Task-Related Modulation of Early Cortical Responses during Language Production: An Event-Related Synthetic Aperture Magnetometry Study

We used whole-head magnetoencephalography measurements to investigate the spatiotemporal pattern of neural activity related to language production. Eight participants overtly responded by repeating aloud or vocalizing an internally generated verb to auditorily or visually presented nouns. Activity peaked within primary sensory (auditory or visual) cortices between 75 and 130 ms after stimulus onset, association cortices (inferior and superior temporal gyri) between 130 and 170 ms, and inferior frontal and premotor areas between 150 and 240 ms. Common to auditory and visual modalities, peak activity at about 220 ms was significantly larger in bilateral inferior frontal and left precentral regions when participants generated a verb than when they repeated a noun. These early differences in frontal regions may reflect the allocation of resources to the processing of low-level perceptions that are projected to the premotor areas early in the preparation of language production.

Keywords: auditory, event-related synthetic aperture magnetometry, language production, magnetoencephalography, visual

Introduction

The neural basis of language production is the focus of many neuroimaging and lesion studies (Petersen et al. 1988; Snyder et al. 1995; Price 2000). Early lesion work identified key structures in the temporal parietal (Wernicke’s Area) and frontal (Broca’s Area) regions that are involved in human receptive and expressive language, respectively (Geschwind 1970). Most neuroimaging studies have since confirmed and expanded on the Wernicke-Geschwind Model (Fiez and Petersen 1998; Petersen et al. 1988; Price 2000). These studies showed that the cortical sources generated in language include the sensory-specific cortices (visual or auditory) for low-level perceptual processing; the superior temporal gyrus (STG), angular gyrus, and supramarginal gyrus for language perception; and the inferior frontal and precentral gyri (preCG) for language and speech production.

The sequence of cortical events is important information for understanding how language is perceived and produced. The excellent spatial information we have gained from functional magnetic resonance imaging and positron emission tomography (PET) studies of language functions is however accompanied with poor temporal information. Electroencephalography (EEG) (intra- and extracerebral) and magnetoencephalography (MEG), on the other hand, provide the precise temporal resolution needed to resolve the sequence of neural activity associated with perceptual and cognitive functions (Hamalainen et al. 1993), such as language perception and production. Halgren et al. (1994) identified that evoked waveforms (N150-P190-N220) recorded using intracerebral EEG in the lateral prefrontal cortices were related to nonspecific sensory processing.

Snyder et al. (1995) also showed larger ERPs at about 200 ms in frontal and temporal parietal electrodes when participants internally generated a verb associated with a presented noun (verb generation) than when they repeated the noun (repeat aloud). These ERP effects were linked to larger PET signals in frontal and temporal regions for the verb generation than the repeat aloud task. Localizations of the ERP effects using dipole modeling were however less precise.

In the present study, we collected MEG data and analyzed it using event-related synthetic aperture magnetometry (ERoSAM) (Cheyne et al. 2006) to determine the spatiotemporal brain dynamics during verb generation and repeat aloud tasks. Additionally, we presented nouns in both auditory and visual modalities in order to determine if frontal differences between tasks are a result of task-related activity and if they are modality-specific. Particularly, previous studies investigating these early frontal differences only presented words in the visual modality (Snyder et al. 1995; Abdullaev and Posner 1998), which might lead one to suggest that such early effects between generate and repeat tasks are related to semantic processing. However, the frontal cortices might not receive sufficient perceptual information to perform semantic processing by 200 ms. Alternatively, such early response differences could reflect differences in setting up the language network to perform the cognitive tasks. In this study, we took advantage of the fact that spoken words take a few hundred milliseconds to be voiced; more time than the 200 ms, it takes for the frontal system to differentially respond between generate and repeat tasks (Halgren et al. 1994; Abdullaev and Posner 1998). We therefore compared cortical responses between repeating nouns and generating verbs for auditory and visual stimuli in order to identify if early frontal response differences reflect semantic processing.

