Short communication

Auditory evoked M100 reflects onset acoustics of speech sounds

Nicole Gage [a], David Poeppel [b,*], Timothy P.L. Roberts [c], Gregory Hickok [a,1]

[a] *Department of Cognitive Sciences, University of California, Irvine, CA 92697, USA*
[b] *Department of Linguistics, Department of Biology, University of Maryland at College Park, 1401 Marie Mount Hall, College Park, MD 20742, USA*
[c] *Department of Radiology, Biomagnetic Imaging Laboratory, University of California, San Francisco 94143, USA*

Abstract

Magnetoencephalography (MEG) was used to investigate the response to speech sounds that differ in onset dynamics, parameterized as words that have initial stop consonants (e.g., /b/, /t/) or do not (e.g., /m/, /f/). Latency and amplitude of the M100 auditory evoked neuromagnetic field, recorded over right and left auditory cortices, varied as a function of onset: stops had shorter latencies and higher amplitudes than no-stops in both hemispheres, consistent with the hypothesis that M100 is a sensitive indicator of spectral properties of acoustic stimuli. Further, activation patterns in response to stops/no-stops differed in the two hemispheres, possibly reflecting differential perceptual processing for the acoustic–phonetic cues at the onset of spoken words. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Auditory; Speech; Language; Magnetoencephalography; Lateralization; Temporal processing; Cognitive neuroscience

Hemispheric asymmetry in the processing of speech and language is a well-documented phenomenon, and has led to a general acceptance of the doctrine of left hemisphere dominance for language. However, there is a fairly large body of evidence suggesting that such asymmetries are not uniform across all classes of linguistic computation [2,12,17], including the perception of classes of speech sounds [5]. For example, studies using the dichotic listening technique have shown that the right ear advantage (REA) for verbal stimuli—which is taken to indicate left hemisphere dominance—varies as a function of the class of speech sound: stop consonants produce a REA while vowels typically produce a smaller, less reliable REA [3,14,15].

Most studies of speech perception using hemodynamic-based neuroimaging methods (e.g., fMRI, PET) have concentrated on relatively coarse comparisons such as speech vs. non-speech sounds (e.g., tones, backwards words) [1,4,18], and not the within-speech contrasts relevant to the present issue. It is worth noting, however, that these studies have consistently shown *bilateral* activation of the superior temporal gyrus for speech and non-speech stimuli, particularly in passive listening conditions, which is consistent with the hypothesis that not all speech perception processes are strongly lateralized.

The lack of obvious hemispheric processing asymmetries in auditory cortex for speech perception in studies using hemodynamic-based methods does not imply that asymmetries do not exist. It could be that speech perception asymmetries are evident on a temporal scale that is not detectable using methods with a temporal resolution of seconds to minutes. Neuroimaging methods such as magnetoencephalography (MEG) reflect neural activity on a millisecond time scale and thus may be better suited to capturing rapid and transient effects. Two MEG studies are relevant. Poeppel et al. [10] evaluated hemispheric asymmetries in the auditory evoked M100 component using synthesized stop consonant CV syllables. M100 latencies in the left hemisphere for stop consonant syllables tended to be longer than those in the right, raising the possibility that the two hemispheres are treating speech stimuli differently in some way. In that study, however, only stop consonants were used and the paradigm tested selective attention, so there is no information about how speech sounds with different onset dynamics may be processed in the two hemispheres. Kuriki et al. [6] used MEG to examine the response to different types of speech sounds. The M100 latencies to stop consonants were shorter than

---

* Corresponding author. Fax: +1-301-405-7104; E-mail: dpoeppel@deans.umd.edu
[1] Also corresponding author. Department of Cognitive Sciences, University of California at Irvine, Irvine, CA 92697, USA. Fax: +1-949-824-2307; E-mail: gshickok@uci.edu.

for other types of speech sounds as measured over the left hemisphere, suggesting that onset dynamics of speech sounds are reflected in the M100 response. Unfortunately for present purposes, they did not record from right hemisphere sites, again leaving open the question of whether there might be hemispheric differences in the response to different types of speech sounds. A more complete examination of within- and across-hemisphere processing is needed in order to fully assess the time course of processing asymmetries for different speech cues.

