Published in final edited form as:

Brain Cogn. 2009 August ; 70(3): 259–266. doi:10.1016/j.bandc.2009.02.008.

Brain oscillations during semantic evaluation of speech

Antoine J. Shahin¹, Terence W. Picton², and Lee M. Miller^{1,3}

¹ Center for Mind & Brain, University of California, Davis

² Rotman Research Institute of Baycrest, University of Toronto, Toronto

³ Department of Neurobiology, Physiology, & Behavior

Abstract

Changes in oscillatory brain activity have been related to perceptual and cognitive processes such as selective attention and memory matching. Here we examined brain oscillations, measured with electroencephalography (EEG), during a semantic speech processing task that required both lexically mediated memory matching and selective attention. Participants listened to nouns spoken in male and female voices, and detected an animate target ($p = 20\%$) in a train of inanimate standards or vice versa. For a control task, subjects listened to the same words and detected a target male voice in standards of a female voice or vice versa. The standard trials of the semantic task showed enhanced upper beta (25–30 Hz) and gamma band (GBA, 30–60 Hz) activity compared to the voice task. Upper beta and GBA enhancement was accompanied by a suppression of alpha (8–12 Hz) and lower to mid beta (13–20 Hz) activity mainly localized to posterior electrodes. Enhancement of phase-locked theta activity peaking near 275 ms also occurred over the midline electrodes. Theta, upper beta, and gamma band enhancement may reflect lexically mediated template matching in auditory memory, whereas the alpha and beta suppression likely indicate increased attentional processes and memory demands.

Keywords

EEG; Gamma band activity; Theta activity; Alpha activity; Beta activity; Semantic; Speech

1. INTRODUCTION

Speech processing entails a spectrum of neural mechanisms ranging from low-level perception to high-level cognition. Low-level mechanisms analyze the various spectrotemporal cues present in the sound, whereas higher-level processes match these acoustic characteristics to learned templates in memory. In addition, high-level processes not specific to language, such as selective attention, are normally recruited during speech tasks to improve sensory processing and control response selection.

Many brain mechanisms elicited during speech and other auditory processing have been related to the event-related potentials (ERPs) and the oscillations of the electroencephalogram (EEG). ERPs are, by definition, phase-locked to experimental events. However, oscillatory brain activity can occur as either phase-locked (“evoked”) or non-phase-locked (“induced”) responses. Like ERPs, evoked oscillations are present in the time-domain average of all trials. Induced oscillations on the other hand, due to the temporal jitter of activity from trial-to-trial,

Corresponding author: Antoine J Shahin, UC Davis Center for Mind and Brain, 267 Cousteau Place, Davis, CA 95618, Tel: 530 297 4426, Fax: 530 297 4603.

Competing interests’ statement: The authors declare that they have no competing financial interests.

may cancel out in the time-domain average and therefore can only be examined in single trial analysis. ERPs, evoked oscillations, and induced oscillations can reflect different functional relevance in auditory perception and cognition (Pantev, 1995; Shahin, Roberts, Chau, Trainor, & Miller, 2008). Their similarities and distinctions help provide a framework for understanding bottom-up (i.e. stimulus driven) and top-down processes during speech processing.

Examples of ERPs related to speech and attention are the N400 and P300 components. The N400 is evoked following a semantic violation and may reflect lexical search (Friederici, 2004; Kutas & Hillyard, 1980), while the P300 reflects selective attention and update of working memory (Picton, 1992; Polich, 2007). Oscillatory brain activity may reflect similar mechanisms or correlate with precursors or consequences of the N400 and P300 processes. For example, theta activity (4–7 Hz) reflects increased language processing demands (Bastiaansen, Oostenveld, Jensen, & Hagoort, 2008) and, like the N400, is induced during semantic violations (Hald, Bastiaansen, & Hagoort, 2006). Bastiaansen et al. (2005) specifically showed that during a lexical-semantic retrieval task, theta power increased over the left hemisphere. This theta enhancement was accompanied by alpha and beta (8–12 Hz and 13–30 Hz, respectively) suppression, which is known to correlate with the level of attentional and memory load (Bastiaansen, van der Linden, Ter Keurs, Dijkstra, & Hagoort, 2005; Fries, Reynolds, Rorie, & Desimone, 2001; Van Winsum, Sergeant, & Geuze, 1984). In contrast, high-frequency oscillations in the gamma band (gamma band activity, or GBA, 30–120 Hz; Pantev, 1995) have been shown to support specific and non-specific neural processes. These include enhanced selective attention (Fries et al., 2001; Snyder & Large, 2005; Sokolov, Pavlova, Lutzenberger, & Birbaumer, 2004), lexical and semantic processing in speech (Braeutigam, Bailey, & Swithenby, 2001; Hannemann, Obleser, & Eulitz, 2007; Pantev, 1995; Pulvermüller et al., 1996) and memory matching (Hannemann et al., 2007; Herrmann, Lenz, Junge, Busch, & Maess, 2004; Lenz, Schadow, Thaerig, Busch, & Herrmann, 2007; Shahin et al., 2008).

