Spectral–temporal analysis of cortical oscillations during lexical processing

Kambiz Tavabi[a], David Embick[b] and Timothy P.L. Roberts[a]

We investigated the oscillatory neural correlates of auditory lexical processing in healthy adults. Synthetic aperture magnetometry was used to characterize the timing of event-related desynchronization (ERD)/event-related synchronization (ERS) in superior temporal gyri following low-frequency and high-frequency words in contrast to nonwords. ERS and ERD responses were found with both word and nonword stimuli. Analysis of power revealed significantly elevated θ–α range (6–14 Hz) ERD in response to words compared with nonwords (left hemisphere: 390–945 ms poststimulus). Furthermore, a burst of ERS in the γ band (40–50 Hz, centered at 410 ms poststimulus) distinguished high-frequency and low-frequency words, and also displayed left-hemispheric enhancement following words. Results demonstrate a clear neural correlate of lexical access and provide a basis for further study of spectral–temporal brain activity during language processing. *NeuroReport* 22:474–478 © 2011 Wolters Kluwer Health | Lippincott Williams & Wilkins.

[a]Lurie Family Foundations' MEG Imaging Center, Department of Radiology, Children's Hospital of Philadelphia and [b]Department of Linguistics, University of Pennsylvania, Philadelphia, Pennsylvania, USA

Correspondence to Professor Timothy P.L. Roberts, Department of Radiology, Children's Hospital of Philadelphia, Second Floor, Wood Building, Suite 2115, 34th Street and Civic Center Boulevard, Philadelphia, PA 19104, USA
Tel: + 1 267 426 0384; fax: + 1 215 590 1345;
e-mail: robertstim@email.chop.edu

Introduction

This study uses magnetoencephalography (MEG) to examine the brain's response to spoken words. Prior studies have identified a number of evoked-response components that correspond to different components of lexical access, including phonological [1,2], associated prelexical [3,4], and lexical stages [5–7]. Less is known about the time-course of oscillatory changes during spoken word recognition. Unlike evoked responses, which are phase-locked to the onset of experimental task or stimulation, changes in ongoing neural oscillations can occur as nonphase-locked (induced) responses. Here we examine the time-course of induced responses to spoken words and unintelligible acoustic control stimuli. The study uses MEG and synthetic aperture magnetometry (SAM), a beamforming approach to magnetic source imaging that has been successfully applied to spatio-temporal characterization of cortical oscillations without the need for time-averaging of stimulus-locked responses [8].

Both event-related desynchronization (ERD) and event-related synchronization (ERS) have been implicated in different cognitive domains, including language [9]. ERD of the α frequency band has been demonstrated with auditory [10] stimulation, as well as during cognitive and attention tasks [11,12]. Gamma-band ERS is found in many cortical areas and is induced by different stimuli or tasks [13]. In the domain of language, γ-band (30–120 Hz) oscillations have been associated with lexical processing in speech [14]. Hannemann *et al.* [14] found that word recognition in the context of acoustically

degraded stimuli produced left-lateralized induced γ-band activity of approximately 30–50 Hz with latency between 300 and 400 ms, interpreted as reflecting successful sound matching to auditory traces and lexical memory. Induced γ-band modulation between 300 and 600 ms has also been reported during semantic violations in sentence comprehension [15], suggesting sentence-level semantic integration. Finally, Shahin *et al.* [16] reported concomitant induced γ (30–60 Hz) enhancement and α (8–12 Hz) suppression in a lexical decision task. This growing body of evidence is developing a spectral–temporal picture of lexical access during speech perception that is consistent with the idea that this is a dynamic and internally complex process [17].