We used MEG to investigate hemispheric processing asymmetries in the perception of speech sounds that differ in onset dynamics, parameterized as individually-presented words with initial consonants that are stop consonants (e.g., /b/, 'stops') or not (e.g., /f/, 'no-stops'). If the cues that occur at the onset of initial consonants provided the basis for processing asymmetries found both in behavioral and neuroimaging studies, then we might expect to see hemisphere differences in auditory evoked neuromagnetic responses that vary as a function of the speech onset properties. Of primary interest in this study, then, is whether the response to stops and no-stops will differ in the two hemispheres. Such differences would corroborate similar effects demonstrated behaviorally.

Five native English speakers (four male, mean age = 27.6 years) participated in the experiment. Each gave informed written consent; none reported any neurological abnormalities or hearing loss. All subjects were strongly right-handed [8].

Forty single-syllable words (duration ~ 500 ms) that varied in initial consonant were presented in a passive listening paradigm. Half of the words had initial stop consonants ('stops': /b/, /p/, /d/, /t/, /g/, /k/), half did not ('no-stops': /f/, /l/, /m/, /r/, /s/) but were otherwise matched for form class, voicing, and rest-of-word rhyme. Individual words were recorded by a male speaker using SoundEdit™16 (Macromedia) software and were presented digitally. Each word was presented 10 times pseudorandomly for a total of 400 epochs of 1 s duration each. Stimuli were presented binaurally at a comfortable listening level (~ 60–70 dB SPL) using Etymotic™ ER-3A earphones and air tubes designed for use with the MEG system. The frequency response of the system was flat (within ± 3 dB) up to 1250 Hz, attenuated by 10 dB at 3000 Hz, with a passband to 6000 Hz.

Neuromagnetic fields were recorded bilaterally using a twin 37-channel biomagnetometer system (Magnes II, BTi, San Diego, CA) in a magnetically-shielded room. Sensor-arrays were placed over the superior temporal lobes. Evoked responses to a 1000-Hz pure tone were evaluated to determine if the sensor arrays were positioned to effectively record the auditory evoked M100 field. Epochs of 1 s duration (100 ms pre-stimulus onset and 900 ms post-stimulus onset) were acquired around each stimulus at a sampling rate of 1041.7 Hz with a bandwidth of 400 Hz and a 1.0-Hz high-pass filter.

The data were selectively averaged by stimulus condition for each hemisphere. Averaged waveforms were digitally filtered using a high cut-off frequency of 40 Hz. The root mean square (RMS) of the field strength across all 37 channels was calculated for each sample point. The M100 latency and amplitude peak served as dependent measures. M100 peak was determined as the peak RMS value across 37 channels in the interval 80–150 ms. RMS values for the M100 latency peak were computed and the peak amplitude was determined as the maximum RMS value that best fit the dipole subject to a single-dipole fit exceeding 0.97 correlation in the 10-ms interval (± 5 ms) surrounding the M100 latency peak.

All stimuli reliably elicited a M100 evoked field response. Dipole fitting of the M100 and coregistration of the dipoles to individual subjects' MRIs showed that the M100 response localized to the upper bank of the superior temporal gyrus in both hemispheres, as has been observed by many investigators [7,9,11,13].

Latency. A 2 (hemisphere) by 2 (consonant) analysis of variance (ANOVA) with M100 *latency* (in ms) as the dependent variable produced a main effect of consonant type ($F_{1,4} = 20.942$, $p = 0.01$), with stops producing shorter latencies than no-stops, but no effect of hemisphere ($F_{1,4} = 0.003$, $p = 0.96$). However, there was a hemisphere $\times$ consonant type interaction which approached significance ($F_{1,4} = 4.785$, $p = 0.09$): stops processed in the right hemisphere had shorter latencies than those processed in the left, while no-stops in the left produced shorter latencies than those in the right (see Fig. 1a). Further analysis showed that consonant type (stops/no-stops) produced a highly reliable difference in the right hemisphere ($F_{1,4} = 15.517$, $p = 0.02$), but less so in the left ($F_{1,4} = 4.545$, $p = 0.10$).

Amplitude. A 2 (hemisphere) by 2 (consonant) ANOVA with M100 *amplitude* (in fT) as the dependent variable produced a main effect of consonant type ($F_{1,4} = 12.462$, $p = 0.02$), stops produced higher amplitudes than no-stops (see Fig. 1b), but no effect of hemisphere ($F_{1,4} = 0.298$, $p = 0.61$). There was no interaction of hemisphere and consonant type ($F_{1,4} = 0.387$, $p = 0.57$).

