

Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition

Carsten Eulitz and Aditi Lahiri

Abstract

■ A central issue in speech recognition is how contrastive phonemic information is stored in the mental lexicon. The conventional view assumes that this information is closely related to acoustic properties of speech. Considering that no word is ever pronounced alike twice and that the brain has limited capacities to manage information, an opposing view proposes abstract underspecified representations where not all phonemic features are stored. We examined this proposal using event-related brain potentials, in particular mismatch negativity (MMN), an automatic change detection

response in the brain that is sensitive to language-specific phoneme representations. In the current study, vowel pairs were presented to subjects, reversed as standard and deviant. Models not assuming underspecification predict equal MMNs for vowel pairs regardless of the reversal. In contrast, enhanced and earlier MMNs were observed for those conditions where the standard is not phonologically underspecified in the mental representation. This provides the first neurobiological evidence for a featurally underspecified mental lexicon. ■

INTRODUCTION

The mental lexicon is a part of the declarative memory containing all information necessary for efficient recognition of speech in humans. A vital issue in linguistics and brain research is the nature of the mental representation involved in the processing of sounds of natural language. From a linguistic point of view, not all features that can be extracted from the acoustic signal are required for recognizing speech sounds. Consequently, all predictable and nondistinctive information can be excluded from the mental lexicon (Kiparsky, 1993; Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002; Keating, 1988). The underspecified approach, spelled out in the featurally underspecified lexicon (FUL) model (Lahiri & Reetz, 2002), serves to explain how the perceptual system handles acoustic variance of single words across speakers and contexts beyond the possibility of merely storing all variance (Bybee, 2001). A word like “ten” can be pronounced as “tem” in “te[m] bags” or “teng” in “te[ŋ] gates” (Fitzpatrick & Wheeldon, 2002). FUL assumes just a single representation for all the variants, namely “ten,” with an /n/ that is not specified for its place of articulation, which is [CORONAL]. This underspecified /n/ does not conflict with the contextual variants [n], [m], and [ŋ]. Such an equivocal representation is, however, not the case for all sounds. Whereas words ending with /n/ have contextual variants,

words ending with /m/ like “cream” (place of articulation [LABIAL]) will not become “crea[ŋ]” in “crea[ŋ] car” or “crea[n]” in “crea[n] dress.” FUL accounts for the asymmetry by fully specifying the place of articulation for /m/ in “cream,” which conflicts with other places of articulation. A similar point can be made for [CORONAL] underspecification of vowels particularly in vowel harmony languages like Finnish (van der Hulst & van de Weijer, 1995). An unresolved question is whether the human brain indeed uses not only specified but also underspecified mental representations during speech recognition.

An ideal methodology for this purpose is mismatch negativity (MMN), an automatic change detection response in the brain, which has been shown to be an index of experience-dependent memory traces and to be sensitive to language-specific phoneme representations (Phillips et al., 2000; Winkler et al., 1999; Näätänen et al., 1997; Näätänen, 2001) and representations for words (Shtyrov & Pulvermüller, 2002; Pulvermüller et al., 2001). MMN is elicited by infrequent, deviant stimuli presented after a random number of frequent, standard stimuli. The standard stimuli create a central sound representation that is more abstract than the sum of perceived acoustic elements and correspond to the information content of the sound perception, the sensory memory, and the long-term memory (Näätänen, 2001; Cowan, 1999). That means that the central sound representation corresponds in part to the long-term memory traces and may thus convey information about the

phonological representation in the mental lexicon, which is, in linguistic terms, the underlying representation. The percept created by the infrequent, deviant stimulus is more low level and has vowel specific information available around 100 msec after stimulus onset (Obleser & Eulitz, 2002; Obleser, Elbert, Lahiri, & Eulitz, 2003; Poeppel et al., 1996; Eulitz, Diesch, Pantev, Hampson, & Elbert, 1995). This percept corresponds in part to the set of phonological features extracted from the acoustic signal, the so-called surface form. In this way the MMN can be used to study the difference between the surface form, extracted from the deviant, and the underlying representation, created by the standard.