Of particular importance to the current study is how oscillatory activity reflects semantic evaluation of speech. Perception is enhanced when the acoustical and higher-level (e.g. phonetic, lexical) representations in memory are well matched. Lenz et al. (2007) found that when compared to novel sounds, familiar sounds produced induced GBA (30–40 Hz) between 300 and 500 ms, reflecting successful sound matching to representations in long-term memory. Similarly, auditory lexical retrieval would require matching a word's acoustical representation and its meaning in memory. Left lateralized induced (non-phase-locked) gamma band activity around 30–50 Hz with a latency between 300 and 400 ms occurs when subjects identify acoustically degraded words as intelligible (Hannemann et al., 2007). Hannemann et al. (2007) attributed their findings to a process whereby speech intelligibility can be accomplished through matching of heard items to traces in auditory and lexical memory. However, enhanced GBA between 300 and 600 ms is also indicative of semantic and world knowledge violations in a sentence (Braeutigam et al., 2001; Hagoort, Hald, Bastiaansen, & Petersson, 2004) reflecting processes other than memory match such as sentence-level semantic integration.

Here we examined oscillatory brain activity present in EEG in a task that required matching speech sounds to prior knowledge or representations in auditory and lexical memory. We hypothesized that if the dynamics of oscillatory brain activity reflect enhanced auditory and lexical template matching in speech, we should observe enhanced theta (4–7 Hz) and gamma (30–50 Hz) activity during lexical-semantic evaluation of spoken words (Hannemann et al., 2007; Pantev, 1995). To test this theory, subjects listened to monosyllabic words, spoken in male and female voices, and discriminated a target word based on its meaning. Because the time to evaluate the meaning of each word varies with the type of word, theta and gamma oscillation should not phase-lock to the onset of the stimulus (i.e. they should be “induced”). To isolate the effect of semantic mechanisms from incidental effects due to differences in

spectrotemporal energy of words, subjects made a similar discrimination based on the voice (male voice in a train of female voices) using identical word stimuli. The difference between the two tasks should then reveal activity specific to semantic evaluation. This effect should be particularly strong later in the words, because semantic evaluation requires more time compared to voice (pitch) identification, which can occur very early after word onset.

2. METHODS

Analysis and results of the event-related potentials of this data set, including the P300 results related to targets, have been previously reported in Shahin et al. (2006). The emphasis in that report was on target effects. Here, entirely new analyses and results are reported. Because of motor response confounds, target effects were excluded.

2.1. Subjects

We examined oscillatory activity in ten normal-hearing subjects (mean age 29 years; range 23–38; 3 females). Nine of the subjects were right-handed, and one ambidextrous (Edinburgh Handedness Inventory). All subjects were fluent in English and gave informed consent prior to the experimental sessions.

2.2. Stimuli

Stimuli were 100 monosyllabic nouns for animate (e.g. ant) and inanimate (e.g. chair) objects (50 animate and 50 inanimate). Each word was recorded in four different voices (two males and two females) using Adobe Audition with a sampling rate of 12207 Hz. Words had a mean duration of 600 ms (range 275 – 984 ms). Fundamental frequencies for the two males were 131 Hz and 155 Hz and for the two females 191 Hz and 203 Hz. Words were presented binaurally with an inter-stimulus interval (offset to onset) of 2 seconds via a Tucker Davis Technologies RP2 stimulus box, through Etymotic E3A insert earphones at a root mean square intensity of 73 dBA SPL (RMS range of 66 – 80 dBA SPL, approximately 60 dB hearing level or HL) calibrated by a Brüel & Kjær sound-level meter. Auditory thresholds at 1000 Hz were less than 20 dB HL in all subjects. The experimental session took place in an acoustically shielded room.

2.3. Task

The experiment consisted of two oddball tasks (targets at $p = 0.2$): a semantic task and a control voice task. Each task consisted of two sessions with 250 words each. In the semantic task, subjects identified words (randomly spoken in a male or female voice) with animate meaning in a train of words classified as inanimate, or vice versa for the second block. In the control task, subjects identified words (randomly animate or inanimate) spoken with a male voice (target) in a train of words spoken in a female voice (standard), or vice versa for the second block. Target words were only repeated once in each session, while standard words were randomly repeated four times. Sessions were presented in random order for each subject. Target and standard words were randomized within each session but randomization was kept constant across all subjects. Participants fixated their vision on a cross displayed on a screen. Subjects pressed a button using their right index finger when they heard the target. A training session preceded the actual experiment.

2.4. Recording

Continuous EEG was recorded from a 64-channel EEG cap (Electro-Cap international, 10–10 system) using a Neuroscan Synamps (El Paso, TX, U.S.A.) amplifier (DC to 100 Hz sampled at 250 Hz) and referenced to Cz with a ground at AFz. The channel configuration was as follows: Frontal, AF3, AF4, AF7, AF8, F1, F2, F3, F4, F5, F6, F7, F8, Fz, Fp1, Fp2, Fpz;

Fronto-central, FC1, FC2, FC5, FC6, FCz; Central, C1, C2, C3, C4, C5, C6; Parieto-central, CP1, CP2, CP5, CP6, CPz; Parietal, P1, P2, P3, P4, P5, P6, P7, P8, Pz; Parieto-occipital, PO3, PO4, POz; Occipital, O1, O2, Oz; Temporal, T7, T8; Temporo-parietal, TP7, TP8; Posterior channels CB1, CB2 and Iz. The following channels were also placed manually on the face and back of the ears, eye channels, IO1, IO2, LO1, LO2, mastoid, TP9, TP10, frontal, F9, F10, and fronto-temporal, FT9, FT10.

2.5. Data analysis

Data analyses were conducted using EEGLAB (Delorme and Makeig, 2004). Continuous EEG files were high-pass filtered at 0.1 Hz and segmented into 1300 ms epochs including a 500 ms pre-stimulus baseline. All channels were re-referenced to an average reference. Any trial with an amplitude of $\pm 75 \mu\text{V}$ at any channel was rejected. To further minimize effects due to muscle activity trials with an amplitude of $\pm 50 \mu\text{V}$ at channels where muscle activity is most prominent (IO1, LO1, IO2, LO2, FP1, FPz, FP2, FT9, FT10, F7, F8, T3, T4, P7, P8, Iz, CB1, CB2) were rejected. To eliminate effects attributed to button press and due to the small number of target trials (low signal-to-noise ratio), only standard trials were analyzed.