Here we report on auditory cortical ERD/ERS in response to spoken words divided into two frequency categories. The auditory modality is the natural one for language, but presents challenges because words are recognized as the speech signal unfolds over time, making it, for example, difficult to distinguish between simultaneously occurring auditory and lexical processing. To facilitate the analysis, the study uses spoken words that are comparable in uniqueness point (the point in the word at which it becomes uniquely identifiable [17,18]). Word frequency is known to affect word evoked neural responses [19]. Thus, we expected low-frequency and high-frequency words to differentially modulate ERD/ERS. We also used nonword control stimuli. Given that left-hemispheric activation is thought to be most relevant for language processing, we hypothesized left hemisphere asymmetry in lexically induced ERD/ERS distinguishing words from nonwords.

Materials and methods

Eighteen healthy human individuals (mean ± standard deviation age: 32 ± 9 years, eight male, self-reported right-handed) volunteered for the experimental procedure after giving an informed consent. The study was approved by the ethics committee of the Children's Hospital of Philadelphia (Pennsylvania, USA). One hundred and twenty monosyllabic English words were cut out from digitized (44.1 kHz) recordings of a female speaker. The length of words ranged from 397–623 ms, with a mean of 504 ± 56 ms. The resulting sound files were faded out with a 50-ms Gaussian ramp, low-pass filtered (12 kHz), and normalized for peak root mean square amplitude (audition 2.0, Adobe Inc., San Jose, California, USA). Spoken word-frequency counts were taken from the CELEX lexical database [20]. The mean Collins Birmingham University International Language Database (COBUILD, Birmingham, UK) spoken frequency per million for lower frequency words was 2.45. For the higher frequency words, the mean frequency per million was 64.43. The approximate recognition time point for each word was measured at the onset of its final phoneme, resulting in an average uniqueness point of 376 ± 73 ms (mean ± standard deviation), with no difference between high-frequency and low-frequency words ($P > 0.05$). The phonemic length of words was either 3 or 4 phonemes (3.2 ± 0.48), and 63.3% of words ended with a stop-consonant. The majority (105) of the words were previously published by Turner *et al.* [18]. The remaining 15 words were chosen using the Medical Research Council Psycholinguistic database [21] to match the psycholinguistic variables characterizing the material in reference. For each word, an acoustic nonword control was created by vocoding the temporal contour of each item by narrow-band white-noise centered at 1 kHz. The resulting items were distinctly word-like in prosody, but unintelligible. Words and nonwords were randomly arranged and separated by pseudorandomized interstimulus interval (2.1 ± 0.2 s). The stimuli were presented binaurally at an intensity of 45 dB above individual sensation level using Presentation software (Neurobehavioural Systems, Inc., California, USA) in an active listening paradigm by TDT Series III attenuation and amplification stages (Tucker-Davis Technologies, Alachua, Florida, USA) and ER3A transducers and ear inserts (Etymotic, Elk Grove, Illinois, USA). To control for changes in vigilance, individuals were required to identify and respond with a left-index button press to target stimuli (words referring to animals) that occurred randomly at a rate of 20%.

MEG measurements were carried out using a 275-channel whole-head biomagnetometer (Omega 275; VSM Medtech Inc., Coquitlam, Canada) in a quiet magnetically shielded room (1200 Hz sampling rate, third-order gradiometer environmental noise reduction). The individuals were seated comfortably in an upright position. They were asked to remain alert and compliance was verified using video-monitoring. Head motion was monitored using three fiduciary detection coils attached to the individuals' forehead and ears. Individuals' heads did not move more than 15 mm during recording. Anatomic, 1-mm isotropic resolution T1-weighted structural magnetic resonance images (magnetization prepared-rapid acquisition with gradient echoes) were obtained for each individual using a 3.0 T Magnetom Verio system (Siemens Medical Solutions, Erlangen, Germany).

Auditory cortex source localization bilaterally was based on a single equivalent current dipole modeling of word onset evoked (M100) response. Prior to source localization, all trials in the raw MEG data containing large amplitude artifacts exceeding 3.5 peak-to-peak threshold were removed. Subsequently, for each word the data were epoched into 800 ms (100 ms prestimulus) segments and averaged. A direct current-offset correction was applied based on the mean value obtained from the prestimulus interval. The word trial averaged data were band-pass filtered between 1.5–25 Hz. Bilateral single equivalent current dipoles (spatiotemporal) in a spherical head model were approximated to the magnetic field distribution around the maximum of the signal root mean square between 90–150 ms after stimulus onset. The target trials requiring a button press response were excluded from MEG data analysis.