Anthony T. Herdman, Elizabeth W. Pang, Volker Ressel, William Gaetz and Douglas Cheyne

1Department of Psychology, Simon Fraser University, Burnaby, Canada, 2Division of Neurology and Department of Paediatrics, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada, 3Department of Pediatric Neurology and Developmental Medicine, Children’s Hospital, University of Tuebingen, Tuebingen, Germany and 4Departments of Diagnostic Imaging and Medical Imaging, The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada

© The Author 2007. Published by Oxford University Press. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org
Materials and Methods

Participants
Eight right-handed volunteers (6 males) between 20 and 42 years of age with normal neurological, visual, and audiological status participated in this study. Informed consent was obtained from each participant after the nature of the experiment was fully explained. The study involved a 60-min session at the MEG facilities in Toronto’s Hospital for Sick Children. The Hospital for Sick Children’s Ethics Review Board approved this study.

Participant’s Task
Participants performed 4 tasks: 1) “Listen Repeat,” listen to the aurally presented noun and repeat it aloud; 2) “Listen Generate,” listen to the aurally presented noun (e.g., ball) and vocalize an internally generated verb (e.g., throw) associated with the noun; 3) “Read Repeat,” read the visually presented noun and repeat it aloud; and 4) “Read Generate,” read the visually presented noun and vocalize an internally generated verb associated with the noun. Tasks were presented as separate blocks in a random order across participants. During each block, participants were asked to verbally respond as quickly and as accurately as possible. Subject compliance was monitored using a video and intercom system.

Stimuli
Auditory stimuli were 90 prerecorded concrete monosyllabic nouns with high familiarity and high frequency, ranging in length from 3 to 5 letters, spoken in English by a native English male speaker. These stimuli ranged in duration from 350 to 550 ms. Auditory stimuli were presented to the participants through insert earphones at an average intensity of 80 dB sound pressure level as measured using a sound level meter at the end of the insert earphone’s tube. The stimuli were presented at an intertrial interval (onset-to-onset) of 3000 ms.

Visual stimuli were the same 90 concrete nouns as the auditory stimuli. They were back-projected onto a screen 70 cm from a participant’s nasion as white letters on black background that subtended 5–8 degrees of visual angle. Visual stimuli were presented for 500 ms with an intertrial interval of 3000 ms.

Measurement
Magnetic fields were recorded in a magnetically shielded room using a 151-channel whole-head CTF MEG system (VSM MedTech Ltd., Coquitlam, Canada) located at Toronto’s Hospital for Sick Children. Sensors were configured as first-order axial gradiometers. The system is configured as an array of first-order axial gradiometers using third-order synthetic gradient noise cancellation with a resulting noise level of approximately 10 fT/√Hz above 1 Hz. For each task block, MEG was continuously sampled at 625 samples/s with an online bandpass of 0–200 Hz.

Participants laid comfortably on their backs with their eyes open for all tasks. Video monitoring verified that they remained alert while performing the task. Each subject was fitted with 3 fiducial localization coils placed at the nasion and preauricular points in order to localize the position of the subject’s head relative to the MEG sensors. The fiducial localization coils were replaced with magnetic resonance (MR)-compatible vitamin E capsules to aid the coregistration of MEG data to each subject’s structural magnetic resonance image (MRI). The head position relative to the MEG sensors was monitored at the start and end of each block. Motion tolerance was <0.5 cm, but this was not exceeded by any subject on any block. $T_1$-weighted structural MR images (3-dimensional [3D] spoiled gradient-recalled sequence) were also obtained for each subject using a 1.5-T Signa Advantage system (GE Medical Systems, Milwaukee, WI) for the purpose of overlaying source images on each subject’s 3D structural MRI.

