Latency of MEG M100 Response Indexes First Formant Frequency

Krishna K. Govindarajan [(a)], Colin Phillips [(b)], David Poeppel [(c)], Timothy P.L. Roberts [(d)], Alec Marantz [(e)]

[(a)] Speech Communication Group, RLE, and [(e)] Dept. of Linguistics and Philosophy, Massachusetts Institute of Technology, Cambridge, MA 02139, [(b)] Dept. of Linguistics, University of Delaware, Newark, DE 19716, [(c)] Dept. of Linguistics and Zoology, University of Maryland, College Park, MD 20742, [(d)] Biomagnetic Imaging Laboratory, Dept. of Radiology, University of California, San Francisco, CA 94143.

Abstract: Magnetoencephalographic recordings from the auditory cortex of subjects show a close correlation between the timing of the evoked M100 response and the first formant frequency (F1) of vowels and vowel-like stimuli. These results are compatible with evoked magnetic field latencies elicited by tone stimuli, which show 100-300 Hz tones associated with latencies up to 30 ms longer than 500-3000 Hz tones. In Experiment 1, three-formant vowels /u,a,i/ were presented at two fundamental frequencies. The M100 latency was a function of the vowel identity and not F0: M100 was significantly shorter for /a/ than /u/. In Experiment 2, single-formant vowels were covaried with two F0 values. M100 latencies were shorter for /a/ (high F1) than for /u/ (low F1), at both F0 values. In Experiment 3, subjects listened to pure tone complexes with frequencies and amplitudes matching the F0 and F1 energy peaks of the stimuli in Experiment 2. M100 latencies showed the same pattern: latency covaried with the energy peak corresponding to F1, suggesting that the sensitivity to the energy in the F1 range is not specific just to speech stimuli.

Recent advances in electrophysiological studies of the brain are helping to provide insights into auditory and speech processing by humans (1, 2, 3, 4, 5). One technique, magnetoencephalography (MEG), is a non-invasive brain recording technique that measures the minute magnetic fields that are generated when large numbers of synchronously-active neurons fire. By taking magnetic field measurements from areas of the scalp close to the left and/or right auditory cortex, one can identify different sources of magnetic activity, and thus, potentially identify different components in speech perception.

One measure of magnetic activity is the evoked M100 response, which is the peak neuromagnetic activity that occurs around 100 ms after the stimulus onset. The M100 response can be characterized by its latency, amplitude, and brain localization. The M100 latency, which is the true time the peak occurs, has been characterized for pure tone stimuli as a decreasing function of the tone frequency (3). The latency is longest for low-frequency tones (approximately 130 ms for a 100 Hz tone), becoming shorter as the tone frequency increases, and finally asymptoting near 1000 Hz, where the latency is near 100 ms. Given that the M100 latency tracks the tone frequency, this paper presents three experiments that investigate the perceptual correlate of the M100 latency in the perception of vowels.

EXPERIMENTS

For all three experiments, subjects were placed on a bed inside a magnetically-shielded room. The first two experiments used a partial-head MEG sensor, which consists of 37 sensors, placed directly above the subject's left or right auditory cortex. The subject passively listened to speech or tone stimuli via an earphone to the ear contralateral to the MEG sensor. Experiment 3 used a whole-head system consisting of 148 sensors, where the subject's head was placed within the MEG sensor, and stimuli were presented to both ears. For each experiment, all stimuli were presented 100 times in pseudorandom order. Recordings were averaged and low-pass filtered at 1-20 Hz prior to further analysis. The source parameters for the M100 were determined by fitting a single dipole model to the sensor data. There were 6, 9, and 4 subjects in Experiments 1, 2, and 3, respectively. All subjects were native speakers of English.

Experiment 1 investigated whether properties of the evoked responses grouped vowel stimuli according to the fundamental frequency or according to the formant frequencies. Three-formant vowels (/a/, /i/, /u/) were syn-thesized at two different fundamental frequencies (Table 1) using the Klatt formant synthesizer (6). The results from Exp. 1 showed no significant separation among the different vowels in terms of the localization of the M100 response. However, there were systematic differences in the M100 latency for /a/, /i/, and /u/ ($F(2,5)=27.2$, $p<0.0001$), while no significant latency differences between the 'Male' and 'Female' vowels. Thus, the data suggests that the M100 latency does not track F0, but rather some aspect of the formant structure. The latency was significantly shorter for /a/ than /u/, consistent with /a/'s higher spectral center-of-gravity (7).