2.5.1. Time-frequency analysis—Time-frequency analyses of single trial data were based on Event-Related Spectral Perturbation (ERSP) and Inter-Trial Phase Coherence (ITPC) techniques, as implemented in EEGLAB (Delorme & Makeig, 2004). ERSP here represents the spectral power difference with respect to time between post-stimulus and the pre-stimulus baseline activity. ERSP is calculated as the log of the ratio of the two activities. This difference is reported in decibel (dB) units. ERSPs contain phase-locked (“evoked”) as well as non-phase-locked (“induced”) oscillatory brain responses on individual trials. The pre-stimulus baseline was limited to the period from -500 ms to -150 ms before sound onset to reduce overlap of pre-stimulus and post-stimulus activity due to windowing. The time-frequency analysis spanned theta, alpha, beta, and lower gamma (30–60 Hz) frequency bands. The frequencies were analyzed in 1.5 Hz increments using a sliding Hanning-windowed 1-cycle sinusoidal wavelet transform of the time-domain signal with a step size of 5 ms. Given the 350ms baseline duration, a 1-cycle wavelet allowed us to examine spectral variation with a floor frequency of 4 Hz (for a 256 ms wavelet). The sliding window was 128 samples (256 ms) long at the lowest frequency and decreased linearly with frequency reaching 64 samples at the highest frequency. To distinguish phase-locked activity from the ERSPs, analysis of phase coherence between individual trials (ITPC) was conducted (Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996). ITPC represents the power of phase-locking between single trials and gives a measure of evoked oscillatory activity. It is noteworthy that phase-locking is a continuous measure, empirically never reaching perfect phase-locking or phase-independence. However, for simplicity we will identify task-related activity as “evoked” if phase-locking is enhanced in the ITPC spectrograms and “induced” otherwise. Time-frequency spectrograms of ERSPs and ITPCs were averaged according to stimulus type and analyzed separately.

Figure 1 illustrates how combining ERSP with ITPC analyses yields clearer interpretations than either alone. It shows the average ERSP (left) and ITPC (right) spectrograms of 10 subjects at channel Cz collapsing across the semantic and voice tasks. From the ERSP plot, we can see several distinct activities that are common to the two tasks, with low- and high-frequency power enhancement represented in red. The low-frequency ($< 20 \text{ Hz}$) power increase spanned a long time-window (0–600 ms), while the high-frequency ($> 25 \text{ Hz}$) power increase was limited to the 30–100 ms window. Power suppression (dark blue) occurred between 15 and 45 Hz and spanned a time window from about 100 to 500 ms. When compared to the ERSP plot, the ITPC shows increased phase-locking (red) for the same spectrottemporal range of ERSP power enhancements (also red) but not for the spectrottemporal range of ERSP power suppression (dark blue). Accordingly, we may conclude that the low- and high-frequency enhancement in

the ERSP plot represents evoked activity while the power suppression represents changes in induced activity. It should be noted that the relationship between the ITPC and ERSP plots must always be interpreted with caution. For example, it is possible that enhanced phase-locking in the ITPC may not be accompanied by an increase of power or may even be accompanied by suppression of power in the ERSP.

2.5.2. Statistical analysis—Permutation tests were used to identify regions that showed significant differences in ERSP or ITPC spectrograms between the semantic and voice tasks (group). Permutation methods do not assume an explicit parametric form for the population distribution, as in parametric methods. Instead, they derive the distribution by resampling the data. For example, under the null hypothesis of no group effect, randomly assigning the group label to the subjects would produce a distribution of observations similar to that of the population (chance) distribution. This distribution is referred to as the null distribution. By comparing the null distribution from resamplings against the observations, one can determine whether to accept the null hypothesis for a given Type I error (Good, 2000). To handle the problem of multiple comparisons in neuroimaging data, permutation tests were applied based on the null distributions of the maximum values obtained in repeated resamplings of the data (Holmes, Blair, Watson, & Ford, 1996). Maximal null distributions were derived from the pre-stimulus period data to improve statistical power (Chau, McIntosh, Robinson, Schulz, & Pantev, 2004).

The following steps were conducted for the permutation tests. 1) The mean ERSP (or ITPC) difference between the semantic and voice tasks was computed. 2) In each resampling step, each subject's data from the semantic or voice task was randomly assigned to either the semantic or voice group. The number of subjects in each group remained unchanged during resampling. 3) For each resampling in step 2, the ERSP or ITPC spectrogram from the pre-stimulus data for all electrodes and time-frequency points was determined. 4) the maximum absolute mean ERSP or ITPC difference between the resampled groups of step (3) was recorded. 5) Following all the resamplings, the recorded maximum absolute differences were pooled together to generate the maximal null distribution. 6) The threshold value for a given p value (e.g. $p = 0.01$ here) from the maximal null distribution was determined. 7) Significant group differences in the mean ERSP or ITPC computed in step (1) based on this threshold were identified.

The number of permutation resamples was restricted by the number of subjects (2^n , where n equals the number of subjects in the comparison; Good, 2000). As such, maximum resampling of $1024 = 2^{10}$ was used for the current contrasts. Time-frequency points exceeding significance denote time-frequency points where spectral power induced by the semantic conditions exceeded that induced by the voice condition (enhancement) or spectral power induced by the voice condition exceeded that induced by the semantic condition (suppression). The threshold of significance was set at $p = 0.01$ for permutation tests.