In each individual, dipole locations were used as virtual sensors for the SAM beamformer. In order to probe ERD/ERS activity SAM was applied to the raw MEG trial data band-passed between 1.5–80 Hz (2.0 s epochs, 0.5 s prestimulus). Extending the epoch ensured the necessary length of baseline used to compute the signal covariance for each stimulus type at the virtual sensors. To visualize the ERD/ERS elicited by stimuli, time-frequency representation (TFR) plots were computed on the single-trial time series of source power estimated at the sensors and averaged across trials. TFR plots were calculated using a Morlet wavelet with center frequencies at 0.5 Hz intervals applied to the 1.5–80 Hz band-pass trial data, and offset corrected using the prestimulus baseline.
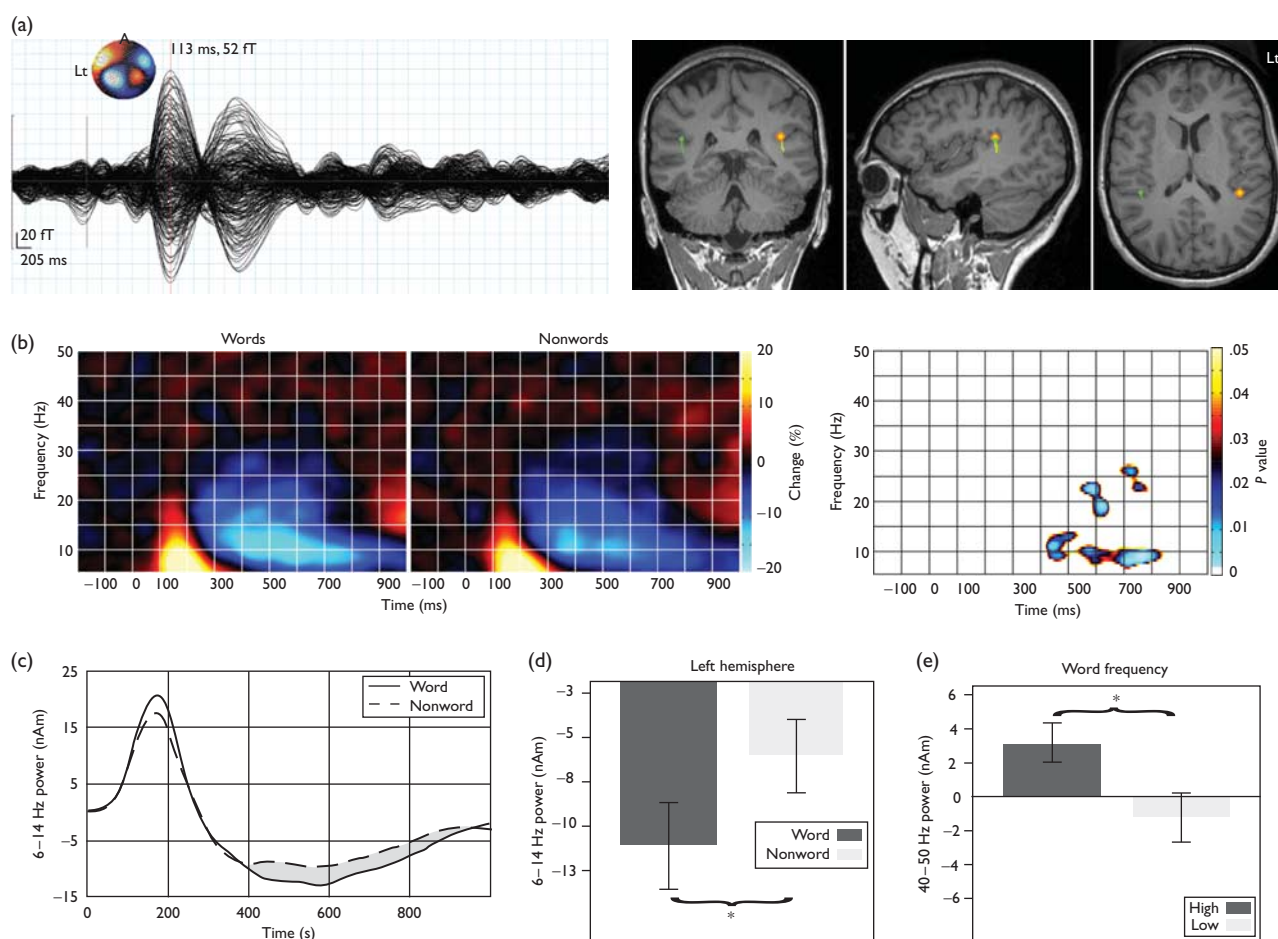
To characterize the time-course of auditory ERD/ERS in the TFR data, samplewise parametric testing was used to determine effects of lexicality (word vs. nonword) and word-frequency (high vs. low). Poststimulus spectral–temporal blocks of interest were identified by paired *t*-test (two-tailed). We computed the mean power for each spectral–temporal region of interest. Subsequently, repeated measures analysis of variance (ANOVA) with factors hemisphere, lexicality, and frequency was used to evaluate differential ERD/ERS. All reported significance levels for post-hoc tests were corrected for multiple comparisons using the Bonferroni adjustment.

