

MEG evidence for phonological underspecification

Mary Ann Walter & Valentine Hacquard

Department of Linguistics & Philosophy, Massachusetts Institute of Technology

Email: {walterma,hacquard}@mit.edu

ABSTRACT

Subjects (n=12) listened to the synthesized syllable tokens *ba*, *da*, *ma*, and *na*. Evoked magnetic fields were recorded using a 93-sensor whole-head biomagnetometer array while subjects performed a same-different judgment task. A model of phonology which does not incorporate underspecification predicts no difference in the MMF for labial/coronal consonant pairs, regardless of which is presented as standard and which as deviant. An underspecification model, on the other hand, predicts an asymmetry – coronal deviants should elicit stronger MMFs than labial deviants. This asymmetry is exactly what we observe ($F(1,11)=7.218$, $p=.02$). This finding provides support for the existence of maximally abstract linguistic mental representations.

1. INTRODUCTION

Underspecification models maintain that phonological representations in the mind are maximally abstract, so that not all phonological features are stored (see Steriade 1995 for review). Such representations result in an efficient organization of the lexicon, and are also held to explain cross-linguistic facts about the behavior of some segments. One such model is the Featurally Underspecified Lexicon (FUL) model proposed by Lahiri and Marslen-Wilson (1991). This model maintains that the phonological features identified in the speech signal are mapped onto a phonologically underspecified mental representation. The unspecified features are not just those which are redundant in the language, but those which are contrastive but inactive.

The English consonant pairs /m/-/n/ and /b/-/d/ provide an example of such a case. The place of articulation feature [CORONAL] is assumed to be unspecified in English in the FUL model. [LABIAL], on the other hand, is assumed to be specified. Thus the presence of the [CORONAL] feature in a surface [d] conflicts with the underlying [LABIAL] specification of /b/. The opposite, however, does not hold – the [LABIAL] specification of surface [b] does not conflict with underlying /d/, since underlying /d/ has no place specification at all.

A sizable body of evidence has accumulated for this viewpoint in the linguistic literature, from investigations into language change and vulnerability to phonological rules as well as from experimental work on language comprehension (Lahiri & Reetz 2002). Eulitz *et al.* (2003) present neurobiological evidence as well, using EEG to measure the mismatch negativity (MMN) response to vowel pairs. They find that MMN responses are significantly earlier and of greater amplitude to vowel pairs involving a feature conflict, as defined by the FUL model, than those separated by equal acoustic distance which do not.

In this study we investigate the same question, using the consonant pairs discussed above and the MEG correlate of the MMN (the mismatch field response, or MMF). The mismatch response is automatic and auditory, and is evoked by a deviant stimulus following a sequence of standards. It peaks ~180-250 ms after the difference point between a standard and a deviant. Its field distribution is shown below in Figure 1.

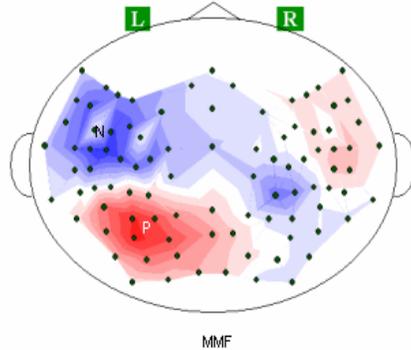


Fig. 1 – Magnetic field distribution of the MMF in response to speech stimuli for one representative subject. The letter P indicates the positive field (emerging from the brain) and the letter N the negative field (entering the brain).

Previous studies have shown that the MMN/F is sensitive to language-specific phoneme representations, and to abstract featural phonological properties in preference to acoustic difference (Näätänen *et al.* 1997, Sharma & Dorman 1999, Phillips *et al.* 2000). In addition, the oddball paradigms used to elicit it are hypothesized to result in a sound representation corresponding more to the underlying representation for the more frequently-played standard stimulus, but to the surface output for the infrequent, deviant stimulus (Eulitz *et al.* 2003; Näätänen 2001). This means that the MMF may also be used as a tool to compare underlying and surface levels of representation.