3. RESULTS

Figure 2A shows the permutation-tested spectrotemporal plots contrasting the semantic to the voice tasks (semantic > voice) for a selected group of channels, which are representative of where the tasks differed most. Supplemental Figure 2 illustrates permutation results at all channels for a more comprehensive account. Also shown is the time course of the various frequency bands based on the permutation-tested temporal dynamics of the semantic > voice contrasts averaged across all channels, revealing the peak-latency of each band (Fig. 2B) and the scalp distribution of these bands at the specified peak-latencies. Notice in the temporal waveforms that the power scale is very small, especially for theta and gamma, due to the more focused scalp distribution of the individual bands. Theta activity (e.g. see FCz channel, arrow),

was enhanced for the semantic task, was mainly localized to the midline fronto-central region (Fig. 2B left panel), and peaked at 200–300 ms. Alpha and lower beta (8–12 and 13–20 Hz, respectively) activities (e.g. see PO3 and PO4 channels) were suppressed for the semantic compared to the voice task. These were distributed across the mid-posterior scalp (Fig. 2B left panel), and spanned an extended time window, with the strongest suppression occurring after 250 ms following speech presentation. Although alpha and beta suppression peaked much later than theta, they commenced about the same time or earlier (see left panel temporal dynamics in figure 2B). Enhanced high-frequency activity spanned the gamma and upper beta bands for the semantic task and occurred bilaterally in small regions of the fronto-central scalp: at channels FC5, C5, C3 in the left hemisphere and F4, F6, FC6 in the right hemisphere (Fig 2A, Fig. 2B right panel). Gamma band activity in the left hemisphere was enhanced, mainly above 30 Hz (see Fig. 2A, FC5, C5), and peaked between 200 and 300 ms. In contrast, GBA above 30 Hz was suppressed at the vertex (Cz, FCz) and right hemisphere (C4), with suppression also occurring in the baseline. However, right hemisphere high-frequency activity enhancement was localized to the 20 to 30 Hz (see Fig. 2A, F4, FC6) band, spanned upper beta and lower gamma band activities and was strongest soon following the onset (95 ms) of the words and at the offset (610 ms) of the words.

There were small but significant differences between the semantic and voice task in the baseline. For example, by examining channels C5 and PO4 of Figure 2, it can be seen that there is gamma band enhancement at C5 and alpha/beta band suppression at PO4 prior to the onset of the stimulus. Activity in the baseline spanned all frequency bands as seen in the temporal waveforms of Figure 2B. These particular baseline activities, as seen in C5 and PO4 were similar in both topography and frequency as the post-stimulus activity, which suggests that they are cortical in origin. However, it should be noted that some GBA in the baseline (around -140 ms, seen for the 26–35 Hz and 36–45 Hz bands) had different scalp distribution than the post-stimulus onset GBA; it was mainly confined to LO1 and CB2 (see supplemental Figure 2), which may reflect residual eye activity unfiltered by artifact rejection.

The results of the spectrotemporal analysis discussed above indicate differences between conditions but do not distinguish whether these differences are phase-locked (“evoked”) or non-phase-locked (“induced”) to sound cues, such as the word onset. To make this distinction, ITPC analysis for each channel for the semantic condition was conducted. Only right hemisphere GBA and theta activity at the midline exhibited phase-locking. For example, the permutation results of the ERSP difference between the semantic and voice conditions at channel F4 and the ITPC mean (not a difference plot) for the semantic condition alone at F4 is shown at Figure 3. If the differences we see in the semantic > voice ERSP comparison of F4 are phase-locked, then we should see enhanced phase-locking in the ITPC plot of F4 occurring at the same frequency/time points. Accordingly, we can conclude that the gamma oscillations occurring at F4 (~ 25–30 Hz) with a latency around 95 ms and also around 430 ms are phase-locked activity.

Similar comparison at Cz revealed that theta oscillations (~ 200–300 ms) were also phase-locked. In contrast, the later gamma (610 ms), seen at F4, F6, FC6, and the beta and alpha activity, seen at posterior electrodes, are not phase-locked and hence are “induced” activity. Although ERSP differences seen for gamma (F4) and theta (Cz) between semantic and voice conditions are phase-locked, this does not necessarily imply that power differences (ERSP) should be accompanied by phase-locking differences (ITPC). Permutation tests conducted on the consistency of phase-locking (ITPC spectrograms) between semantic and voice conditions for all channels indicated that only theta activity exhibited enhanced phase-locking for the semantic compared to the voice conditions. Figure 4 depicts this relationship. Notice the phase-locking effects at Cz in the ITPC plot were localized to the same time and frequency window of the low-frequency activity depicted in the ERSP plot.

4. DISCUSSION

EEG activity exhibited distinct spectral and temporal dynamics during semantic evaluation of spoken words. Evoked theta and induced gamma band activity was accompanied by induced alpha and beta suppression. These findings can be neurophysiological correlates of 1- attributes specific to speech, such as *template matching in auditory and lexical memory*, and also 2- attributes non-specific to speech but essential to speech comprehension such as *selective attention*. In addition to auditory template matching and selective attention, numerous other processes such as wordform detection were surely executed by the listener as well. Such intermediate processes, however, likely occurred in both tasks and therefore should not play a strong role in our between-condition contrasts.