Results

Overall, individual source models accounted for 80–90% of variance in the fit interval for the evoked response in superior temporal gyrus. In three individuals, single equivalent dipole fits were not possible for both hemispheres due to weak responses; for these a symmetric model was used. Furthermore, for four individuals with bilaterally weak responses we used a model that consisted of coordinates that were averaged across individuals with bilateral fits. Figure 1a shows an example of single individual-evoked MEG data and bilateral source model.

Figure 1b shows the grand average, collapsed across individuals and hemispheres, TFR plots of ERD/ERS in response to word and nonword stimuli. Oscillations were characterized by an evoked low-frequency ERS 100 ms poststimulus onset lasting for approximately 200 ms. This was followed by prolonged differential ERD in the upper-θ, α, and β ranges. The result of samplewise TFR contrasts indicated significant ERD (6–14 Hz) power modulation dissociating word from nonword stimuli approximately 400–900 ms poststimulus onset (Fig. 1b). On the basis of the time-course of mean 6–14 Hz power throughout the stimulus epoch (Fig. 1c), repeated

(a) Word average evoked-response magnetoencephalography sensor data and bilateral auditory source localization results in typical individuals. Cursor indicates maximal dipolar activity occurring at 113 ms (52 fT root mean square). (b) Time-frequency representation plots for increases (yellow/red) and decreases (blue) in auditory cortical source power following words and nonwords. The stimulus data were grand averaged across individuals (n = 18), over all trials and hemispheres in units of percent change with respect to a 200 ms baseline. On the right is the threshold samplewise paired t-test comparison between the mean responses after the stimulus types. An extended spectrotemporal region of significant difference can be seen ranging from approximately 6 to 14 Hz and commencing approximately 400 ms poststimulus onset and remaining significant for approximately 500 ms. Additional spectral–temporal regions of significance in the β range, with shorter duration appear between 600–800 ms. These appear to be spectrotemporally contiguous with the reported θ–α ERD. (c) Time-course of grand-average mean of 6–14 Hz activity defined by the spectral–temporal region of interest. Activity is plotted throughout the duration of stimulus types, and the shaded area indicates the interval where mean source power significantly differed between stimuli. (d) Results of analysis of variance for mean power. Posthoc analysis revealed (i) left-lateralized word greater than nonword ERD activity between 390–945 ms poststimulus *(P < 0.01), and (e) bilateral differential γ-band (40–50 Hz) enhancement after high-frequency words as opposed to low-frequency words in the interval 385–435 ms poststimulus *(P < 0.05). Post-hoc significance levels were corrected for multiple comparisons using Bonferroni adjustment.

