



HHS Public Access

Author manuscript

Lang Cogn Neurosci. Author manuscript; available in PMC 2017 January 01.

Published in final edited form as:

Lang Cogn Neurosci. 2016 ; 31(7): 841–855. doi:10.1080/23273798.2015.1029498.

Sentential influences on acoustic-phonetic processing: A Granger causality analysis of multimodal imaging data

David W. Gow Jr.^{1,2,3,4,*} and Bruna B. Olson^{1,3}

¹Neuropsychology Laboratory, Massachusetts General Hospital, 175 Cambridge St., CPZ S340, Boston, MA 02114

²Department of Psychology, Salem State University, 352 Lafayette St., Salem, MA 01970

³Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, 149 Thirteenth St., S2301, Charlestown, MA.02129

⁴Harvard-MIT Division of Health Sciences and Technology, 77 Massachusetts Ave., E25-519, Cambridge, MA 02139

Abstract

Sentential context influences the way that listeners identify phonetically ambiguous or perceptual degraded speech sounds. Unfortunately, inherent inferential limitations on the interpretation of behavioral or BOLD imaging results make it unclear whether context influences perceptual processing directly, or acts at a post-perceptual decision stage. In this paper, we use Kalman-filter enabled Granger causation analysis of MR-constrained MEG/EEG data to distinguish between these possibilities. Using a retrospective probe verification task, we found that sentential context strongly affected the interpretation of words with ambiguous initial voicing (e.g. DUSK-TUSK). This behavioral context effect coincided with increased influence by brain regions associated with lexical representation on regions associated with acoustic-phonetic processing. These results support an interactive view of sentence context effects on speech perception.

Keywords

speech perception; semantic context effect; sentential context effect; modularity; top-down processing

Language is famously ambiguous. We see this at all levels of representation in phenomena including the lack of invariance problem in the mapping between speech sounds and phonetic categories, in the interpretation of homophones, in syntactic surface forms that have multiple potential structural mappings, and in the distinction between literal and figurative language. Paradoxically, we generally experience language as being relatively unambiguous. This ability may be largely attributable to the mutual constraint between

Corresponding author: David W. Gow Jr., Neuropsychology Laboratory, Massachusetts General Hospital, CPZ S340, 175 Cambridge St., Boston, MA 02114, Tel: 617-726-6143, gow@helix.mgh.harvard.edu.

Second author: Bruna B. Olson, Neuropsychology Laboratory, Massachusetts General Hospital, CPZ S340, 175 Cambridge St., Boston, MA 02114, Tel: 617-724-8846, bbolson@partners.org

The authors and their institutions have no conflicts of interest related to this work.

different levels of representation. Sounds must make sense in the context of words. Words must make sense in the context of sentences. Sentences must make sense in the context of discourse. In this paper, we explore the question of how sentential context influences the perception of speech sounds. Using Granger causality analysis of MR-constrained MEG/EEG data, we will ask whether this influence reflects post-perceptual selection (feedforward processing) or top-down influences on acoustic-phonetic representation (interactive processing).

A large body of behavioral work dating back to the 1950's demonstrates that sentential context influences the perception of ambiguous, or perceptually degraded speech (c.f. Abada, Baum, & Titone, 2008; Boothroyd & Nittrouer, 1988; Borsky, Shapiro, & Tuller, 2000; Borsky, Tuller, & Shapiro, 1998; Connine, 1987; Connine, Blasko, & Hall, 1991; Kalikow, Stevens, & Elliot, 1977; G. A. Miller, Heise, & Litchen, 1951; G. A. Miller & Isard, 1963; J. L. Miller, Green, & Schermer, 1984). In this work, target words were either made phonetically ambiguous through voice onset time manipulations (e.g. GOAL-COAL), or were presented in noise and embedded at the ends of lexically constraining sentences. Listeners heard these sentences, and then were typically asked to make either a 2AFC phonemic (/g/ versus /k/) or lexical (*goal* versus *coal*) judgment about the final word. In all cases, subjects showed a strong tendency to interpret ambiguous or degraded speech in a way that produced an meaningful, well-formed sentence. These effects appear to be the result of both semantic (G. A. Miller et al., 1951) and syntactic (G. A. Miller & Isard, 1963) constraints. This work clearly shows that perceptual and higher-level contextual representations interact during spoken language perception. However, it does not resolve *how* these representations interact.