4.1. Template matching in auditory memory

Theta and gamma enhancement may reflect enhanced matching of acoustical cues to representations in auditory and lexical memory (Bastiaansen et al., 2008; Bastiaansen et al., 2005; Hannemann et al., 2007; Schneider, Debener, Oostenveld, & Engel, 2008). Using a lexical decision task, Bastiaansen et al., (2008) revealed an increase in theta power during the retrieval of lexical information in vision and audition. Hannemann et al. (2007) revealed in a degraded word identification task that induced GBA is enhanced at left centro-temporal sites – as in the current findings – when a successful lexically matched word is deemed intelligible. Furthermore, in a lexical decision task (Lutzenberger, Pulvermuller, & Birbaumer, 1994), it was shown that correct identification of the word versus non-words resulted in enhanced 30 Hz gamma band activity. In principle, theta and gamma band activities seen here could reflect template matching or they could be related to lexical-semantic processing in general. For example, although semantic violation in a sentence reflects semantic processing, it does not imply template matching; rather it is a failure of template matching. Hald et al. (2006) revealed that theta activity is enhanced when semantic violation in a sentence is detected. They did not show an increase in GBA; instead, they reported a decrease. Semantic violation requires lexical access and implies failure of template matching in lexical memory. Hence, we are led to believe that GBA here does not necessarily indicate a lexical-semantic process per se; rather it reflects how well the stimulus is matched in memory regardless of stimulus type or modality (Hannemann et al., 2007; Herrmann et al., 2004; Lenz et al., 2007; Shahin et al., 2008). Nonetheless, the matching of stimulus characteristics in memory may be lexically mediated. In contrast to GBA, theta activity may index a general mechanism essential to any lexical-semantic process including or supporting template matching in lexical-memory.

Due to the low spatial resolution of EEG, it is difficult to compute source locations of the various oscillatory activities. However, the theta activity had fronto-central topography consistent with sources originating in the supratemporal plane (Picton et al., 1999), while GBA was maximally exhibited in the left centro-temporal region which could imply sources more lateral in the temporal cortex than the theta sources. The lateral portion of the temporal cortex (belt and parabelt) can be recruited for more complex sounds (e.g. speech or music) (Rauschecker, Tian, & Hauser, 1995; Shahin, Roberts, Miller, McDonald, & Alain, 2007) especially for preferred sounds (Tian, Reser, Durham, Kustov, & Rauschecker, 2001). Furthermore, the temporal dynamics and topography of GBA differed between the hemispheres. First, right GBA mostly peaked shortly following the onset (95 ms) and at the offset (610 ms) of the speech stimuli, suggesting that right hemisphere GBA reflects the “on” and “off” response of the stimuli. The “on” and “off” auditory responses are expected to be phase-locked (Pantev, Eulitz, Hampson, Ross, & Roberts, 1996), which is consistent with the GBA at 95 ms but not the one at 610 ms. However, because its topography and frequency bandwidth highly resembles that of the GBA at 95 ms, it is likely that the GBA at 610 ms is phase-locked as well but the phase-locking was naturally smeared due variability of the

different sound offsets of the different words. At channel F4, larger phase-locked activity around 430 ms also occurred for the semantic compared to the voice condition. This activity exhibited similar topography and frequency bandwidth as the activity at 95 ms and 610 ms, suggesting an intermediate process. Second, GBA was more anterior in the right than the left hemisphere, consistent with the asymmetry of the human primary auditory cortices (Makela, Hamalainen, Hari, & McEvoy, 1994; Penhune, Zatorre, MacDonald, & Evans, 1996; Rademacher et al., 2001; Shahin et al., 2007). Thus GBA here may be auditory in nature, with the left hemisphere GBA reflecting higher-level auditory and lexical mechanisms, specific to word-level processing seen with non-phase-locked activity (Pantev, 1995).

It is worth noting that recent findings suggests that induced gamma band activity may be generated by muscle activity (Whitham et al., 2008; Whitham et al., 2007), and the muscle activity might be correlated with experimental tasks. GBA due to muscle activity is usually broadband, maximally exhibited at electrode locations where muscle activity is most evident such as at the forehead (e.g. FP1, FP2), temporal sites (e.g. T3, T4), and neck (e.g. Iz, CB1, CB2), and symmetrical in topography. Also, it has been reported that induced gamma band activity (Yuval-Greenberg, Tomer, Keren, Nelken, & Deouell, 2008) can be produced by cortical activity associated with miniature saccades. The GBA representing saccade activity is usually limited to the 200–300 ms, spans a wide frequency bandwidth, and has a posterior and symmetrical topography. These characteristics of artifactual induced GBA sources differ from the current findings as well as prior work on induced GBA in audition (Hald et al., 2006; Hannemann et al., 2007; Schneider et al., 2008; Shahin et al., 2008). The current gamma band activity was asymmetrical, confined to a narrow frequency band, and had centro-temporal topography. Furthermore, by examining GBA distribution in supplemental Fig. 2 it can be seen that GBA was absent or least exhibited at sites where eye and muscle activity are largest.