TABLE 1. Fundamental and formant frequency values for the stimuli in Experiment 1

Vowel	F0 (Hz)		F1 (Hz)		F2 (Hz)		F3 (Hz)	
	Male	Female	Male	Female	Male	Female	Male	Female
/a/	100	200	710	850	1100	1220	2540	2810
/i/	100	200	280	310	2250	2790	2890	3310
/u/	100	200	310	370	870	950	2250	2670

An analysis of F1 and the M100 latency shows a high correlation ($r = -0.97$): the higher the F1 frequency, the shorter the M100 latency, which shadows Roberts and Poeppel (1996)'s results for the pure-tone stimuli.

In Exp. 1, the formant frequencies covaried with F0. Thus, in order to better test whether M100 latency variation tracks F1, Experiment 2 used four single-formant vowels, where F0 and F1 were varied independently: F0 was either 100 or 170 Hz, and the F1 values were 720 Hz for /a/ and 300 Hz for /u/. Since a single-formant /i/ sounds very unnatural, it was not included in the stimulus set. Although the single-formant /a/ and /u/ did not sound entirely natural, subjects reported having no difficulty perceiving them as speech, and no difficulty in successfully categorizing them as /a/ or /u/. The results from Exp. 2 showed no latency effect for F0 variation, but a significant effect due to differences in F1 ($F(1,8)=4.2$, $p < .05$). In addition, the timing difference between responses to /a/ and /u/ was 5.2 ms, which is extremely close to what the pure-tone results predict, based on the F1 values for /a/ and /u/.

To verify that the M100 latency effects are due to auditory mechanisms and not speech-specific mechanisms, Experiment 3 used four two-tone stimuli that matched the center frequencies and amplitudes of the F0 and F1 energy peaks of the Exp. 2 stimuli. For all four stimuli, the amplitude of the tone corresponding to F1 was 8-10 dB greater than the tone corresponding to F0. The M100 latencies for the stimuli in Exp. 3 showed the same pattern as in Exp. 2: the latency covaried with the tone corresponding to F1, and did not covary with the tone corresponding to F0. The results support the conclusion that the latency sensitivity to the energy in the F1 range is not specific to just speech. In addition, these results are consistent with the results of Roberts et al. (1998), who have shown that for two tones at 100 and 1000 Hz, the M100 latency corresponds to the tone with the greater amplitude.

DISCUSSION

These experiments suggest that the M100 latency tracks the most psychophysically prominent peak in the auditory spectrum between 100 and 1000 Hz, whether it is a tone complex or a speech stimulus. Since the first formant frequency of a vowel primarily resides in this frequency range and usually has a greater amplitude than the fundamental frequency, the M100 latency tends to index F1. Moreover, the M100 timing does not seem to be based on a broader range of information about the vowel spectrum. If it did, then one might expect to find that M100 latencies would reflect some composite of multiple spectral peaks in the speech signal, including at least F0 and F1, and maybe even the higher formants. Further experiments are planned to help clarify if the M100 latency tracks this prominent peak based on a temporal or spectral representation.

ACKNOWLEDGMENTS

We would like to acknowledge the support of the following organizations for this research: NEC Corporation, NSF SGER grant #SBR-9423242, the McDonnell-Pew Center for Cognitive Neuroscience at MIT, the NSF-funded RTG Program 'Language: Acquisition and Computation' (#DIR9113607), NIH grant DC00194-03, and the JST funded project 'Mind Articulation' (co-PIs W. O'Neil, MIT and Y. Miyashita, U. of Tokyo).

REFERENCES

1. Diesch, E., Eulitz, C., Hampson, S., and Ross, B., *Brain and Language* **53**, 143-168 (1996).
2. Ragot, R. and LePaul-Ercole, R., *NeuroReport* **7**, 905-909 (1996).
3. Roberts, T. P. L. and Poeppel, D., *NeuroReport* **7**, 1138-1140 (1996).
4. Eulitz, C., Diesch, E., Pantev, C., Hampson, S., and Elbert, T., *Journal of Neuroscience* **15**, 2748-2755 (1995).
5. Poeppel, D., Phillips, C., Yellin, E., Rowley, H. A., Roberts, T. P. L., and Marantz, A., *Neuroscience Letters* **221**, 145-148 (1997).
6. Klatt, D.H. and Klatt, L. C., *Journal of the Acoustical Society of America* **87**, 820-857 (1990).
7. Chistovich, L., *Journal of the Acoustical Society of America* **77**, 789-805 (1985).
8. Roberts, T. P. L., Ferrari, P., and Poeppel, D., To be presented at the Cognitive Neuroscience Society meeting (1998).