2. MATERIALS AND METHODS

Seventeen subjects, all students and employees at MIT, all of whom gave informed consent, participated in the experiment. The sample included 9 males and 8 females. All subjects were right-handed and had no history of hearing or neurological disorders. Eleven were native speakers of English. Five were native speakers of languages which also have the relevant contrasts (Czech (2), Spanish,

Russian, and Persian). They were also fluent in English. One subject was later excluded due to technical problems that arose during the experiment, three because they did not display the field distribution associated with the MMF during the appropriate time window, and one (Argentinian Spanish speaker) because his mismatch peak fell outside of two standard deviations from the mean.

Stimuli consisted of one token each of four CV syllables – [ba], [da], [ma], [na]. Stimuli were synthesized using the formant-based synthesizer HLSyn (Stevens & Bickley 1991). The HLSyn pseudo-articulatory system uses a small set of high-level (HL) parameters, including aerodynamic parameters, to control a low-frequency model of the human vocal tract and acoustic parameters to represent vocal tract shapes. A set of mapping relations automatically calculates the low-level control parameters of a formant synthesizer. Formant-based synthesizers have the benefit of providing direct, detailed control of the acoustic characteristics of the synthesized speech signal. In formant-based speech synthesizers, spectral and temporal parameters can be specified to control the generation of the synthesized speech signal. It has been demonstrated that highly intelligible speech that is nearly indistinguishable from recorded speech can be generated by copy-synthesis methods using a formant-based synthesizer (Hanson & Stevens 2002; Klatt & Klatt 1990). The stimulus syllables were also correctly identified and discriminated by speakers of languages with the relevant contrasts to ensure that they were good exemplars.

Subjects lay supine in a magnetically shielded room while stimuli were presented binaurally over earphones. Evoked magnetic fields were recorded using the MIT/KIT whole-head biomagnetometer array, with 93 axial gradiometers.

Before beginning the experiment, subjects listened to a 1kHz tone presented 100 times. Subjects then performed a mismatch detection task (oddball paradigm) where a series of four identical precursor stimuli, separated by inter-stimulus intervals of 400 ms, was presented followed by a fifth

stimulus, either identical to the previous four (control condition or *standard*), or different from it with respect to place of articulation of the initial consonant only (*deviant* condition). The stimuli were arranged into the eight conditions illustrated in Figure 2. The subjects heard each condition 30 times, in random order.

1)	ba	(400 ms)	da	deviant						
2)	da	(400 ms)	da	standard						
3)	da	(400 ms)	ba	deviant						
4)	ba	(400 ms)	ba	standard						
5)	ma	(400 ms)	na	deviant						
6)	na	(400 ms)	na	standard						
7)	na	(400 ms)	ma	deviant						
8)	ma	(400 ms)	ma	standard						

Fig. 2 – Stimulus conditions

Trial presentation was randomized by Pyscope script (Cohen *et al.* 1993). Subjects were instructed to press one button when the fifth item of each set was the same as the previous ones, and a second button if it was different. Trials were divided into six blocks of forty trials. Between blocks, subjects were given a break of self-determined duration.

Data were sampled at 1000 Hz, with acquisition between 1 and 200 Hz. The recording for each participant lasted approximately 25 minutes. The raw data was then subjected to a noise reduction routine to eliminate measured magnetic activity from external sources. Responses to stimuli were averaged by stimulus condition separately, in 700 ms windows keyed to the onset of the stimulus: 100 ms pre-, 600 ms post onset. The averaged signal was subjected to a low-pass filter below 30 Hz and adjusted to baseline using a 100 ms pre-stimulus interval.