Analyses
To estimate the evoked activity within the brain, we used the erSAM method previously described by Cheyne et al. (2006). We calculated weighting factors based on the single-state, pseudo Z-synthetic aperture magnetometry (SAM) spatial filter for single-trial epochs of ~500 to 2000 ms, filtered between 0 and 30 Hz. These weighting factors were then applied to the average of the single-trial data from all 151 MEG sensors to obtain virtual channel waveforms throughout the whole brain for each task. Each participant’s pseudo ZSAM map (referred to as erSAM maps) for each poststimulus sample point was transformed into common anatomical space using a transformation matrix calculated from transforming each participant’s structural MRI into common Montreal Neurological Institute anatomical space using SPM99 (Statistical Parameter Mapping; Wellcome Trust Functional Imaging Laboratory 1999). These spatially normalized erSAM maps were then averaged across participants and projected onto MRI slices of SPM99’s template brain.

In order to determine a significance threshold for erSAM activity, we estimated a null distribution for the erSAM noise by permutation analyses (Chau et al. 2004) of the absolute pseudo $Z$ values for all voxels at 0, 4, and 8 ms. The 99.9% largest value in the pseudo $Z$ null distribution was 0.95 and was considered the statistical threshold for erSAM maps. Thresholded erSAM maps were averaged across participants and then projected onto the surface of a 3D structural MRI for visualization.

In order to reduce the erSAM data, we focused on the latencies of the maxima within the global field power of the group-averaged evoked response that occurred before participants’ verbal responses (i.e., maxima between 0 and 300 ms). Latencies of field power maxima occurred at 76, 92, 132, 168, 208, 216, 236, and 264 ms (see Fig. 1). Thresholded erSAM maps at each of these latencies were then used to determine the spatial locations of the erSAM maxima with the criteria that erSAM peaks were spatially separated by a minimum distance of 1 cm. We then extracted the virtual channel waveforms at the location for each group-averaged erSAM maxima for all participants across the 4 tasks. For each virtual channel, we obtained 99.9% confidence limits for amplitude differences between the listening tasks (Listen Generate minus Listen Repeat) and the reading tasks (Read Generate minus Read Repeat) from null distributions that were calculated by bootstrap sampling 1024 times the difference waveform amplitudes between -200 and 0 ms.

Results
The brain regions where pseudo $Z$ values exceeded the statistical permutation threshold for preverbal response maxima are shown in Figure 2 for Listen Generate and Listen Repeat tasks and in Figure 3 for Read Generate and Read Repeat tasks. For the listening tasks, erSAM maps showed that cortical activation begins within the transverse temporal gyri (TTG) between 76 and 132 ms, followed by a spread of activity to the posterior superior temporal gyrus (pSTG), insula/frontal operculum (Ins/FO), and precuneus between 132 and 264 ms. ErSAM maps

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Grand-mean global field power for Listen (top traces) and Read (bottom traces) conditions by Generate (thick lines) and Repeat (thin lines) tasks. Arrowheads designate peaks in the global field power prior to verbal response onset at 92, 132, and 236 ms for the Listen condition and at 76, 168, 208, 216, and 264 ms for the Read condition. Time zero designates stimulus onset.
revealed larger pseudo Z values for the Listen Generate than the Listen Repeat task within the pSTG and FO/Ins between 208 and 216 ms. For the reading tasks, erSAM maps showed that cortical activation begins within the cuneus between 76 and 132 ms, followed by activation within the pSTG, fusiform gyrus, right superior parietal lobule (SPL), Ins/FO, and preCG between 168 and 264 ms. The main observable differences between erSAM maps for Read Generate and Read Repeat are larger
pseudo Z maxima in the Ins/FO, pSTG, and preCG during Read Generate.

Virtual channel waveforms generally showed early peak activity within sensory cortices followed by response peaks within superior temporal and frontal regions for all Read and Listen conditions. For the Listen conditions (Fig. 4), pseudo Z values peaked within bilateral TTG and STG around 130 ms, followed by peaks within Ins/FO and preCG around 150 ms, and
then by peaks within bilateral STG, Ins/FO, and preCG at about 230 ms. All virtual channels, except the cuneus, had sustained pseudo Z values after 250 ms. Difference waveforms (Listen Generate minus Listen Repeat) revealed significantly larger pseudo Z values in bilateral TTG between 100 and 130 ms for the Listen Repeat than Listen Generate. Conversely, larger values occurred within bilateral STG, Ins/FO, and preCG between 190 and 280 ms for Listen Generate than Listen Repeat and between 400 and 600 ms within bilateral TTG.