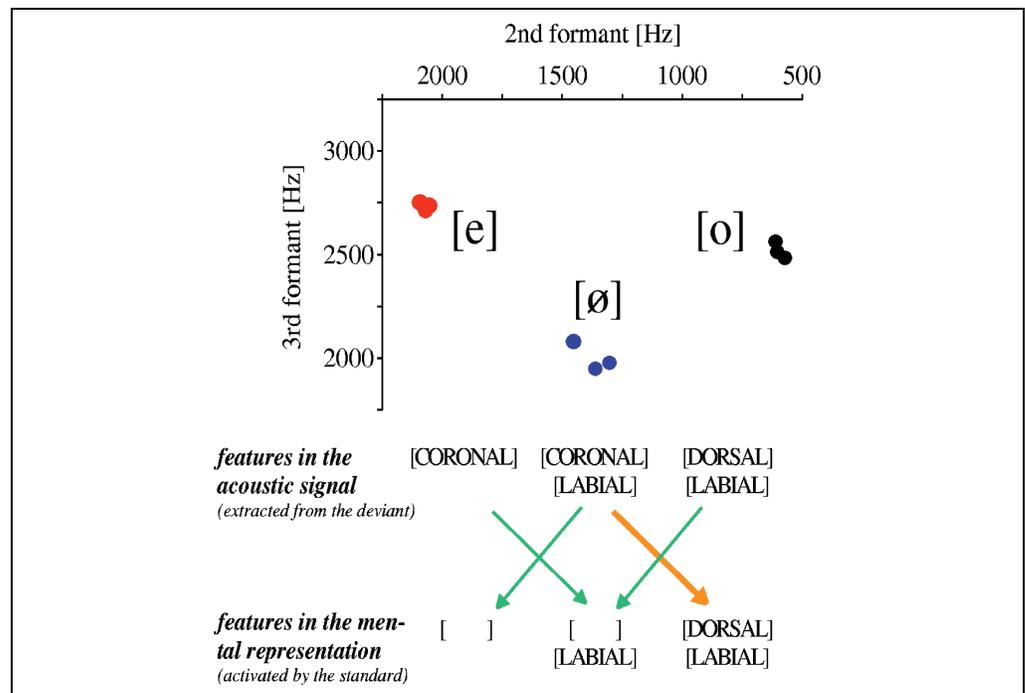
Our standard and deviant stimuli were three German vowels [e], [ø], and [o]. The acoustic differences between [e] and [ø] were similar to the difference between [ø] and [o]. The relevant phonological features (cf. Ghini, 2001; Halle, 1992; Lahiri & Evers, 1991) in the surface form are given in Figure 1. According to FUL, not all of these features are specified in the mental representation. The surface forms and mental representations are thus asymmetric for [CORONAL] vowels but symmetric for [DORSAL] vowels. This difference in symmetry and the fact that the place features [CORONAL] and [DORSAL] are mutually exclusive for vowels in any natural language (Lahiri & Evers, 1991) was exploited in the following way: Where features from the surface form and the mental representation are identical or not mutually exclusive, there is no conflict between the surface form and the representation. On the other hand, where they are mutually exclusive, for example, [CORONAL] and [DORSAL], the two features are in conflict. Depending

on the vowels chosen as deviant and standard stimuli in an oddball paradigm, the mapping of the surface features to the representation results in a conflict or nonconflict situation. The results of the mapping process are thus asymmetric as well and give clear predictions regarding the expected MMN.

Our hypothesis is that the MMN would be of a higher magnitude and show an earlier peak latency if the mapping involves a conflict rather than a nonconflict situation (Näätänen & Alho, 1997). The critical comparison is between /o/ and /ø/: A conflict occurs when /o/ is the standard and /ø/ the deviant, but not vice versa. The conflict occurs when the feature [CORONAL] extracted from the deviant stimulus [ø] maps onto the feature [DORSAL] in the mental representation created by the standard /o/, which will be referred to as [ø]_{/o/}. In contrast, there is no such conflict with the inverse mapping, when the feature [DORSAL] from the deviant [o] maps onto the underspecified representation of the standard /ø/, referred to as [o]_{/ø/} (see also lower part of Figure 1). In a parallel argument, we do not predict differential MMN when we consider the vowels /ø/ and /e/. The acoustic difference between these two vowels is about the same as between /ø/ and /o/. Nevertheless, there is no conflict in this pair of inversion, that is, in [ø]_{/e/} and [e]_{/ø/}. That is, for both vowels used as deviants, the feature [CORONAL] is extracted and does not conflict with the respective mental representations created by the standards. Hence, no difference in MMN is expected in this pair of inversion.