measures ANOVA revealed a significant interaction between factors hemisphere and lexicality [F(1,17) = 11.42, $P < 0.01$; $\eta p^2 = 0.40$] in the interval 390–945 ms after stimulus onset. Posthoc examination of marginal means revealed significant left hemisphere ERD enhancement after words (-11.13 ± 2.5 nAm; mean $\pm$ standard error of the mean) compared with nonwords (-6.04 ± 2.1 nAm), ($P < 0.01$).

The samplewise TFR contrast between high-frequency and low-frequency words indicated a short burst of differential γ ERS (40–50 Hz), lasting approximately for 150 ms with onset of approximately 320 ms poststimulus. The time-course of ERS power throughout the stimulus epoch showed that the difference between high-frequency and low-frequency words was maximal at 410 ms. Therefore, ANOVA of mean 40–50 Hz power was carried out in the interval 385–435 ms poststimulus onset. ANOVA revealed a significant interaction between factors lexicality and frequency [F(1,17) = 9.16, $P < 0.01$; $\eta p^2 = 0.35$], as well as an interaction between hemisphere and lexicality [F(1,17) = 5.16, $P < 0.05$; $\eta p^2 = 0.23$]. Posthoc comparison of marginal means revealed: (i) a significantly enhanced ERS to high-frequency (3.2 ± 1.2 nAm) compared with low-frequency (-1.2 ± 1.5 nAm) words ($P < 0.05$) (Fig. 1e) and (ii) a significantly stronger left-hemispheric ERS after words (1.2 ± 1.3 nAm) compared with nonwords (-2.7 ± 1.3 nAm), ($P < 0.05$).

Discussion

The main findings of this study are (i) a significantly elevated θ–α (6–14 Hz) ERD for words compared with nonwords that is markedly evident in the left superior temporal gyrus 390–945 ms and (ii) a short burst of differential ERS in the γ-band (40–50 Hz) distinguishing high-frequency and low-frequency words, as well as words and nonwords in the left hemisphere between 385–435 ms poststimulus.

The word-related enhancement of θ–α power suppression is in agreement with a priori expectation of predominantly left-lateralized activity during language processing. Moreover, the finding that ERD time-course correlates with lexicality is consistent with evoked-response (e.g. M350, N400) studies of phonological and semantic processing of words [7,22]. The left-lateralized difference in α ERD between word and nonword stimuli appears to reflect a difference in left hemisphere processing according to lexicality. In contrast, approximately equal θ–α ERD is observed in the right hemisphere to both word and nonword stimuli, perhaps reflecting processing of specific acoustic properties shared between both stimulus types.

The left-lateralized word enhancement of γ-ERS, along with the dissociation between high-frequency and low-frequency words, appear to reflect neural correlates of lexical processing [14]. Moreover, the fact that γ-ERS differentiates between high-frequency and low-frequency words provides further evidence for the functional role of coherent auditory γ activity during language comprehension.

An important question for auditory presentation is how variable recognition timing can be controlled so that it is possible to look in finer detail at the time-course of spectral–temporal activity. Here we used words controlled for their uniqueness point, in a way that allows for an initial characterization of ERD/ERS activity with respect to spoken word onset; the particular uniqueness point definition used in this study is in agreement with a prior analysis of spoken word recognition [17], and stimulus selection criteria used in [18]. An important result of this study is that peak γ-ERS differentiation between high-frequency and low-frequency items occurred at c. 410 ms poststimulus onset, which is within 35 ms of average word uniqueness point for the stimuli in this study. Thus, within 35 ms of the point in time a word spoken in isolation can be considered uniquely identifiable; activity in superior temporal gyrus is differentially responsive to lexical attributes, such that items with a relatively higher frequency of occurrence induce stronger γ synchrony. The finding that high-frequency words elicit stronger induced γ is in agreement with the idea that increased spreading activation during lexical competition may be associated with more familiar items [17].