For the Read conditions (Fig. 5), pseudo Z values peaked within bilateral middle occipital gyrus (MOG), and lingual gyri between 120 and 150 ms, followed by peaks within left fusiform, left STG, and right SPL at about 170 ms. Pseudo Z values then peaked in bilateral Ins/FO and preCG between 200 and 250 ms. Frontal and occipital virtual channels had sustained pseudo Z values greater than baseline or long-duration response peaks between 250 and 600 ms. Larger pseudo Z values occurred for Read Generate than Read Repeat between 100 and 130 ms within bilateral MOG, left STG, and right SPL; between 190 and 210 ms within right lingual and SPL and bilateral Ins/FO and preCG; and between 390 and 500 ms within bilateral Ins/FO and right MOG. Conversely, larger pseudo Z values occurred within the left STG at between 185 and 205 ms.

Figure 6A shows that during Listen and Read conditions pseudo Z values between 180 and 240 ms in the Ins/FO are larger for Generate than Repeat tasks; Pseudo Z values are also larger for the Generate than Repeat task at 192 and 180 ms for Listen and Read conditions. Larger pseudo Z values for the Generate than Repeat task are evident in the Ins/FO for both Listen and Repeat conditions. Additionally, the inferior frontal cortex has larger pseudo Z values for Listen Generate than Listen Repeat, whereas the posterior inferior temporal cortex has larger pseudo Z values for Read Repeat than Read Generate.

Discussion

Results from this study show that auditory and visual word presentations during verb generation and repeat aloud tasks produce large activations within sensory cortices immediately followed by activation of the pSTG and frontal cortices. These brain regions are consistent with classical language areas (Geschwind 1970; Price 2000). Moreover, we found a fast sequence of cortical activations among these areas. Activity peaked within primary sensory (auditory or visual) cortices between 75 and 130 ms after stimulus onset, association cortices (inferior temporal gyrus and STG) between 130 and 170 ms, and inferior frontal and premotor areas between 150 and 240 ms. This fast transition of activity from sensory to association to frontal systems suggests that a fast feed-forward route exists between sensory and frontal networks. This is consistent with previous findings in the visual system for word perception (Halgren et al. 1994), an intermodal selective attention task (Foxe and Simpson 2002), and in the auditory system for a mismatch negativity language perception task (Pulvermüller et al. 2005). These routes likely follow known...
Figure 5. Same as in Figure 4, except that Read Generate is designated by black thick lines and Read Repeat by black thin lines.

Figure 6. (A) Comparison of grand-mean Ins/FO virtual channel waveforms (top and middle) for the Listen and Read conditions. Difference waveforms (Generate minus Repeat) are shown in the bottom traces for Listen (thick black line) and Read (thin black line) conditions. Time zero designates stimulus onset. (B) Pseudo Z value difference maps (Generate minus Repeat) are shown at 192 and 180 ms for Listen and Read conditions, respectively.
anatomical connectivity between auditory frontal and visual frontal cortices (Pandya and Knypers 1969; Deacon 1999). We believe that this feed-forward system projects sensory information to frontal regions in order to prepare them for subsequent cognitive processing, such as internally generating a verb. This hypothesis is consistent with Foxe and Simpson’s (2002) alternative interpretation of their data suggesting that a rapid input to the frontal cortex could serve as an alerting function to prepare the frontal cortex for subsequently slower computations. Results from our study support such a feed-forward mechanism in that the peak activity in the Ins/FO at approximately 200 ms is at least 150 ms before sufficient auditory information was obtained (shortest auditory stimulus duration was 350 ms) in order to perform the cognitive process of internally generating a verb. We cannot however rule out that this early activity might be the beginning of semantic retrieval because there is sufficient time for perception of the first phoneme to reach the frontal cortex. However, we can be confident that this early frontal activity does not index semantic word retrieval because there is insufficient perceptual information entering the auditory system about the whole word by this time.