In contrast, the conventional models assume that the underlying representation of speech sounds corre-

Figure 1. Acoustic and phonological characteristics of the natural German vowels used. The upper part shows the locations of all vowels in the F2–F3 space. The lower part lists the corresponding sets of phonological features (except for the common feature [VOCALIC]) extracted from the incoming stimulus, that is, the surface form, and those in the underlying mental representation, assuming a featurally underspecified lexicon. The arrows illustrate the main statistical model. Green arrows indicate combinations of standard and deviant stimuli in nonconflicting conditions and the orange arrow the conflicting condition.



sponds directly to their acoustic signal properties (Klatt, 1989) or to closely related surface forms (Steriade, 2000; Ohala & Ohala, 1995). In that case, words are stored with all phonetic and redundant details, including variant pronunciations, such that all mental representations are fully specified (Bybee, 2001). These models make distinct predictions from FUL. For them, both inversion pairs of standard and deviant stimuli would have exactly analogous acoustic differences as well as levels of conflict, and no MMN difference between $[\emptyset]_{/o/}$ and $[o]_{/o/}$ or between $[\emptyset]_{/e/}$ and $[e]_{/o/}$ would be expected. The distinct predictions of the two opposing views on mental representation were tested by neurobiological means.

RESULTS

As shown in Figure 2, there is a clear MMN component in the grand average difference waveforms in all experimental conditions. Topographical information about the MMN is displayed in the lower part of Figure 2. Both the amplitude maps as well as the current source

density (CSD) maps show typical MMN topographies (Näätänen, 2001) with a predominant influence of left and right hemispheric temporal generators on the MMN. Based on the maps, the prominent differences between experimental conditions were found in amplitude and not in topography. Consequently, we restricted further analyses to the difference between Fz and linked mastoids as a transparent parameter for the integrated brain activity related to the ongoing change detection processes. The MMN components in the upper part of Figure 2 show a larger and earlier peak in the conditions where there is a conflict between surface form and underlying representation. This difference predicted by the FUL model was paralleled by similar MMNs in the pair of inversion with the two nonconflicting conditions. This difference pattern was evident for the frontal electrode position as well as for the root mean square (rms). Mean latencies and amplitudes of the MMN are summarized in Table 1. Two-way ANOVA revealed statistically significant interactions for the MMN latency, $F(1/11) = 10.53$; $p < .01$, the MMN amplitude at the

Figure 2. MMN waveforms (upper part) and the corresponding MMN topographies (lower part) for all vowel pairs reversed as standard and deviant are shown. Data corresponding to the main statistical model are presented in the first two columns. The third pair of inversion, with a considerably larger acoustic difference, is illustrated in the third column. The upper row shows the waveforms at the frontal electrode position (Fz) and the next row, the rms waveforms. The color of the waveforms indicates the deviant vowel, red for [e], blue for $[\emptyset]$, and black for [o]. Solid lines indicate the conflicting conditions and dotted lines, the nonconflicting conditions. The superimposed rectangle illustrates the time window for calculating the MMN amplitudes. The lower part of the figure shows the MMN topographies for the corresponding experimental conditions. The first row displays the spline interpolated amplitude maps (average reference) with a contour step of $0.25 \mu\text{V}$. Blue lines indicate negative potentials, red lines positive potentials. The second line shows the corresponding current source density (CSD) maps using a contour step of $0.025 \mu\text{V}/\text{cm}^2$.

