Obleser†, Carsten Eulitz†‡, Henning Reetz‡ and Aditi Lahiri‡

† Department of Psychology, University of Konstanz, Germany

‡ Department of Linguistics, University of Konstanz, Germany

E-mail: {jonas.obleser,carsten.eulitz,henning.reetz,aditi.lahiri}@uni-konstanz.de

ABSTRACT

Two studies were conducted to elucidate the influence of phonological properties of speech segments on their cortical representation. The aim was to show that timing and mapping of the N100m brain response elicited by natural speech reflect the detection of phonological features. Auditory evoked fields were recorded while subjects performed target token detection tasks. In the first study using natural German vowels and in a consecutive study using voiced CV-syllables, the processing of place of articulation was the main focus. Results of both experiments suggest that timing and topography of the cortical N100m response reflect properties of speech stimuli that correspond to abstract phonological features. Mutually exclusive places of articulation such as CORONAL and DORSAL are reflected in separate cortical centers of gravity in the auditory cortex. This holds true when vowels are perceived in isolation as well as when they are preceded by voiced stop consonants. N100m peak latency delivers additional and complementary insight in the dynamics of speech segment processing in the human brain.

1. 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 [1] recently proposed a model of speech recognition that emphasizes the role of phonological features instead of phoneme categories in the pre-lexical processing of speech. Their approach of feature-based automatic speech recognition led to the question whether abstract phonological features are also utilized by the auditory cortex when processing speech. Here, we used magnetoencephalography (MEG) to analyze responses of the human auditory cortex to German vowels (study I) and CV-syllables (study II).

We asked to what extent auditory evoked fields (AEF) mirror the extraction of phonological features from the signal, and to what extent these features drive the cortical mapping of speech sounds in terms of timing and topography. Specifically, we analyzed the magnetic counterpart of the most prominent brainwave deflection after stimulus onset, the N100m [2] (Fig.1), and investigated its modulation through systematic variations in the phonological feature place of articulation.

2. METHODS AND MATERIALS

Both studies reported here were conducted with healthy monolingual German right-handed subjects. In Study I, we used natural German vowels with differing place of articulation (CORONAL [i], [e], CORONAL-LABIAL [y], [ø], DORSAL-LABIAL [u], [o]) spoken by a male speaker and cut out of spoken words (10 kHz sampling rate, 350 ms duration, gaussian 50 ms fade in and 150 ms fade out). Study II examined the same place features in CV-syllables (LABIAL [b], CORONAL [d] and DORSAL [g] combined with CORONAL [ø] and DORSAL [o], respectively) spoken by a female speaker (22.05 kHz sampling rate, 350 ms, no fade in, gaussian 50 ms fade out).

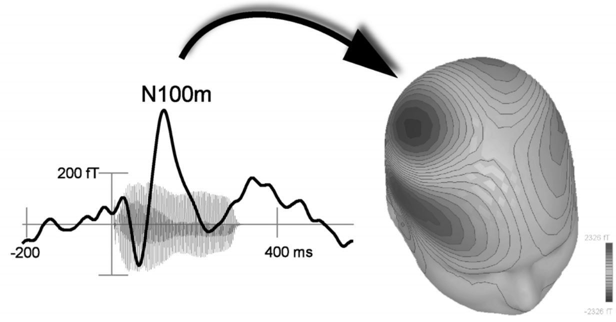

Fig. 1 A typical auditory evoked field from a channel over the right hemisphere is shown (left panel). The N100m is the most prominent response, appearing approximately 100 ms after syllable or vowel onset. Plausible source models of cortical activity can then be calculated from the field distribution of the N100m (right panel).

Design and data analysis were very comparable in studies I and II and are described in great detail in refs. [3;4]. In both studies, stimuli were presented in pseudo-randomized order via a non-magnetic auditory stimulus delivery system to both ears with 50 dB above individual hearing thresholds. To maintain subjects' attendance to the stimuli, a target

ISBN 1-876346-48-5 © 2003 UAB

detection task was employed in both studies: In every sequence, two given vowel or syllable categories had a lowered probability (10 %) and served as targets. Subjects had to press a button with their right index finger whenever a target occurred. As all stimulus categories consisted of six acoustically diverse tokens (they varied in pitch and formant frequencies), subjects had to map stimuli on phonetic categories to decide whether a given stimulus is a target or not, i.e., subjects had to maintain a phonological processing mode throughout the experiment. Subjects watched silent videos in order to maintain constant arousal and to reduce excessive eye movements.

Auditory magnetic fields (AEFs) evoked by vowel stimuli were recorded simultaneously from both hemispheres using a whole head neuromagnetometer (MAGNES 2500, 4D Neuroimaging) in a magnetically shielded room (Vaccumschmelze). Artifact-free and non-target epochs were averaged and filtered with a 20 Hz lowpass filter. Further analysis was confined to the rising slope and peak of the N100m component defined as the prominent deflection in the time range between 90 and 160 ms.