4.2. Selective attention

Selective attention would imply directing attention to one stimulus among others (Fries et al., 2001; Sokolov et al., 2004) or selecting a target based on prior expectations as in the classical P300 design (Donchin & Coles, 1988; Picton, 1992; Polich, 2007). In the current design, subjects selected the word based on its meaning in the semantic task or the pitch in the voice task. The design here included targets ($p = 20\%$) and standards, but the targets, where selective attention is most manifested, were not analyzed due to low signal to noise ratio. Nevertheless, it is highly likely that subjects had to exercise selective attention for the standards as well by deciding whether to ignore or respond. Selective attention has been associated with alpha and beta suppression in vision (Fries et al., 2001; Gomar, Althaus, Wijers, & Minderaa, 2006) and audition during semantic processing tasks (Bastiaansen et al., 2005; Klimesch, Doppelmayr, Pachinger, & Russegger, 1997; Mazaheri & Picton, 2005). Alpha suppression can be interpreted as the reduction of an idling process in a brain network when it becomes active (Pfurtscheller, Stancak, & Neuper, 1996), as in selective attention. Naturally, the selection process is more delayed and spans a longer duration for the semantic than the voice task, because identifying the voice (e.g. pitch) occurs earlier and more rapidly than identifying the meaning. Therefore, we should expect neurophysiological markers, such as alpha and beta oscillations, of selective attention to accompany or follow in time the neurophysiological correlates of matching to prior expectations (template matching). Moreover, these neurophysiological markers should occur later and sustain longer for the semantic compared to the voice task. Indeed, alpha and beta suppression reached their maxima between 500 and 600 ms while gamma and theta activity peaked much earlier (~250 ms). This is consistent with a selective attention process extending through template matching, which, as we propose, is indexed by theta and gamma oscillations.

As we noted earlier, the alpha and beta suppression, although it peaked late, commenced very early (~150 ms) suggesting that processes other than selective attention may also have contributed to semantic evaluation. It is possible that sustained alpha/beta suppression commencing early in the semantic task is related to a memory search or memory scanning demands (Kaufman, Curtis, Wang, & Williamson, 1992; Rojas, Teale, Sheeder, & Reite, 2000), which should precede and overlap with memory match and selection. We reported previously, using the same data set, that reaction times were significantly longer in the semantic than the voice task (Shahin, Alain, & Picton, 2006) consistent with enhanced memory operations (Sternberg, 1969). Also worth noting is that the alpha/beta suppression (e.g. channel PO4) and gamma band enhancement (channel C5) prior to stimulus onset may reflect anticipatory attention (Bastiaansen, Bocker, & Brunia, 2002; Bastiaansen, Bocker, Cluitmans, & Brunia, 1999; Onoda et al., 2007; Widmann, Gruber, Kujala, Tervaniemi, & Schroger, 2007). The presence of anticipatory processes is not surprising here given that the voice and semantic tasks were presented in a blocked design, where subjects knew prior to the stimulus onset which cue (semantic or pitch) they should anticipate. The constant interstimulus interval was another constraining factor that may have allowed the subjects to predict the timing of stimulus presentations.

In addition to its role in template matching, gamma band activity has also been implicated during tasks requiring selective attention (Fries et al., 2001; Snyder & Large, 2005; Sokolov et al., 2004). Findings from animal data showed that neurons oscillating in the gamma band frequency of the visual receptive fields synchronize in phase during selective attention. An increase in gamma band activity is accompanied by a decrease in low frequency oscillations (< 20 Hz). It has been suggested that suppression of low frequencies can enhance postsynaptic efficiency by reducing the co-occurrence of high-frequency (e.g. GBA) spikes within specific time windows governed by low-frequency activity, hence, high frequency adaptation is minimized (Fries et al., 2001).

We should note that beta oscillations here may represent the 2nd harmonic of the alpha band. This is likely, given that the topography of the beta peak around 500 ms resembles that of the alpha scalp distribution. However, suppression of the beta activity also peaked at 250 ms, around the same time as the theta and gamma peaks, and differed in topography from that of alpha. Hence, the earlier beta suppression may indicate an independent process. Even though a button press was not required for the standards, in light of its central and focal topography, early beta can be associated with motor activity preceding a likely response (Kaiser, Birbaumer, & Lutzenberger, 2001; Pfurtscheller & Aranibar, 1977). The semantic task was more demanding than the voice task (Shahin et al., 2006), and hence subjects might have exercised enhanced motor preparation in anticipation to response.

4.3 Conclusion and significance

Enhanced theta and gamma oscillations represent lexically-mediated template matching in auditory memory, and this template matching is supported by top-down selective attention. The current findings provide evidence for the network dynamics of semantic evaluation of speech and offer further testable hypotheses relating to where these activities originate and how different frequency bands might modulate one another.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

This research was funded by the Canadian Institutes of Health Research (CIHR, TWP) and by the National Institutes of Health: National Institute on Deafness and other Communication Disorders (NIH/NIDCD, LMM). Dr. Wilkin Chau provided expert advice.