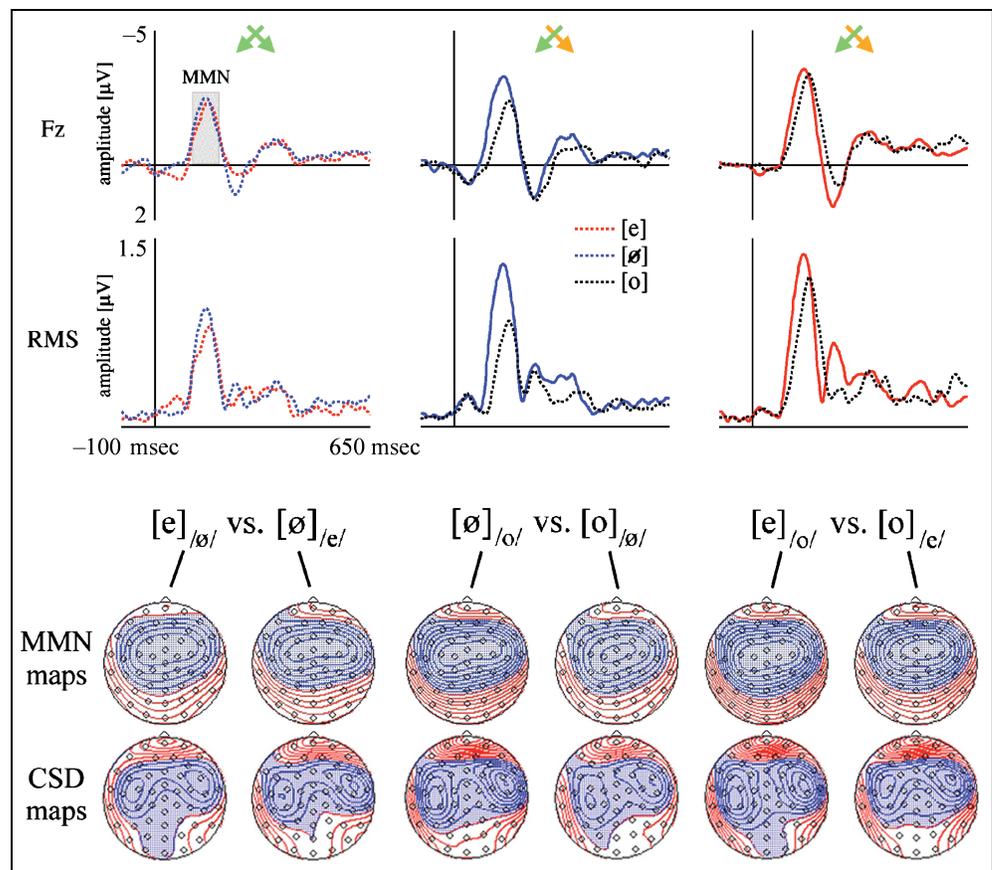


Table 1. Mean Peak Latencies and Amplitudes of the MMN for all Experimental Conditions

Experimental Condition	Peak Latency \pm SEM (msec)	Mean Fz Amplitude \pm SEM (μ V)	Mean rms Amplitude \pm SEM (μ V)
[e] _{/o/}	157.5 \pm 3.3	-1.83 \pm 0.30	0.75 \pm 0.09
[ø] _{/e/}	158.2 \pm 2.4	-1.95 \pm 0.27	0.83 \pm 0.11
[ø] _{/o/}	148.2 \pm 2.7	-2.80 \pm 0.31	1.17 \pm 0.13
[o] _{/ø/}	164.0 \pm 2.4	-1.77 \pm 0.18	0.72 \pm 0.07
[e] _{/ø/}	149.2 \pm 3.9	-2.94 \pm 0.21	1.22 \pm 0.09
[o] _{/e/}	167.0 \pm 3.6	-2.66 \pm 0.37	1.01 \pm 0.14

frontal electrode position, $F(1/11) = 5.21$; $p < .05$, as well as the rms amplitude, $F(1/11) = 8.17$; $p < .05$. Post hoc comparisons for the two pairs of inversion showed significant differences for the conflict versus nonconflict pair (all parameters $p < .001$) and no significance for any of the parameters in the other pair of inversion (all p values $> .47$). Post hoc tests comparing the conflicting pair to the three nonconflicting conditions (μ : -3, 1, 1, 1) revealed significant differences as well [latency: $F(1/11) = 36.02$; Fz amplitude: $F(1/11) = 19.38$; rms amplitude: $F(1/11) = 21.00$; all p values $< .005$].

Using a one-way ANOVA, the third pair of inversion ([o]_{/e/} vs. [e]_{/o/}) revealed main effects for the MMN latency, $F(1/11) = 24.02$; $p < .001$, and the rms MMN amplitude, $F(1/11) = 5.90$; $p < .05$, but no significance for the amplitude difference at the frontal electrode Fz.