Both button-press responses and reaction times were recorded. Reaction times were calculated from the onset of the fifth stimulus. Results for selected conditions are reported in Hacquard &

Walter (2003). In the analysis of the MEG data, averaged signals were first visually inspected to identify dipolar field distributions that appeared consistently across experimental conditions and across participants. Such distributions were identified in two time windows: the M100 window (150-170 ms) and the MMF window (225-275 ms). As shown in Figure 1, the MMF, being evoked by speech stimuli, was strongly left-lateralized (an MMF response to non-speech stimuli is more evenly distributed across both hemispheres). The amplitudes and latencies of these components were recorded by calculating the root mean square (RMS) field strength from sensors covering the field pattern of the MMF/M100 in the left hemisphere. The sensors used for the RMS analysis were selected for each subject by creating a grand average of all 8 conditions and choosing those sensors that showed the clearest dipolar distribution and held constant across conditions within a subject. The number of sensors used was 6, divided evenly between source and sink. We then compared the RMSs of the selected sensors for the response to the two coronal deviant conditions to those of the corresponding standards (e.g., the *da* response following a series of *ba* to the *da* response following a series of *da*), and likewise for the labial conditions.

3. RESULTS

A one-way repeated measures ANOVA showed that there was no significant difference in the average amplitude of the RMSs of selected sensors for the labial versus coronal mismatch condition in the 0-199 ms post-stimulus onset time window ($F(1,11)=.57, p=.47$), as shown in Figure 3 below.

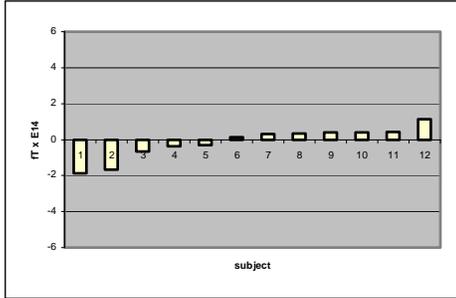


Fig. 3 – Coronal minus labial subtraction of per-subject average amplitudes of RMSs of selected sensors over the 0-199 ms post-stimulus onset time window.

However, in the MMF time window of 200-300 ms, a significant difference did appear ($F(1,11)=7.22, p=.02$).

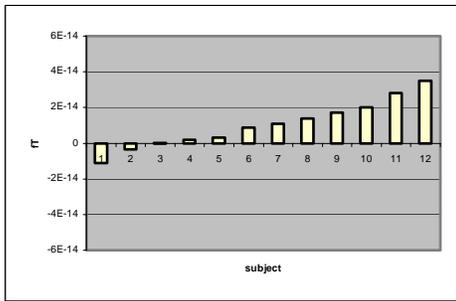


Fig. 4 – Coronal minus labial subtraction of per-subject average amplitudes of RMSs of selected sensors over the 200-300 ms post-stimulus onset time window.

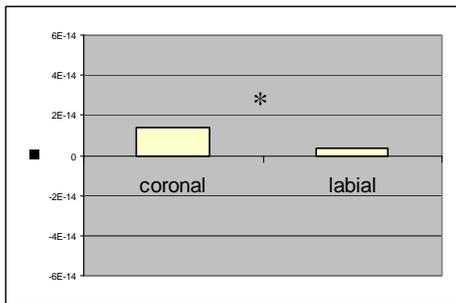


Fig. 5 – Cross-subject average amplitudes of RMSs of selected sensors over the 200-300 ms post-stimulus onset time window.

Finally, latency of the mismatch peak showed a trend toward being earlier for coronal deviants than for labial deviants, though this tendency did not reach

significance ($F(1,11)=2.72, p=.13$). Results per subject and across subjects are displayed in Figures 6 and 7, respectively.

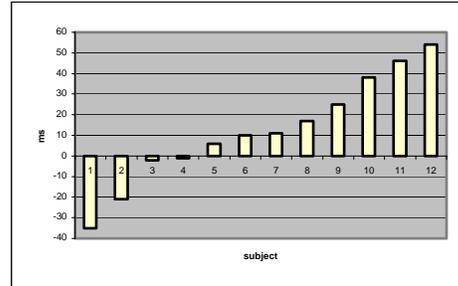


Fig. 6 – Labial minus coronal subtraction of MMF peak amplitudes, per subject.