Two broad functional architectures have been proposed to explain how representations interact. Feedforward models (c.f. Fodor, 1983; Massaro, 1989; Norris, McQueen, & Cutler, 2000) frame spoken language as a purely bottom-up process in which perceptual or phonological processes are not directly influenced by higher-level representations. In such a system, the output of speech analyses becomes the input to lexical and then sentence level representations. If the output of these analyses is ambiguous, higher-level representations may be used to select the most contextually appropriate representation. In contrast, models that embrace top-down or interactive mechanisms (c.f. Gagnepain, Henson, & Davis, 2012; Hannagan, Magnuson, & Grainger, 2013; Marslen-Wilson & Tyler, 1980; McClelland & Elman, 1986) allow higher level representations to directly influence speech processing.

While strong claims are frequently made about the interactive or feedforward nature of language processing, it is extremely difficult to distinguish between these accounts empirically. Behavioral data are at a disadvantage because they typically depend on overt judgments performed after either interaction or selection has taken place. It remains unclear then whether any interaction takes place at a perceptual level, prior to overt judgment, or decision level after perceptual processes have been completed (Dennett, 1992; Gow & Caplan, 2012; Norris et al., 2000).

Connine (1987) addressed these limitations by comparing the patterning of sentential context effects on speech categorization with the patterning found in independent studies

that explicitly depended on post-perceptual bias. She found that sentential context influenced the categorization of voice onset time continua of word initial stop consonants (e.g. a TENT-DENT continuum). Sentential context had the strongest influence on identification for (phonetically ambiguous) boundary tokens. This result is widely replicated across different experimental paradigms and using both phonetically ambiguous and perceptual degraded speech (c.f. Abada et al., 2008; Boothroyd & Nittrouer, 1988; J. L. Miller et al., 1984; Obleser, Wise, Dresner, & Scott, 2007; Zekveld, Heslenfeld, Festen, & Schoonhoven, 2006). However, analyses of reaction time data showed an effect of context at the endpoints of the continua, but not for boundary tokens. This pattern matches one found for voice onset continua in nonword minimal pairs in a categorization task in which post-perceptual response bias was manipulated through a monetary incentive (Connine & Clifton, 1987). Connine argued that the parallel between these results suggests that sentential influences on speech perception are also the result of post-perceptual response biases. However, other explanations are available. Under an interactive reading, the reaction time effect might reflect processing devoted to semantic integration. A perceptually unambiguous, but contextually inappropriate word might take longer to integrate, resulting in detectable context effects only at the endpoints of continua.

There are several problems with this interpretation. First, it is notable that the reaction time effects involved (endpoint) tokens that were not subject to sentential influences on phoneme categorization. Also, a subsequent study by Borsky and colleagues (1998) failed to replicate the earlier reaction time effects. They found a significant reaction time effect for boundary items that were sensitive to sentential context, but not for endpoint tokens that were not. It is not clear why the two studies found different effects, but it is significant that the critical comparison is between a null reaction time effect for boundary items in the Connine (1987) study, and a significant effect for the same items by Borsky et al. (1998). These factors undermine one argument for a feedforward explanation of sentential context effects, but do not provide direct evidence in favor of interactivity. Available behavioural results simply fail to resolve the debate.

Given the challenges of discriminating between interactive and feedforward processing using behavioral techniques, a number of researchers have turned to BOLD imaging methods to determine whether these effects depend on brain structures associated with acoustic-phonetic processing, or perhaps higher lexical or decision making processes. Several studies have compared BOLD activation patterns for the recognition of perceptually degraded or phonetically ambiguous words in sentences that provide different levels of contextual constraint or predictability. Most of these studies have involved the collection of fMRI data while participants listened to sentences in which the predictability of a target word was manipulated, and intelligibility was manipulated parametrically using noise vocoding techniques. In some studies the task was attentive listening (Obleser & Kotz, 2010; Obleser et al., 2007), while in other work subjects were asked to determine retrospectively whether a particular word occurred in the sentence (Davis, Ford, Kherif, & Johnsrude, 2011). A study by Guediche and colleagues (Guediche, Salvata, & Blumstein, 2013) took a different approach, manipulating word initial VOT in otherwise clear speech to create potential lexical ambiguity as in the behavioral studies of Connine (1987) and Borsky et al. (2000). Consistent with the behavioral literature, the studies of Davis et al. (2011) and

Guediche et al. (2013) showed the strongest sentential context effects on phonemic interpretation for speech that was either phonetically ambiguous or moderately degraded by noise vocoding. All of these studies show effects of sentence predictability on frontal activation, with speech intelligibility modulating superior temporal activation. Left parietal regions including the angular gyrus and supramarginal gyrus were specifically implicated in predictability manipulations for those stimuli that showed the strongest context effects.