DISCUSSION

Our results clearly support the predictions of the FUL model, which claims that the same vowel pair may be conflicting or not conflicting depending on which phonological features are specified in the underlying representation. Thus, if we take the [o]_{/ø/} versus [ø]_{/o/} pair, when /ø/ is the standard, and thus taps the underlying representation that is not specified for its place feature [CORONAL], there is no conflict with the deviant [o]'s surface representation [LABIAL]. However, if [ø] is the deviant, the surface representation does contain [CORONAL] and conflicts with the underlying representation [LABIAL] of /o/ when it is the standard, the underlying representation contains the specified feature [LABIAL]. According to these assumptions, the same acoustic contrasts trigger differential MMNs when they are reversed as standard and deviant in conflict situations, but not when both conditions of the pair of inversion are nonconflicting. In the nonconflict cases, [ø]_{/e/} and [e]_{/ø/}, the acoustic difference between standard and deviant is the same, and, correspondingly, the MMN does not vary. However, in the asymmetric cases [ø]_{/o/} versus [o]_{/ø/}, the MMN is significantly earlier and of larger amplitude when there is a conflict ([ø]_{/o/}) compared to when there is no conflict ([o]_{/ø/}). This MMN change in a conflict/nonconflict pair was corroborated by the third pair of inversion when

using the latency and the rms amplitude. The reduced amplitude effect for the frontal electrode position can be explained by ceiling effects. Obviously one of the two parallel processes contributing to the MMN (Näätänen, 2001), the acoustic change detection, was too dominant relative to the phoneme-specific processes.

This illustrates potential methodological difficulties in running these kinds of experiments. There are multiple stimulus characteristics, such as frequency, intensity, or timing (for reviews, see Näätänen, 2001; Picton, Alain, Otten, Ritter, & Achim, 2000), influencing the latency and the amplitude of the MMN in parallel. Moreover, when using spectrally complex stimuli it is almost impossible to control for all possible factors of influence at the same time in a perfect way. Therefore, we used a number of methodological solutions going beyond the standard MMN designs. First, we introduced acoustic variability within the vowel categories for both the standard and the deviant. This was done not only to simulate more natural speech perception conditions and to force the processing system to map the incoming acoustic signals onto more abstract representations, but also to avoid that the change detection as reflected in the MMN will be based on just one or a few particular acoustic features, which may even be unimportant for verbal processing. Second, critical comparisons were made between pairs of inversion, for example, [ø]_{/e/} and [e]_{/ø/}. This guarantees that the acoustic contrast between standard and deviant is equal and just the directionality is different. To our knowledge there is no study showing that the MMN is sensitive to the directionality of change in formant frequencies. The approach with the inversion of the role of standard and deviant worked out in the present study where the stimuli varied mainly in the F2 dimension, which is the most critical acoustic parameter when varying the place of articulation. But, given the amount of nonlinearities in the auditory system and the acoustic complexity of the speech sounds, it may well be that other contrasts that are not confined to an acoustic difference in the F2 dimension (e.g., vowel height) may not show symmetric MMNs after inversion even when both conditions are conflicting or both are nonconflicting. This possibility is, however, hypothetical in nature

and has to be investigated in further studies. Third, difference waveforms are based on within category differences (see Methods). This is important because event-related brain responses to different vowels show different waveform morphologies that may have nothing to do with any change detection processes. This is at least known for the N100 component (Obleser & Eulitz, 2002; Roberts, Ferrari, Stufflebeam, & Poeppel, 2000), which can be close to the MMN in latency and shows different latencies and amplitudes for different vowels. Consequently, when calculating the within-block difference waveforms (as done in most of the MMN studies, but not here) the difference between standard and deviant waveforms beyond the change detection will be superimposed in the difference waveforms. Last but not least, the present study shows that the different pairs of inversion for comparison should have acoustic contrasts as similar as possible (which were given for two pairs of inversion in the present study). This seems to be important to avoid ceiling effects that may lead to false negative amplitude results. In sum, we recommend that this set of special methodological solutions should be seriously considered to enable us to come to reasonable conclusions based on MMN differences when studying the processing of speech sounds or other acoustically complex stimuli.

The emphasis in this study has been on place features for vowels. The results indicate that the mental representation of phonological place features of vowels are not isomorphic with the acoustic information available to the listener, in the sense that its dimensions cannot be transformed into a space that exactly maps the acoustics of speech signals. What is represented is determined by both universal linguistic principles of underspecification and language-specific contrasts. This implies that the mental representation is more abstract and sparse than assumed by theories suggesting the storage of all acoustic information (Bybee, 2001). It may well be that on occasion listeners do store more information in the long-term memory than demanded by the phonological system, but it is equally obvious that an abstract representation (or at least a set of abstract features in the central sound representation) is the only way to account for the present results. More neurobiological experiments on different phonological dimensions across languages are naturally necessary to better understand the functional architecture of the mental lexicon.