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. A median solution representing the center of gravity in cortical N100m activity was calculated for statistical analyses. Only significant interactions, main effects and post-hoc tests are reported here.

3. RESULTS AND DISCUSSION

Study I: Place of articulation in vowels is mirrored both in latency and topography of the N100m.

We analyzed N100m peak latency, i.e. the point of maximal cortical activation in this time range (*N*=20), as well as source topography, i.e. the relative spatial displacement of the cortical generator of N100m activity (*N*=14), due to the possible influences hemisphere (left, right), tongue height (HIGH, non-HIGH) and place of articulation (CORONAL, CORONAL-LABIAL, DORSAL-LABIAL) in a 2x2x3 repeated measures Analysis of Variance design.

Along the anterior-posterior axis, we found that place of articulation influenced the spatial mapping of cortical activity irrespective of the feature tongue height: sources of DORSAL vowels [u], [o] were located more posterior and separately from rounded and un-rounded CORONAL vowels [3]. As these two place features are mutually exclusive and never co-occur in time in the speech signal, their spatial separation in neural tissue is coherent with neural plasticity research [5] as well as with developmental research [6]. No vowel-specific hemispheric differences appeared.

N100m latency has been repeatedly proposed to be sensitive to changes in F_1 frequency, and the vowel [u] has been repeatedly found to elicit later N100m peaks compared to high-F_1 [a] [7].

We partly corroborated these findings (DORSAL vowels [u], [o] elicited by far the latest peaks), but also found latency differences that are not compatible with a simple F_1 explanation: Response latencies to un-rounded CORONAL vowel categories [i] and [e] differed highly significant by 6 ms – categories, that were not separable on a topographical level (cf. Fig. 2). This suggests that N100m peak latency delivers additional and complementary information in speech segment processing, especially when the spatial distinctiveness of vowel categories is low.

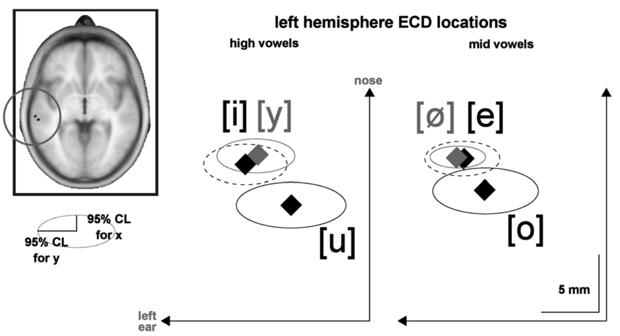

Fig. 2 ECD source locations in the left hemisphere are shown separately for HIGH vowels (left) and non-HIGH vowels (right). Ellipsoids indicate the mean 95% confidence regions across individual source locations. Note, that the spatial separation is most pronounced when just DORSAL is the distinctive feature, i.e. [u] vs. [y], [o] vs. [ø].

Study II: How is place of articulation in CV-syllables processed and reflected in N100m latency and topography?

Here, we orthogonally varied place of articulation both in the stop consonant and the vowel of CV-syllables (table I), and analyzed N100m peak latency (*N*=22), source location and orientation (*N*=16) in a 2x3x2 repeated measures Analysis of Variance with factors hemisphere (left, right), stop consonant place (LABIAL, CORONAL, DORSAL) and vowel place (CORONAL, DORSAL).

Table I. Charted are the phonetic features conventionally assigned to the vowels and stop consonants used.

		vowel place of articulation	
		CORONAL	DORSAL
stop place of articulation	LABIAL	[bø]	[bo]
	CORONAL	**[dø]**	**[do]**
	DORSAL	**[gø]**	**[go]**

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<.005] (Fig. 3b). This effect was of very comparable

ISBN 1-876346-48-5 © 2003 UAB

magnitude in both hemispheres. That is, a mapping of place of articulation driven by DORSAL vs. CORONAL vowels in auditory cortex was very comparable to the results of study I. Irrespective of the preceding stop consonant, the vowel which constitutes the articulatory target state [8] appeared to influence spatial mapping of speech sounds.

However, stop consonants also affected the sources of the N100m: The rotation of the ECD source in the posterior-anterior plane was by 6° more horizontal for CORONAL [d] syllable onsets than for DORSAL [g] onsets [F(1,15)=9.96, p<.01](Fig. 2a). LABIAL [b] onsets differed from neither category. The rotation of an equivalent current dipole can indicate that the underlying activity configuration 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. Interestingly enough, it is again the mutually exclusive place features CORONAL vs. DORSAL that show this significant change in source configuration.