The key issue addressed in this study was whether hemispheric processing asymmetries would be found in the M100 in response to words that differed in onset dynamics. The results provide evidence for that asymmetry: latency of the M100 component for no-stops was longer in the right hemisphere than in the left and the latency difference for stops vs. no-stops was larger in the right hemisphere than in the left. Several additional findings are of note: latency for stops was shorter than latencies for no-stops in both hemispheres, replicating the findings of Kuriki et al. [6] for the left hemisphere and extending them to the right, and latency for stops was longer in the left hemisphere than in the right, replicating Poeppel et al. [10].

The amplitude analysis revealed that stops produced a larger response than no-stops in both hemispheres with no
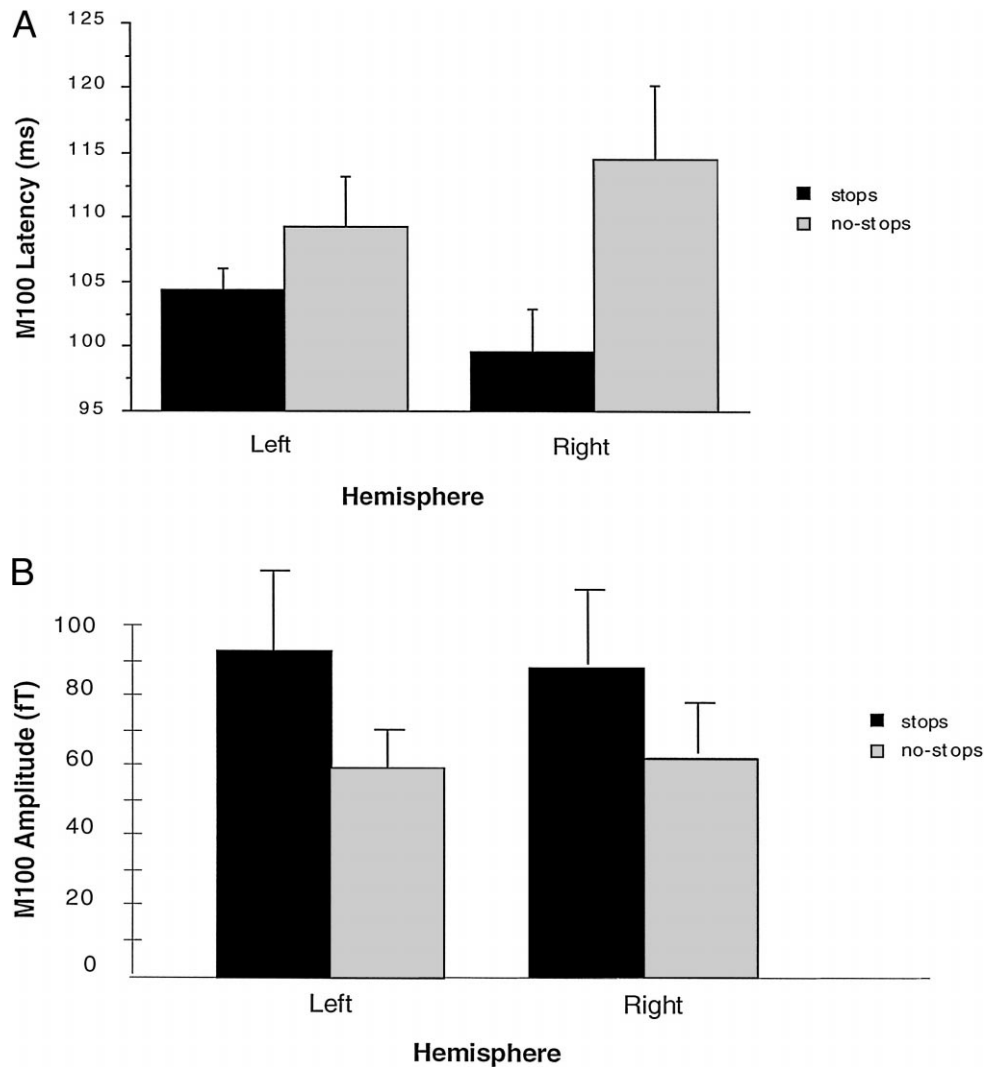
Fig. 1. (a) M100 latency (mean and S.E.) by hemisphere for stops and no-stops and (b) M100 amplitude (mean and S.E.) by hemisphere for stops and no-stops.

interaction of hemisphere and consonant type. This suggests that the activation levels were symmetric in the hemispheres. The fact that we only observed hemispheric asymmetries in latency and not amplitude may explain why similar asymmetries have not been observed using hemodynamic methods.