We found a task-related modulation of this fast feed-forward activity in that there was larger frontal activity for the Generation than Repeat task for both auditory and visual stimuli. This is consistent with ERP results showing larger frontal ERPs at 200 ms for verb generation than repeat aloud task (Snyder et al. 1995). Although they localized this effect to the anterior cingulate, we suggest that the task-related differences in our bilateral frontal sources reflect an interaction between lateral frontal sources and the anterior cingulate source for allocating attentional resources during the Generate and Repeat tasks. Thus, the greater activity in the Ins/FO for the Generate than Repeat task at about 200 ms could be explained as a difference in attention. Because this frontal region is known to be involved in verb generation (Geschwind 1970), more attention is likely to be given to incoming sensory information during the Generate than Repeat task. During the Repeat task, more attention is likely given to the phonemic information of sensory input and thus greater activity would occur in and around the auditory and association (e.g., pSTG) cortices, as our results show (Fig. 4). Moreover, the later larger responses for Generate than Repeat (between 400 and 600 ms) could reflect the cognitive processing differences between the tasks rather than differences in attention because sufficient perceptual information about the word has reached the frontal cortices by this time to allow for verb generation.

Additionally, our findings that show greater responses at about 220 ms in the precentral gyrus during the Generate than Repeat task could indicate a similar type of early attentional modulation of incoming perceptual information to that in the Ins/FO. Response amplification for the Generate condition could also be caused by mirror neuron activity associated with observed or conceptual processing of actions (Pulvermuller et al. 1999; Aziz-Zadeh et al. 2006). We are unable to conclude from our results if this is the case. However, the premotor response difference occurs before semantic word processing is complete. Thus, we believe that this difference is more likely a result of attentional modulation than mirror neuron activation associated with generating the action word. Future studies that compare nouns and verbs for Generate and Repeat tasks would help assess these possibilities.

In this study, we also showed that the greater early frontal activity for the Generate than Repeat task is not modality-specific. The fact that difference waveforms in the Ins/FO (Fig. 6A) showed very similar patterns between aurally and visually presented nouns provides further evidence that there is an underlying task-related process that is not specific to either stimulus modality. Top–down/executive processes, for example attention, could drive such task-related activity and prepare the frontal system for incoming information from either of the visual or of the auditory senses. Thus, the top–down mechanisms reconfigure the neural networks to process the bottom-up (sensory) information in different ways in order to maximize performance on the task at hand. Furthermore, the peak latencies of the difference responses to auditory or visual stimuli were similar. This indicates that modulation of sensory information occurs on a similar time frame regardless of stimulus modality.

Conclusion
Our results show that the timing of early evoked activity from sensory-to-frontal systems during verb generation and repeat aloud tasks occurs within 250 ms. We showed that early differences in frontal regions might reflect the allocation of attentional resources to processing low-level perceptions, which are projected to the premotor areas early in preparation for language production. These findings suggest that a fast feed-forward, modality-unspecific, sensory-to-motor processing exists within the neural network used for expressive language tasks.

Notes
Grants from the Canadian Institutes of Health Research and the Hospital for Sick Children supported this research. Conflict of Interest: None declared.

Address correspondence to Dr Elizabeth W. Pang, PhD, Division of Neurology and Department of Paediatrics, The Hospital for Sick Children, University of Toronto, 555 University Avenue, Toronto, Ontario M5G 1X8, Canada. Email: elizabeth.pang@utoronto.ca.

References

Wellcome Trust Functional Imaging Laboratory. 1999. Statistical parameter mapping [Internet]. Available from: http://www.fil.ion.ucl.ac.uk/spm/.