References

- Bastiaansen MC, Bocker KB, Brunia CH. ERD as an index of anticipatory attention? Effects of stimulus degradation. *Psychophysiology* 2002;39(1):16–28. [PubMed: 12206292]
- Bastiaansen MC, Bocker KB, Cluitmans PJ, Brunia CH. Event-related desynchronization related to the anticipation of a stimulus providing knowledge of results. *Clinical Neurophysiology* 1999;110(2):250–260. [PubMed: 10210614]
- Bastiaansen MC, Oostenveld R, Jensen O, Hagoort P. I see what you mean: theta power increases are involved in the retrieval of lexical semantic information. *Brain and Language* 2008;106(1):15–28. [PubMed: 18262262]
- Bastiaansen MC, van der Linden M, Ter Keurs M, Dijkstra T, Hagoort P. Theta responses are involved in lexical-semantic retrieval during language processing. *Journal of Cognitive Neuroscience* 2005;17(3):530–541. [PubMed: 15814011]
- Braeutigam S, Bailey AJ, Swithenby SJ. Phase-locked gamma band responses to semantic violation stimuli. *Brain Research Cognitive Brain Research* 2001;10(3):365–377. [PubMed: 11167062]
- Chau W, McIntosh AR, Robinson SE, Schulz M, Pantev C. Improving permutation test power for group analysis of spatially filtered MEG data. *Neuroimage* 2004;23(3):983–996. [PubMed: 15528099]
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods* 2004;134(1):9–21. [PubMed: 15102499]
- Donchin E, Coles MGH. Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences* 1988;11:357–374.
- Friederici AD. Event-related brain potential studies in language. *Current Neurology and Neuroscience Reports* 2004;4(6):466–470. [PubMed: 15509448]
- Fries P, Reynolds JH, Rorie AE, Desimone R. Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 2001;291(5508):1560–1563. [PubMed: 11222864]
- Gomarus HK, Althaus M, Wijers AA, Minderaa RB. The effects of memory load and stimulus relevance on the EEG during a visual selective memory search task: an ERP and ERD/ERS study. *Clinical Neurophysiology* 2006;117(4):871–884. [PubMed: 16442346]
- Good, P. *Permutation tests: a practical guide to resampling methods for testing hypotheses*. New York: Springer-Verlag; 2000.
- Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. *Science* 2004;304(5669):438–441. [PubMed: 15031438]
- Hald LA, Bastiaansen MC, Hagoort P. EEG theta and gamma responses to semantic violations in online sentence processing. *Brain and Language* 2006;96(1):90–105. [PubMed: 16083953]
- Hannemann R, Obleser J, Eulitz C. Top-down knowledge supports the retrieval of lexical information from degraded speech. *Brain Research* 2007;1153:134–143. [PubMed: 17451657]
- Herrmann CS, Lenz D, Junge S, Busch NA, Maess B. Memory-matches evoke human gamma-responses. *BMC Neuroscience* 2004;5:13. [PubMed: 15084225]
- Holmes AP, Blair RC, Watson JD, Ford I. Nonparametric analysis of statistic images from functional mapping experiments. *Journal of Cerebral Blood Flow & Metabolism* 1996;16(1):7–22. [PubMed: 8530558]
- Kaiser J, Birbaumer N, Lutzenberger W. Event-related beta desynchronization indicates timing of response selection in a delayed-response paradigm in humans. *Neuroscience Letters* 2001;312(3):149–152. [PubMed: 11602332]
- Kaufman L, Curtis S, Wang JZ, Williamson SJ. Changes in cortical activity when subjects scan memory for tones. *Electroencephalography and Clinical Neurophysiology* 1992;82(4):266–284. [PubMed: 1372548]

- Klimesch W, Doppelmayr M, Pachinger T, Russegger H. Event-related desynchronization in the alpha band and the processing of semantic information. *Brain Research Cognitive Brain Research* 1997;6(2):83–94. [PubMed: 9450602]
- Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. *Science* 1980;207(4427):203–205. [PubMed: 7350657]
- Lenz D, Schadow J, Thaerig S, Busch NA, Herrmann CS. What's that sound? Matches with auditory long-term memory induce gamma activity in human EEG. *International Journal of Psychophysiology* 2007;64(1):31–38. [PubMed: 16959348]
- Lutzenberger W, Pulvermuller F, Birbaumer N. Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neuroscience Letters* 1994;176(1):115–118. [PubMed: 7970225]
- Makela JP, Hamalainen M, Hari R, McEvoy L. Whole-head mapping of middle-latency auditory evoked magnetic fields. *Electroencephalography and Clinical Neurophysiology* 1994;92(5):414–421. [PubMed: 7523085]
- Mazaheri A, Picton TW. EEG spectral dynamics during discrimination of auditory and visual targets. *Brain Research Cognitive Brain Research* 2005;24(1):81–96. [PubMed: 15922161]
- Onoda K, Okamoto Y, Shishida K, Hashizume A, Ueda K, Yamashita H, et al. Anticipation of affective images and event-related desynchronization (ERD) of alpha activity: an MEG study. *Brain Research* 2007;1151:134–141. [PubMed: 17408598]
- Pantev C. Evoked and induced gamma-band activity of the human cortex. *Brain Topography* 1995;7(4):321–330. [PubMed: 7577330]
- Pantev C, Eulitz C, Hampson S, Ross B, Roberts LE. The auditory evoked "off" response: sources and comparison with the "on" and the "sustained" responses. *Ear and Hearing* 1996;17(3):255–265. [PubMed: 8807267]
- Penhune VB, Zatorre RJ, MacDonald JD, Evans AC. Interhemispheric anatomical differences in human primary auditory cortex: probabilistic mapping and volume measurement from magnetic resonance scans. *Cerebral Cortex* 1996;6(5):661–672. [PubMed: 8921202]
- Pfurtscheller G, Aranibar A. Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalography and Clinical Neurophysiology* 1977;42(6):817–826. [PubMed: 67933]
- Pfurtscheller G, Stancak A Jr, Neuper C. Event-related synchronization (ERS) in the alpha band—an electrophysiological correlate of cortical idling: a review. *International Journal of Psychophysiology* 1996;24(1–2):39–46. [PubMed: 8978434]
- Picton TW. The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology* 1992;9(4):456–479. [PubMed: 1464675]
- Picton TW, Alain C, Woods DL, John MS, Scherg M, Valdes-Sosa P, et al. Intracerebral sources of human auditory-evoked potentials. *Audiology and Neurotology* 1999;4(2):64–79. [PubMed: 9892757]
- Polich J. Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology* 2007;118(10):2128–2148. [PubMed: 17573239]
- Pulvermuller F, Eulitz C, Pantev C, Mohr B, Feige B, Lutzenberger W, et al. High-frequency cortical responses reflect lexical processing: an MEG study. *Electroencephalography and Clinical Neurophysiology* 1996;98(1):76–85. [PubMed: 8689998]
- Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, et al. Probabilistic mapping and volume measurement of human primary auditory cortex. *Neuroimage* 2001;13(4):669–683. [PubMed: 11305896]
- Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. *Science* 1995;268(5207):111–114. [PubMed: 7701330]
- Rojas DC, Teale PD, Sheeder JL, Reite ML. Neuromagnetic alpha suppression during an auditory sternberg task. Evidence for a serial, self-terminating search of short-term memory. *Brain Research Cognitive Brain Research* 2000;10(1–2):85–89. [PubMed: 10978695]
- Schneider TR, Debener S, Oostenveld R, Engel AK. Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *Neuroimage* 2008;42(3):1244–1254. [PubMed: 18617422]