In sum, these data show that in addition to mere acoustic changes, what influences MMN is the conflict between phonological features extracted from the deviant vowel with those specified in the mental representation created by the standard. This can be interpreted as neurobiological evidence that the human brain uses phonologically underspecified mental representations during vowel perception. Thus, the asymmetry predicted by the phonologically underspecified mental

representation of the FUL model is borne out by the neurobiological evidence.

METHODS

Stimulus Description

The standard and deviant stimuli were three different tokens of the three German vowels [e] (as in *bay*), [ø] (as in *Goethe*) and [o] (as in *go*), spoken by a male speaker. Acoustically, the vowels differed in F2 and F3, the tongue height, corresponding to F1, having been kept almost constant. The F0 (109–111 Hz) and the F1 (301–316 Hz) were close to each other for all vowel categories. As shown in the upper part of Figure 1, the [ø] and [e] had smaller F2 but larger F3 differences than [ø] and [o], resulting in relatively close overall acoustic differences in both pairs of stimuli. Stimuli of 200-msec duration (50-msec onset and offset ramps) were presented every 700 msec with a fixed intertrial interval binaurally via headphones. Stimuli were subjectively rated as equally loud and had equalized sound energy (root mean square with equally weighted frequencies). Measured sound pressure levels (linear scale) were 51 dB for the [ø] and [o], and 51.5 dB for the [e]. By using three tokens of each vowel, acoustic variability was introduced to simulate more natural speech perception conditions and thereby force the processing system to map the incoming acoustic signals onto more abstract representations.

Subjects and Task

Twelve right-handed undergraduate students of psychology (50% women; aged 21–30 years) underwent an electroencephalographic (EEG) study where two vowel categories were presented as standard and deviant stimuli in a passive oddball paradigm. The EEG was recorded from 65 electrode positions (Electrocap, Germany) against Cz as a reference between 0.1 and 30 Hz using a sampling rate of 500 Hz. The electrooculogram was coregistered and used to correct the EEG raw data for eye movements. During the recordings subjects were reading a self-chosen book and were asked to be quiet and to avoid excessive eye movements. During the study, 850 standards and 150 deviants per vowel category and block were presented. Standards immediately after the deviant were excluded from the averages. Epochs with artefacts exceeding 100 μ V or containing artefacts found by visual inspection were discarded.

In each experimental session the three vowel categories were combined in all possible pairs, with each vowel serving as a standard as well as a deviant, resulting in six blocks. The order of blocks was counter-balanced across subjects.

Data Analysis

The averaged standard waveforms of the same vowel in different blocks were statistically not different and therefore grand averaged across two corresponding blocks before being used for further analyses. These standards were the basis to calculate the difference waveforms that are used to extract the MMN. As there are slight changes in topography and timing of the components of the event-related potentials between vowels, we calculated the within-vowel-category differences. For instance, the MMN waveform to [e]-deviant measured in a block where /ø/ was the standard was calculated as [e]-deviant minus /e/-standard and will be reported as [e]_{/ø/}.

From the difference waveforms we derived the dependent variables. These were (i) the MMN latency measured at the maximum amplitude of the MMN at the Fz electrode position (rereferenced against linked mastoids) in the latency range from 90–210 msec post-stimulus onset and (ii) the MMN amplitude at Fz position (rereferenced against linked mastoids) measured as the mean amplitude across 80 msec centered at the mean MMN latency across subjects in the corresponding experimental condition (for latencies, see Table 1). As a more general amplitude measure, we used (iii) the rms amplitude across all electrodes, which was calculated for the same adjusted latency windows. These parameters were subjected to a two-way repeated measures ANOVA with the factor pair of inversion showing an equalized acoustic change (called later on pair of inversion), that is, [ø]_{/e/} versus [e]_{/ø/} and [ø]_{/o/} versus [o]_{/ø/}, and direction of change of the F2 frequency between deviant and standard (ascending: [e]_{/ø/} and [ø]_{/o/} vs. descending: [ø]_{/e/} and [o]_{/ø/}). Planned comparisons were used for post hoc testing. The statistical model was restricted to the two pairs of inversion with an almost equal acoustic change. The third pair of inversion ([o]_{/e/} vs. [e]_{/o/}) parallels the prediction for the conflict/nonconflict pair but shows a markedly larger acoustic difference between standard and deviant and was therefore statistically tested separately.