Obleser et al. (2007) suggest that this pattern of activation, and an observed increase in correlated activity in the parietal and frontal lobes related to predictability effects, reflect a role of attention and memory processes in the effortful processing of speech. Obleser and Kotz (2010) argue that these results reflect an interactive mechanism in which successful bottom-up perceptual processing in the superior temporal regions regulates activity in the frontal and parietal lobes, which is mirrored by context driven top-down regulation of superior temporal activity by regions involved in semantic analysis. This suggests that contextual constraints influence the strategic allocation of attention to acoustic-phonetic processing.

Guediche et al. (2013) examined these issues in the context of more naturalistic stimuli that required less effortful processing than the perception of noise vocoded speech. In their study, listeners were asked to identify the final word in sentences ending in a token taken from a COAT to GOAT voice onset time phonetic continuum. Despite the reduced processing load incurred by these stimuli, they found an interaction between phonetic ambiguity and sentential constraint involving the posterior superior temporal gyrus (pSTG) and middle temporal gyrus (MTG). The STG is implicated in acoustic-phonetic processing, and the MTG is implicated in lexico-semantic processing by a convergence of BOLD imaging and neuropsychological results (cf. Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Gow, 2012; Hickok & Poeppel, 2007; Price, 2012). They found greater activation in the left pSTG and a second, more anterior cluster involving left MTG extending into STG in neutral as compared to biased sentence contexts for unambiguous speech tokens. This context effect was reversed for phonetically ambiguous items. Guediche et al. interpret this interaction between context and ambiguity as evidence that lexically-mediated sentential context shapes phonetic processing. Like the previous BOLD studies, perceptual ambiguity was related to an increase in activation in frontal areas associated with effortful linguistic retrieval and response selection. This is consistent with the idea that frontal regions regulate phonetic processing.

None of these imaging results are uniquely consistent with interactive processing. Increased activation in the pSTG in degraded or phonetically ambiguous tokens could reflect a phonetic locus for the application of sentential constraints. However it could also reflect the need for additional processing of hard to resolve speech tokens, which also happen to be the same tokens whose interpretation is most strongly influenced by context. Similarly, frontal and temporal activation may reflect increased word and sentence level processing at critical points in sentence processing. Within a feedforward processing architecture where the output of perceptual and sentential analyses is integrated at a post-perceptual decision level increased sentential context may decrease the need for exhaustive perceptual analysis. The combination of neutral sentence contexts and perceptual ambiguity creates a special

challenge, because it leaves the listener with no principled way of interpreting what they hear. In these cases, listeners may devote fewer resources to both perceptual and higher level analysis, moving the focus of processing to higher level response selection. Under this reading, frontal activation seen in these studies would reflect response selection and competition rather than top-down regulation of phonetic processing.

The idea that the same set of BOLD results is interpretable in either a feedforward or interactive processing framework is not unique to these studies. Low temporal resolution spatial patterns of BOLD activation are fundamentally unable to support strong inferences about functional architecture, in part because functional architecture is a description of causal structure, and causation is an inherently temporal phenomenon (Gow & Caplan, 2012; Gow, Segawa, Ahlfors, & Lin, 2008).