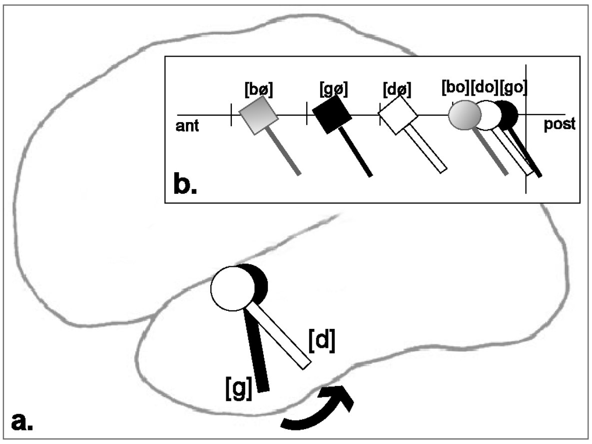

Fig. 3 a. A schematic illustration of the source orientation difference we found for CORONAL (white ECD symbol) vs. DORSAL (black) syllable onsets. **b.** Empirical data shows the combined orientation difference between CORONAL (white ECD symbols) and DORSAL (black ECD symbols) syllable onsets with the posterior-anterior location difference between CORONAL and DORSAL vowel syllables (i.e. square symbols are more anterior than circle symbols in all colors). LABIAL onsets (gray symbols) do not differ in orientation.

N100m latency was also sensitive to the combination of place features in the CV-syllables: Whereas the feature DORSAL led to prolonged responses in study I, here, the combination of a DORSAL consonant [g] with a DORSAL vowel [o] led to the latest N100m latency. On the other hand, when comparing [go] to [dø], [do], and [gø], it appears that the mere presence of the feature CORONAL elicits a faster N100m response - thus extending similar observations for the first syllable in non-words by Vihla & Eulitz [9].

4. CONCLUSIONS

Through the comparison of six HIGH and non-HIGH vowel categories of differing place of articulation, study I delivers conclusive evidence that the human auditory cortex may utilize more abstract feature information instead of phoneme categories to map the incoming speech signal: Neither tongue height nor labiality influenced the topographical differences found. These features may be mirrored in the temporal domain, i.e. differences in response latency, instead. How temporal and spatial information are integrated in the further processing of the speech signal remains to be examined more thoroughly.

Study II gives us an interesting idea how the interaction of differing place information in the continuous speech signal may be processed in auditory cortex: Features of articulatory target states which are available from formant transitions (which in turn are pre-processed in the peripheral auditory system and in primary auditory cortex before the signal enters the processing stage mirrored in the N100m, cf. [10]) may influence the center of cortical activity, whereas features inherent to the onset of a syllable may influence the extent and orientation of the active cortical patch.

Taken together, the studies show that a cortical mechanism of basic speech processing that relies on phonological features rather than phoneme categories is supposable and is a worthwhile subject to both psycholinguistic and neuroscientific research.

REFERENCES

[1] Lahiri, A. and Reetz, H., "Underspecified Recognition," in Gussenhoven, C. and Warner, N. (eds.) *Laboratory Phononology VII* Berlin: Mouton, 2002, pp. 637-675.

[2] Naatanen, R. and Winkler, I., "The concept of auditory stimulus representation in cognitive neuroscience," *Psychol.Bull.*, 125(6), pp. 826-859, 1999.

[3] Obleser, J., Lahiri, A., and Eulitz, C. "Magnetic brain response mirrors extraction of phonological features from spoken vowels", *J.Cogn.Neurosci.*, in press, 2003.

[4] Obleser, J., Elbert, T., Lahiri, A., and Eulitz, C., "Cortical representation of vowels reflects acoustic dissimilarity determined by formant frequencies," *Brain Res.Cogn Brain Res.*, 15(3), pp. 207-213, 2003.

[5] Buonomano, D. V. and Merzenich, M. M., "Cortical plasticity: from synapses to maps," *Annu.Rev.Neurosci.*, 21, pp. 149-186, 1998.

[6] Kuhl, P. K., "A new view of language acquisition," *Proc.Natl.Acad.Sci U.S.A*, 97(22), pp. 11850-11857, 2000.

 ISBN 1-876346-48-5 © 2003 UAB

[7] Roberts, T. P., Ferrari, P., Stufflebeam, S. M., and Poeppel, D., "Latency of the auditory evoked neuromagnetic field components: stimulus dependence and insights toward perception," *J.Clin.Neurophysiol.*, 17(2), pp. 114-129, 2000.

[8] Sussman, H. M., Fruchter, D., Hilbert, J., and Sirosh, J., "Linear correlates in the speech signal: the orderly output constraint," *Behav.Brain Sci.*, 21(2), pp. 241-259, 1998.

[9] Vihla, M., and Eulitz, C., "Topography of the auditory evoked potential in humans reflects differences between vowels embedded in pseudo-words", *Neurosci. Lett.*, 338, pp. 189-192, 2003

[10] Rauschecker, J. P., "Parallel processing in the auditory cortex of primates," *Audiol.Neurootol.*, 3(2-3), pp. 86-103, 1998.