Acknowledgments

The research was supported by the Deutsche Forschungsgemeinschaft including the Leibniz funds. We thank B. Awiszus and A. Kupka for technical assistance, and T. Elbert, J. Fitzpatrick, A. Keil, J. Obleser, F. Plank, H. Reetz, B. Rockstroh, and L. Wheeldon for comments and discussions.

Reprint requests should be sent to Carsten Eulitz, Departments of Psychology and Linguistics, University of Konstanz, PO Box D25, 78457 Konstanz, Germany, or via e-mail: Carsten.Eulitz@Uni-Konstanz.de.

REFERENCES

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Cowan, N. (1999). An embedded-processes model of working

memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.

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.

Fitzpatrick, J., & Wheeldon, L. R. (2002). Phonology and phonetics in psycholinguistic models of speech perception. In N. Buton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological knowledge* (pp. 131–159). Oxford: Oxford University Press.

Ghini, M. (2001). *Asymmetries in the phonology of Miogiolia*. Berlin: Mouton.

Halle, M. (1992). Phonological features. In W. Bright (Ed.), *International encyclopedia of linguistics* (Vol. 3, pp. 207–212). Oxford: Oxford University Press.

Keating, P. A. (1988). Underspecification in phonetics. *Phonology*, *5*, 275–292.

Kiparsky, P. (1993). Blocking in nonderived environments. In S. Hargus & E. M. Kaisse (Eds.), *Studies in lexical phonology* (Vol. 4, pp. 277–314). New York: Academic Press.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. D. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 162–226). Cambridge: MIT Press.

Lahiri, A., & Evers, V. (1991). Palatalization and coronality. In C. Paradis & J.-F. Prunet (Eds.), *The special status of coronals: Phonetics and phonology 2* (pp. 79–100). San Diego: Academic Press.

Lahiri, A., & Marslen-Wilson, W. D. (1991). The mental representation of lexical form: A phonological approach to the recognition lexicon. *Cognition*, *38*, 245–291.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 637–676). Berlin: Mouton.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.

Näätänen, R., & Alho, K. (1997). Mismatch negativity (MMN)—The measure for central sound representation accuracy. *Audiology and Neuro-otology*, *2*, 341–353.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmonemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language specific phoneme representations revealed by electric and magnetic brain responses. *Nature*, *385*, 432–434.

Obleser, J., Elbert, T., Lahiri, A., & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by format frequencies. *Cognitive Brain Research*, *15*, 207–213.

Obleser, J., & Eulitz, C. (2002). *Extraction of phonological features from spoken vowels is mirrored by the MEG response*. Paper presented at the 13th International Conference on Biomagnetism, Jena.

Ohala, J., & Ohala, M. (1995). Speech perception and lexical representation: the role of vowel nasalization in Hindi and English. In B. Connell & A. Arvaniti (Eds.), *Papers in laboratory phonology IV: Phonology and phonetic evidence* (pp. 41–60). Cambridge: Cambridge University Press.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, *12*, 1038–1055.

- Picton, T. W., Alain, C., Otten, L., Ritter, W., & Achim, A. (2000). Mismatch negativity: Different water in the same river. *Audiology and Neuro-otology*, *5*, 111–139.
- Poeppl, 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. *Cognitive Brain Research*, *4*, 231–242.
- Pulvermüller, F., Kujala, T., Shtyrov, Y., Simola, J., Tiitinen, H., Alku, P., Alho, K., Martinkauppi, S., Ilmoniemi, R. J., Näätänen, R. (2001). Memory traces for words as revealed by the mismatch negativity. *Neuroimage*, *14*, 607–616.
- Roberts, T. P., Ferrari, P., Stufflebeam, S. M., & Poeppl, D. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology*, *17*, 114–129.
- Shtyrov, Y., & Pulvermüller, F. (2002). Neurophysiological evidence of memory traces for words in the human brain. *NeuroReport*, *13*, 521–525.
- Steriade, D. (2000). Paradigm uniformity and the phonetics–phonology boundary. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Acquisition and the lexicon* (pp. 313–334). Cambridge: Cambridge University Press.
- van der Hulst, H., & van de Weijer, J. (1995). Vowel harmony. In J. A. Goldsmith (Ed.), *The handbook of phonological theory* (pp. 495–534). Oxford: Blackwell.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*, 357–369.