Davis et al. (2011) addressed this limitation in part through the use of time-resolved fMRI data. They found that when listening to noise vocoded tokens of meaningful sentences, the posterior superior temporal gyrus, associated with acoustic-phonetic processing, is activated before frontal regions associated with sentential context processing, even when prior sentential context should constrain speech interpretation. Based on these results, they concluded that sentential context effects on speech perception occur within a feedforward framework. However, a subsequent study using higher temporal resolution MEG and EEG data showed that inferior frontal activation may precede superior temporal activation when a strong non-sentential manipulation (cross-modal repetition) is used to predict the identity of a noise vocoded word. (Sohoglu, Peelle, Carlyon, & Davis, 2012). These results are consistent with those of a number of electrophysiological studies that show early perceptual components including the N1 and N200 are sensitive to semantic congruency or predictability manipulations in spoken sentence processing (c.f. Diaz & Swaab, 2007; Groppe et al., 2010; Sivonen, Maess, Lattner, & Friederici, 2006; van den Brink, Brown, & Hagoort, 2001). In all of these cases, evidence that contextual processing occurs before phonetic processing might be considered to support an interactive processing framework (*post hoc ergo propter hoc*). However, while precedence is a necessary condition for causation, it does not necessary establish causation. Listeners may detect a semantic incongruence during a perceptual component without using this information to drive perceptual processing. Furthermore, sensor space analyses of scalp electrode data may be unable to differentiate temporally overlapping signals from independent localized processors. As a result, evoked contextual processes may influence the amplitude or latency of components associated with perceptual processing without directly influencing those processes.

In summary, while it is clear that sentential context influences the intelligibility of degraded speech and the interpretation of ambiguous speech, existing data do not support strong inferences about whether these influences are due to interactive or feedforward functional architectures. Data from behavioral, BOLD imaging and electrophysiological studies have often been contradictory, and the results of individual studies have been consistent with both feedforward and interactive mechanisms.

We have adopted an alternate strategy for examining functional architecture that relies on effective connectivity analyses of high spatiotemporal resolution MR-constrained source space reconstructions of MEG and EEG data. Our approach relies on a variant of multivariate Granger causality analysis (Geweke, 1982; Granger, 1969; Milde et al., 2010) that allows us to infer patterns of directed influence between brain regions. This approach can be performed over large sets of regions of interest (ROIs), and is entirely data-driven. It is built on the inference that the strongest predictions about effects are made based on the unique information carried by their causes. This is operationalized by making predictive models for the upcoming activation of one brain region based on activation timeseries data collected from all non-redundant, potentially causal brain regions including the predicted region. Granger causality is computed by comparing the error associated with models based on all measured brain regions with that associated with models created by eliminating timeseries information from a single region. To the extent that removing information from one brain region increases prediction error, that region is said to Granger cause changes in activity in region whose activity is being predicted. To the degree that localized activation can be assigned a functional interpretation based on evidence from BOLD imaging, pathonormal inference and other functional localization techniques, we can use these analyses to discriminate between top-down and bottom-up processing architectures (Gow & Caplan, 2012). Using this approach, we have identified functional architectures associated with various phenomena including lexical influences on speech categorization (Gow et al., 2008), motor mediation of perceptual compensation for coarticulation (Gow & Segawa, 2009), and phonotactic influences on speech perception (Gow & Nied, 2014).

We collected simultaneous MEG and EEG data while participants listened to meaningful sentences and performed a retrospective phoneme probe verification task (Gow & Gordon, 1993). In experimental trials, word-initial stop consonants were digitally manipulated to create phonetically ambiguous phonemes that were potentially consistent with alternate lexical interpretations (e.g. DUSK-TUSK). The same tokens appeared in alternate sentence contexts that biased interpretation towards voiced (*dusk*) or unvoiced (*tusk*) interpretations. This task produced a robust sentential influence on phoneme categorization.

The purpose of this work is to determine whether sentential context directly influences acoustic-phonetic processing. Accordingly, our predictions focus on interactions involving the pSTG, the hypothesized hub of acoustic-phonetic processing (cf. Chang et al., 2010; Dronkers et al., 2004; Gow, 2012; Hickok & Poeppel, 2007; Mesgarani, Cheung, Johnson, & Chang, 2014; Price, 2012). Our first prediction was that our data-driven algorithm for identifying ROIs (described in the methods section) would identify the pSTG as a component of the processing network. We hypothesized that sentential context might constrain lexical candidates, which in turn might influence acoustic-phonetic processing. Thus, we predicted that if sentential context influences phonetic representation or processing, we should see causal influence by areas associated with lexical processes on activation of the pSTG during word recognition. We predicted that these influences would be absent in a feedforward system. Instead, we predicted that a feedforward system would show feedforward connections from the same regions converging on frontal regions associated with decision-making and response selection including the middle frontal gyrus and cingulate. Our predictions were based primarily on current dual stream neuroanatomical

models of speech processing (Gow, 2012; Hickok & Poeppel, 2004, 2007) and a model of the semantic system based on a metanalysis by Binder and colleagues (Binder, Desai, Graves, & Conant, 2009).