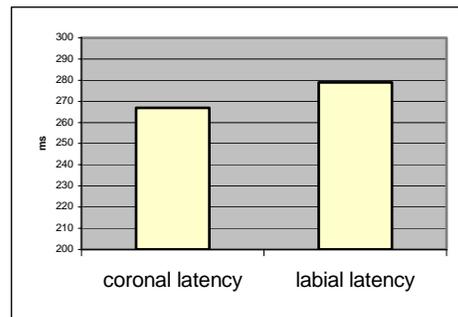


Fig. 7 – Average MMF peak latencies, across subjects.

4. DISCUSSION

The stimulus pair *ba/da* is acoustically equidistant regardless of the order in which it is presented – likewise for *ma/na*. They might be expected, therefore, to evoke a mismatch response of similar amplitude and latency regardless of ordering. However, instead we find that a significant difference does appear for amplitude, and a trend towards a distinction in latency. Following Eulitz *et al.* (2003), we hypothesize that these differences are due to comparisons being made between representations that differ not just in place of articulation specification, but in whether such a specification is present at all.

5. CONCLUSION

Models of phoneme representation in the mind which depend solely on acoustics and/or fully-specified representations fail to account for the asymmetry observed here in perception of consonant pairs. This suggests that differences in degree of phonological specification exist and are relevant in speech processing.

References

- Cohen, J.D., B. MacWhinney, M. Flatt & J. Provost. 1993. "Psyscope: A new graphics interactive environment for designing psychology experiments." *Behavioral Research Methods, Instruments and Computers*, **25**(2), 257-271
- Dehaene-Lambertz, E. Dupoux & A. Gout. 2000. "Electrophysiological correlates of phonological processing: a cross-linguistic study" *Journal of Cognitive Neuroscience* 12:4, 635-647
- Eulitz, C., A. Lahiri, J. Obleser, H. Reetz. 2003. "Brain electric activity reflects the underspecification of phonological features in the mental lexicon" *Proceedings of the 15th International Congress of the Phonetic Sciences*, 1631-1634
- Hacquard, V. & M.A. Walter. 2003. "Consonant confusability: An MEG Study." *Proceedings of the 29th Annual Symposium of the Berkeley Linguistics Society*
- Hanson, H. & K.N. Stevens. 2002. "A quasi-articulatory approach to controlling acoustic source parameters in a Klatt-type formant synthesizer using HLSyn." *J. Acoust. Soc. Am.* **112**, 1158-1182
- Klatt, D.H. & L.C. Klatt. 1990. "Analysis, synthesis, and perception of voice quality variations among female and male talkers." *J. Acoust. Soc. Am.* **87**, 820-857
- Lahiri, A. & W.D. Marslen-Wilson. 1991. "The mental representation of lexical form: A phonological approach to the recognition lexicon" *Cognition* 38, 245-291
- Lahiri, A. & H. Reetz. 2002. "Underspecified recognition" in C. Gussenhoven & N. Warner (eds.) *Laboratory Phonology VII*. Berlin: Mouton. 637-676
- 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., A. Lehtokoski, M. Lennes, M. Cheour, M. Huottilainen, A. Ilvonen, M. Vainio, P. Alku, R. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen & K. Alho. 1997. "Language-specific phoneme representations revealed by electric and magnetic brain responses." *Nature* 385, 432-434
- Phillips C., T. Pellathy, A. Marantz, E. Yellin, K. Wexler, D. Poeppel, M. McGinnis & T. Roberts. 2000. "Auditory cortex accesses phonological categories: an MEG mismatch study." *Journal of Cognitive Neuroscience*, 12
- Sharma A. & M. F. Dorman. 1999. "Cortical auditory evoked potential correlates of categorical perception of voice-onset time." *Journal of the Acoustical Society of America*, 106, 1078-1083
- Steriade D. 1995. "Underspecification and markedness" in John A. Goldsmith (ed.), *Handbook of Phonology*. Oxford: Blackwell. 114-174
- Stevens, K.N., & C. Bickley. 1991. "Constraints among parameters simplify control of Klatt formant synthesizer." *J. Phonetics* **19**, 161-174