- Shahin AJ, Alain C, Picton TW. Scalp topography and intracerebral sources for ERPs recorded during auditory target detection. *Brain Topography* 2006;19(1–2):89–105. [PubMed: 17136467]
- Shahin AJ, Roberts LE, Chau W, Trainor LJ, Miller LM. Music training leads to the development of timbre-specific gamma band activity. *Neuroimage* 2008;41(1):113–122. [PubMed: 18375147]
- Shahin AJ, Roberts LE, Miller LM, McDonald KL, Alain C. Sensitivity of EEG and MEG to the N1 and P2 auditory evoked responses modulated by spectral complexity of sounds. *Brain Topography* 2007;20(2):55–61. [PubMed: 17899352]
- Snyder JS, Large EW. Gamma-band activity reflects the metric structure of rhythmic tone sequences. *Brain Res Cognitive Brain Research* 2005;24(1):117–126.
- Sokolov A, Pavlova M, Lutzenberger W, Birbaumer N. Reciprocal modulation of neuromagnetic induced gamma activity by attention in the human visual and auditory cortex. *Neuroimage* 2004;22(2):521–529. [PubMed: 15193580]
- Sternberg S. Memory-scanning: mental processes revealed by reaction-time experiments. *American Scientist* 1969;57(4):421–457. [PubMed: 5360276]
- Tallon-Baudry C, Bertrand O, Delpuech C, Pernier J. Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *Journal of Neuroscience* 1996;16(13):4240–4249. [PubMed: 8753885]
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP. Functional specialization in rhesus monkey auditory cortex. *Science* 2001;292(5515):290–293. [PubMed: 11303104]
- Van Winsum W, Sergeant J, Geuze R. The functional significance of event-related desynchronization of alpha rhythm in attentional and activating tasks. *Electroencephalography and Clinical Neurophysiology* 1984;58(6):519–524. [PubMed: 6209102]
- Whitham EM, Lewis T, Pope KJ, Fitzgibbon SP, Clark CR, Loveless S, et al. Thinking activates EMG in scalp electrical recordings. *Clinical Neurophysiology* 2008;119(5):1166–1175. [PubMed: 18329954]
- Whitham EM, Pope KJ, Fitzgibbon SP, Lewis T, Clark CR, Loveless S, et al. Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clinical Neurophysiology* 2007;118(8):1877–1888. [PubMed: 17574912]
- Widmann A, Gruber T, Kujala T, Tervaniemi M, Schroger E. Binding symbols and sounds: evidence from event-related oscillatory gamma-band activity. *Cerebral Cortex* 2007;17(11):2696–2702. [PubMed: 17272264]
- Yuval-Greenberg S, Tomer O, Keren AS, Nelken I, Deouell LY. Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron* 2008;58(3):429–441. [PubMed: 18466752]

Fig. 1.

The mean Event-Related Spectral Perturbation (ERSP) and Inter-Trial Phase Coherence (ITPC) at channel Cz for 10 subjects. In the ERSP plot, red indicates enhancement of activity (increase of power) with respect to the pre-stimulus baseline and blue indicates suppression of activity with respect to pre-stimulus baseline. In the ITPC plot, red indicates the power of phase-locking between trials and blue indicates no phase-locking. The term “induced” indicates that the suppressed activity in blue was not phase-locked to stimulus onset and therefore only appears in the ERSP plots. The term “evoked” indicates that the enhanced activity in red was phase-locked to stimulus onset and therefore appears in both ERSP and ITPC plots.

Fig. 2.

A) ERSP plots after permutation tests for selected channels showing the largest significant differences between the semantic and voice tasks. In these plots, non-significant activity was deliberately set to zero (green color). **B)** The temporal dynamics for the mean across all channels after permutation tests of the theta, alpha, beta, and gamma band activity and corresponding scalp distributions at the peaks identified in temporal waveforms. Significance was set at $p = 0.01$.

Fig. 3.

Top) ERSP plot of the semantic vs. voice comparison, after permutation tests for the F4 channel contrasted to **Bottom)** ITPC plot for mean activity of the semantic task across 10 subjects for the same channel. For the ITPC analysis, one sample t-tests indicated that enhancement of the theta activity (4–7 Hz, 200–400 ms), upper beta and gamma activity (30–50 Hz, 80–130 ms and 25–35 Hz, 350–450 ms) were highly significant across the 10 subjects ($p < 0.001$).

Fig. 4.

Top) ERSP plot after permutation tests for the semantic vs. voice comparison at Cz channel contrasted to **Bottom)** ITPC plot after permutation tests for the semantic vs. voice comparison at the same channel. In these plots, non-significant activity was deliberately set to zero (green color for ERSP and dark blue for ITPC). Significance was set at $p = 0.01$.