Materials and Methods

Participants

Fourteen right-handed people (7 female) with a mean age of 21.8 years participated in this study. All were native English speakers who reported no auditory processing deficits. All provided written informed consent to participate under a protocol approved by the Partners Healthcare Institutional Review Board.

Materials

The stimuli are illustrated in Table 1. They consisted of a total of 674 short (5–6 word) meaningful sentences. The sentences were digitally recorded in a quiet room with 16-bit sound at a sampling rate of 44.1 kHz. These included 54 pairs of experimental items (108 sentences) in which a final monosyllabic word began with a stop consonant (/t/, /d/, /p/ or /b/) with a minimal pair that differed only in voicing (e.g. TUSK-DUSK) (see Appendix). Sentences were constructed to make one member of the pair contextually appropriate and the other inappropriate. The voiceless member of the pair was excised from an unused sentence token, and made perceptually ambiguous using an automated procedure that identified the offset of stop release bursts and the onset of voicing, and cut out segments of this interval at ascending zero-crossings to create a voice onset time of 20 ms. This value is at the low end of typical estimates of boundary VOT for stop voicing contrasts, but was judged by two phonetically trained listeners to produce the greatest subjective voicing ambiguity in our test materials for word-initial labial and coronal oral stops. This same ambiguous token (^d*usk*) was then used to replace both the sentence final voiced (*dusk*) and unvoiced (*tusk*) tokens in their original constrained sentence contexts, with all splices made at ascending zero crossings to minimize acoustic discontinuity. These stimuli were selected after behavioral piloting (beginning with 164 sentence pairs) established that listeners interpreted all ambiguous tokens in a contextually appropriate manner at least 90% of the time in both voiced and unvoiced contexts. In addition, 108 baseline items were created in which a final monosyllabic word began with the same voiced and unvoiced segments in equal proportion. These items were phonetically unambiguous, and thus predicted to reflect the weakest influence of potential sentential context based on prior behavioral results (c.f. Abada et al., 2008; Boothroyd & Nittrouer, 1988; J. L. Miller et al., 1984; Obleser et al., 2007; Zekveld et al., 2006). In experimental and baseline contexts, the duration of the sentence was digitally manipulated using the pitch synchronous overlap-add utility (with a pitch range of 75–600 Hz) within PRAAT (<http://www.praat.org>) to make the duration of the final word 300 ms. An additional 458 distractor sentences were also recorded. All sentences were normalized for mean amplitude. An uppercase consonant letter probe (T, D, P or B) was presented on a large projection screen after each sentence. In experimental trials, the probe was the onset of the contextually appropriate final word. In baseline items, it was the onset of a contextual appropriate and phonetically unambiguous final word. Distractor items included 337 trials in which the phoneme target did not occur. In the remaining distractor

items targets appeared randomly in all word positions except sentence finally. As a result, only half of all trials contained the phoneme probe, and the position of targets was variable.

Procedure

By using a large number of distractor trials, a subtle phonetic speech manipulation to produce ambiguity, and no trials in which sentences were unambiguously ungrammatical or nonsensical, we hoped to make the task simple and naturalistic, focusing attention as much as possible on automatic sentence comprehension rather than effortful acoustic-phonetic processing. In a single testing session, participants completed a retrospective probe-monitoring task while magnetoencephalographic (MEG) and electroencephalographic (EEG) data were acquired in a three-layer room that was magnetically shielded. In each trial, participants heard a sentence. Four hundred milliseconds after the sentence ended and 300 ms after the end of the period we examined, a visual letter probe appeared on a screen. Participants were instructed to determine if that phoneme appeared in the sentence, and to indicate their decision by keypress, using their left hands. No sentences were repeated, and all conditions and probes were presented in pseudorandomized order in 674 trials, which were divided into six blocks (four blocks of 112 trials and two blocks of 113 trials). Subjects were given 1–3 minute breaks between blocks. During this break head position data were acquired and neural data files were saved. Of the fourteen subjects, 3 reported awareness of phonetic ambiguity on some trials when questioned during debriefing. None of the subjects reported any difficulty in determining whether targets were present.