Conclusion

Our findings demonstrate recording and detection of the spectrotemporal dynamics associated with spoken word processing. The data reveal γ (40–50 Hz) synchrony at approximately the time of word recognition, overlapping and followed by extended event-related θ–α (6–14 Hz) desynchrony. Not only do the results show a signature of lexicality (in both the θ–α and γ domains) but also differential sensitivity to a property of the words, namely their frequency in the English language. Taken together, these findings motivate and provide a foundation for further studies of induced brain activity during linguistic processing.

References

1 Naatanen R, Lehtokoski A, Lennes M, Cheour M, Huotilainen M, Iivonen A, et al. Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature 1997; 385:432–434.

2 Parviainen T, Helenius P, Salmelin R. Cortical differentiation of speech and nonspeech sounds at 100 ms: implications for dyslexia. Cereb Cortex 2005; 15:1054–1063.

3 Mitterer H, Blomert L. Coping with phonological assimilation in speech perception: evidence for early compensation. Percept Psychophys 2003; 65:956–969.

4 Tavabi K, Elling L, Dobel C, Pantev C, Zwitserlood P. Effects of place of articulation changes on auditory neural activity: a magnetoencephalography study. PLoS One 2009; 4:e4452.

5 Friederici AD. Event-related brain potential studies in language. Curr Neurol Neurosci Rep 2004; 4:466–470.

6 Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, et al. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. Neuroimage 2002; 17:1101–1116.

7 Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. Nat Rev 2008; 9:920–933.

8 Barnes GR, Furlong PL, Singh KD, Hillebrand A. A verifiable solution to the MEG inverse problem. Neuroimage 2006; 31:623–626.

9 Hirata M, Goto T, Barnes G, Umekawa Y, Yanagisawa T, Kato A, et al. Language dominance and mapping based on neuromagnetic oscillatory changes: comparison with invasive procedures. J Neurosurg 2010; 112:528–538.

10 Krause CM, Lang H, Laine M, Kuusisto M, Porn B. Cortical processing of vowels and tones as measured by event-related desynchronization. Brain Topogr 1995; 8:47–56.

11 Freunberger R, Fellinger R, Sauseng P, Gruber W, Klimesch W. Dissociation between phase-locked and nonphase-locked alpha oscillations in a working memory task. Hum Brain Mapp 2009; 30:3417–3425.

12 Sauseng P, Klimesch W, Freunberger R, Pecherstorfer T, Hanslmayr S, Doppelmayr M. Relevance of EEG alpha and theta oscillations during task switching. Exp Brain Res 2006; 170:295–301.

13 Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends Cogn Sci 2005; 9:474–480.

14 Hannemann R, Obleser J, Eulitz C. Top-down knowledge supports the retrieval of lexical information from degraded speech. Brain Res 2007; 1153:134–143.

15 Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. Science 2004; 304:438–441.

16 Shahin AJ, Picton TW, Miller LM. Brain oscillations during semantic evaluation of speech. Brain Cogn 2009; 70:259–266.

17 Marslen-Wilson WD. Functional parallelism in spoken word-recognition. Cognition 1987; 25:71–102.

18 Turner JE, Valentine T, Ellis AW. Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. Mem Cognit 1998; 26:1282–1291.

19 Embick D, Hackl M, Schaeffer J, Kelepir M, Marantz A. A magnetoencephalographic component whose latency reflects lexical frequency. Brain Res Cogn Brain Res 2001; 10:345–348.

20 Baayen RH, Piepenbrock R, Gulikers L. The CD-ROM Version of the CELEX Lexical Database. Nijmegen: Max Planck Institute for Psycholinguistics; 1995.

21 Coltheart M. The MRC psycholinguistic database. Q J Exp Psychol 1981; 33A:497–505.

22 Pylkkanen L, Marantz A. Tracking the time course of word recognition with MEG. Trends Cogn Sci 2003; 7:187–189.