It remains to be determined exactly which stimulus feature(s) are driving these effects, and what the underlying processing mechanisms are doing. We can, however, propose the following hypothesis: the overall difference in latency to stops vs. no-stops reflects acoustic differences in the onset dynamics of these two classes of stimuli—stops contain more energy at onset than no-stops. That M100 latency is sensitive to stimulus amplitude differences has been shown for tonal stimuli [16] and is now extended to signal differences across speech sounds. The more pronounced difference in latencies in the right hemisphere for stops vs. no-stops may indicate that right hemisphere auditory cortex tracks the acoustic variation more 'di-

rectly' than does left hemisphere auditory cortex [17]. Left hemisphere auditory systems may reflect an analysis of auditory stimuli based not solely on acoustic properties but also on the phonetic restructuring [3,15,17] or temporal processing [4,14] mechanisms that have been proposed as necessary for their decoding.

References

[1] J.T. Binder, J.A. Frost, T.A. Hammeke, S.M. Rao, R.W. Cox, Function of the left planum temporale in auditory and linguistic processing, Brain 119 (1996) 1239–1247.

[2] C. Chiarello, C. Burgess, L. Richards, A. Pollock, Semantic and associative priming in the cerebral hemispheres: Some words do, some words don't… sometimes, someplaces, Brain and Language 38 (1990) 75–104.

[3] J. Dwyer, S. Blumstein, Ryalls, The role of duration and rapid temporal processing on the lateral perception of consonants and vowels, Brain and Language 17 (1982) 272–286.

[4] J.A. Fiez, M.E. Raichle, F.M. Miezin, S.E. Miezin, P. Miezin, P. Tallal, W.F. Katz, PET studies of auditory and phonological processing: effects of stimulus characteristics and task demands, Journal of Cognitive Neuroscience 7 (1995) 357–375.

[5] R. Ivry, P.C. Lebby, The neurology of consonant perception: specialized module or distributed processors? in: M. Beeman, C. Chiarello (Eds.), Right Hemisphere Language Comprehension: Perspectives from Cognitive Neuroscience, Lawrence Erlbaum, Mahwah, NJ, 1998.

[6] S. Kuriki, Y. Okita, Y. Hirata, Source analysis of magnetic field responses from the human auditory cortex elicited by short speech sounds, Experimental Brain Research 104 (1995) 144–152.

[7] N. Nakasato, T. Kumabe, A. Kanno, S. Ohtomo, K. Mizoi, T. Yoshimoto, Neuromagnetic evaluation of cortical auditory function in patients with temporal lobe tumors, Journal of Neurosurgery 86 (4) (1997) 610–618.

[8] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, Neuropsychologia 9 (1971) 97–113.

[9] C. Pantev, Evoked and induced gamma-band activity of the human cortex, Brain Topography 7 (4) (1995) 321–330.

[10] D. Poeppel, E. Yellin, C. Phillips, T.P.L. Roberts, H. Rowley, K. Wexler, A. Marantz, Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds, Cognitive Brain Research 4 (1996) 231–242.

[11] M. Reite, J.T. Zimmerman, J. Edrich, J.E. Zimmerman, Auditory evoked magnetic fields: response amplitude vs. stimulus intensity, Electroencephalography and Clinical Neurophysiology 54 (2) (1982) 147–152.

[12] E.D. Ross, R.D. Thompson, J. Yenkosky, Lateralization of affective prosody in brain and the callosal integration of hemispheric language functions, Brain and Language 56 (1) (1997) 27–54.

[13] M. Sams, R. Hari, Magnetoencephalography in the study of human auditory information processing, Annals of the New York Academy of Sciences 620 (1991) 102–117.

[14] J. Schwartz, P. Tallal, Rate of acoustic change may underlie hemispheric specialization for speech perception, Science 207 (1980) 1380–1381.

[15] M. Studdert-Kennedy, D. Shankweiler, Hemispheric specialization for speech perception, Journal of the Acoustical Society of America 48 (1970) 579–594.

[16] S. Stufflebeam, D. Poeppel, H.A. Rowley, T.P.L. Roberts, Perithreshold encoding of stimulus frequency and intensity in the M100 latency, NeuroReport 9 (1) (1998) 91–94.

[17] E. Zaidel, Language in the right hemisphere, in: D.F. Benson, E. Zaidel (Eds.), The Dual Brain: Hemispheric Specialization in Humans, Guilford Press, New York, NY, 1985.

[18] R.J. Zatorre, A.C. Evans, E. Meyer, A. Gjedde, Lateralization of phonetic and pitch discrimination in speech processing, Science 256 (1992) 846–849.