MEG and MRI Acquisition and Reconstruction

This work employs a measure of localized neural activation based on reconstructions of the density of ionic currents produced by the synchronized activity of populations of large cortical neurons. These currents are measured directly via EEG, and indirectly by MEG, which measures magnetic fields induced by these currents. Our entire processing stream is summarized in Figure 1. All aspects of signal analysis, ROI identification, and Granger analysis were completed using GPS, a freely available GUI-based analysis and visualization tool created in our lab (<https://www.martinos.org/software/gps>).

MEG and EEG data were filtered between 0.1 and 400 Hz and recorded at a sampling rate of 1209.83 Hz. A 306-channel whole head Neuromag Vectorview system (Elekta, Helsinki, Finland) containing 204 planar gradiometers and 102 magnetometers was used to collect MEG data. This system includes a 70-electrode EEG cap with nose reference and horizontal and vertical electro-oculogram (EOG) recording. Electrode positions followed the international 10–10 system. The maximum allowable impedance for EEG electrodes was 10 kV. At the beginning of each block and at the end of the testing section, measurements of the head position within the MEG sensor array were taken. A Fast-track 3D Digitizer (Polhemus, Colchester, VT) was used to measure the position of scalp electrodes, nasion and two auricular anatomical landmarks as well as four head position indicator coils. We collected structural MRI data immediately after completion of MEG testing using a 1.5 T Avanto 32 channel “TIM” system with a MPRAGE sequence (TR = 2730 ms, T1 = 1000 ms, TE = 3.31 ms, flip angle = 7u, FOV = 256 mm, slice thickness = 1.33 mm) allowing the acquisition of a high-resolution 3D T1-weighted structural head only image. All

reconstructions of activity and Granger analyses were timelocked to the onset of the final word in each sentence.

In traditional component-based analyses researchers compare the timecourse of measures collected at the same electrode or sensor in two different conditions. In our approach we localize all MEG and EEG measures to cortical surfaces and reconstruct activation patterns within a condition by comparing localized MEG and EEG collected during task performance with baseline activity collected in the 300 msec interval preceding stimulus onset. Source localization depends on the solution of two problems. The forward problem involves modelling the pattern of EEG and MEG signals (measured by sensors adjacent to the scalp) that would be produced by hypothetical sources at specific locations in the cortex. This is facilitated by the use of boundary element methods (BEM) that use MRI reconstructions of brain, skull and scalp features to accurately predict conduction patterns. The inverse problem involves identifying cortical sources based on sensor data. Because there is no unique solution to the unstrained inverse problem, it can only be solved by leveraging a variety of constraints derived from anatomy, physiology and the physical behaviour of magnetic and electrical signals. There are a number of techniques available to solve the localization pattern, but our relied on the Minimal Norm Estimate technique (Dale & Sereno, 1993), which creates physiologically plausible and temporally smooth reconstructions of cortical activity.

We calculated an optimal inverse operator to localize the source of combined MEG/EEG activity using individual subjects' anatomical MRI data to constrain the forward solution in minimum norm estimate reconstructions of cortical surface currents (Dale & Sereno, 1993). Tessellated cortical surfaces were reconstructed based on individual MRI data and then decimated to create a cortical surface representation consisting of approximately 12,000 vertices, which served as dipole locations. We used three-layer BEM models (Hämäläinen & Sarvas, 1989; Oostendorp & Oosterom, 1989) to generate the forward model, modelling the dipole produced by each cortical vertex. This type of MRI-constrained MEG/EEG activation produces more accurate source localization than unimodal MEG or EEG analyses, while providing submillisecond temporal resolution that cannot be achieved using BOLD imaging techniques (Sharon, Hämäläinen, Tootell, Halgren, & Belliveau, 2007). MRI-constrained MEG/EEG data provide sufficient spatial resolution to localize sources at roughly the scale of Brodmann areas and sufficient temporal resolution to support meaningful timeseries analyses, which are essential to Granger analysis in order to examine temporally evolving event-related brain activity (Gow & Caplan, 2012). The combination of MEG and EEG improves source localization mainly by creating an extremely dense sensor array (70 EEG electrodes plus 306 MEG channels), and taking advantage of the fact that unlike EEG signals, MEG signals are not conducted by tissue and fall off quickly as a function of distance.

All neural analyses focused on the period between 100–500 ms after the onset of target words. We chose this range to encompass the time periods associated with the N1 and N200, which are early perceptual evoked components that show sensitivity to sentential context (Diaz & Swaab, 2007; Groppe et al., 2010; Sivonen, Maess, Lattner, & Friederici, 2006; van

den Brink, Brown, & Hagoort, 2001), and the N400, which is sensitive to cloze probability (Kutas & Hillyard, 1984).

ROI Identification

Regions of interest (ROIs) were generated from the bottom-up using an algorithm based on the mean strength and similarity of activation over time courses from all the vertices on the cortical surface. Activity maps from all subjects were spherically averaged and then morphed into an average brain (Fischl, Sereno, Tootell, & Dale, 1999). Vertices were selected based on the average activity map and were expanded outward to form continuous ROIs. The localization of vertices was based on the averaged cortical surface, but was verified based on projections back to subjects' individual automatically parcellated cortical surfaces. The identification of ROIs was divided into 3 dynamic steps. First, vertices with mean activation over the 95th percentile during the analysis period were identified as potential ROI centroids or seeds for identify ROIs. Any vertices within 5mm of these centroids were excluded as potential centroids, generating approximately 55 to 150 potential centroids. Second, we determined the spatial extent of ROIs by iteratively examining the similarity (Euclidean distance between) of normalized timecourses associated with contiguous vertices. Contiguous vertices that differed from the timeseries associated with an ROI's centroid by less than 0.5 standard deviations were included in the ROI. This ensured that ROIs showed homogenous activation timecourse structure, and provided a fuller description of ROI localization. A third step used similar comparisons of normalized timeseries similarity to eliminate redundant ROIs. This is an essential requirement of Granger analysis, which identifies causal influences based on the unique predictive information carried by each timeseries (Gow & Caplan, 2012). When two non-contiguous ROIs differed by less than 0.9 standard deviations, the ROI with the weaker (non-normalized) signal was eliminated. ROIs were labelled using an automatic parcellation procedure (Fischl, 2004) to identify primary units (e.g. MTG) coupled with visual inspection to identify subregions within those units (e.g. pMTG).

Kalman Filter-based Granger Analysis

Effective connectivity analyses relied on a Kalman-filter based implementation of Granger causality analysis (Gow & Caplan, 2012; Milde et al., 2010). We applied this to individual subjects' averaged timecourses of activation (reconstructed from combined MEG and EEG data) at each ROI. Granger analysis is a data-driven statistical technique that infers patterns of causal interaction based on the premise that causes carry information that uniquely predict their own effects (Granger, 1969). This focus on unique predictive information makes Granger immune to spurious correlation results if all nonredundant, potentially causal variables are examined. It can be applied to multivariate cases, and may be used to identify patterns of simultaneous reciprocal influence between variables (Geweke, 1982). We use Kalman filter techniques to estimate the coefficients for time-varying multivariate autoregressive prediction models, eliminating the usual need for unrealistic assumptions about the stationarity of evoked neural signals. Predictions were made at every timepoint (0.83 ms sample) over 5 lagged timepoints. This prediction lag or model order was assigned heuristically based on other studies with similar numbers of ROIs after the Aikaike and Bayesian Information Criteria analyses failed to identify a unique optimal model order for

these data. When estimating these models, Kalman filters assign less weight to anomalous (unusually high or low) observations. This property makes Kalman filters robust estimators, even given noisy input measures. The error terms associated with these models were subsequently used to estimate Granger causality. The use of Kalman filters in Granger analysis has been shown to produce robust multivariate analyses of networks of up to 58 nodes (Milde et al., 2010).

The strength of Granger Causation was estimated at each timepoint, for each directed interaction between ROIs using the Granger Causality Index or GCi (Milde et al., 2010). The GCi is the log ratio of the standard prediction error for a complete model of future activity at a single ROI versus that of a model that excludes one ROI (the potential causal signal). Once GCis were obtained, their significance was calculated through a bootstrapping method, which generated a null distribution of GCi values for that relationship and timepoint through 2000 iterations of reconstructed data created by the original Kalman filter results. We zeroed the coefficients for the hypothesized causal variable in the Kalman models and shuffled residuals from different timepoints to generate simulated data. These reconstructed data were then submitted to Granger analysis to create the distribution of GCi values associated with null effects. We compared the strength of GCi between conditions (over the same set of ROIs) using a binomial test (Cho & Campbell, 2000) that compared the number of timepoints that achieve a $p < 0.05$ in each condition over a predefined interval.

Results

Behavioral analyses showed a robust effect of sentential context on the identification of word-initial stop consonants. Subjects were more likely to categorize targets as voiced in contexts consistent with the voiced (93.9%, $SD = 3.1$) versus unvoiced interpretation (6.3 %, $SD = 4.7$) ($t(13) = 52.4$, $p < 0.00001$, $d_{\text{Cohen}} = 22.0$). All subjects heard each ambiguous word token in both contexts favouring voiced and unvoiced interpretations. Overall, they classified ambiguous items in a contextually consistent manner in 93.5% of all trials. Significantly, individual subjects interpreted 87.6% of the same individual ambiguous tokens (e.g. *^tdusk*) as voiced (DUSK) in contexts favouring the voiced interpretation, and unvoiced (TUSK) in contexts favouring the unvoiced interpretation. This suggests that categorization biases cannot be attributed to the phonetic properties of the stimuli we used. Signal detection theory analyses of performance on baseline and distractor condition trials showed a weak response bias ($c = -0.28$).

In the experimental condition, analyses were limited to trials in which an individual classified the same phonetically ambiguous token as voiced in contexts that encourage the voiced interpretation and unvoiced in contexts that favor the unvoiced interpretation. For purposes of comparison, the baseline condition consisted of trials in which a phonetically unambiguous word appeared in the same sentence-final position.

Using the combined activation data from experimental and baseline trials, we identified 32 ROIs (Table 2 and Figure 2). Critically, these included the left posterior superior temporal gyrus (pSTG), which is reliably sensitive to speech degradation in related studies (Davis et al., 2011; Guediche et al., 2013; Obleser & Eisner, 2009; Obleser & Kotz, 2010). As a

hypothesized hub of acoustic-phonetic processing (Chang et al., 2010; Gow, 2012; Hickok & Poeppel, 2007; Mesgarani et al., 2014; Price, 2012), the left pSTG is the focus of our hypothesis about top-down versus bottom-up accounts of sentential influences on speech categorization. The other ROIs are distributed over both hemispheres, and include two regions implicated in lexical representation, the left posterior middle temporal gyrus (pMTG) and supramarginal gyrus (SMG) (Gow, 2012; Hickok & Poeppel, 2007). They also included several regions thought to play a role in semantic representation and sentence integration including ROIs in the left and right middle temporal gyri (MTG), left temporal pole (TP), superior frontal gyrus (SFG), and left pars triangularis (ParsTri) (Binder et al., 2009; Price, 2012).

Figure 3 shows the relative influence of ROIs on left pSTG activation in experimental versus baseline trials. Additional visualizations of these data including separate visualizations of baseline and experimental Granger patterns and the timecourse of Granger causation can be found in the supplementary materials. All effects reported here were significant ($\alpha = 0.05$) after correction for multiple comparisons (Benjamini & Hochberg, 1995). Consistent with the predictions of an interactive framework, a number of regions showed strong top-down influences on pSTG activation. These influences were stronger in cases where sentential context influenced phoneme categorization for 10 ROIs including left SMG ($p < 0.05$), anterior MTG ($p < 0.05$), ParsTri ($p < 0.001$), two portions of left SFG ($p < 0.05$), and MFG ($p < 0.05$), as well as right posterior ITG ($p < 0.05$), two right SFG regions ($p < 0.001$), and MFG ($p < 0.005$). The baseline condition produced stronger influences on pSTG activation by a region in left SFG ($p < 0.001$), right AG ($p < 0.05$) and a region in right ITG ($p < 0.05$).

Several models hypothesize that the pMTG serves as an interface between acoustic-phonetic and lexico-semantic/lexico-syntactic representation (Gow, 2012; Hickok & Poeppel, 2007; Snijders, Petersson, & Hagoort, 2010). Under this interpretation, the pMTG might be expected to directly influence pSTG activation during interactive processing. Left pMTG did exert stronger influences on left pSTG activation in the experimental condition, but this effect did not reach statistical significance ($p = 0.11$). However, additional analyses showed that the left pMTG indirectly influenced left pSTG activation via the left SMG, which had strong bidirectional interactions with pMTG (pMTG to SMG, $p < 0.001$; SMG to pMTG, $p < 0.001$). Comparison of the time course of Granger causation suggests that MTG influences on SMG activation largely precede SMG influences on STG in the experimental condition (Figure 4). This pattern is most pronounced after 200 msec when the majority of SMG influences on STG activation are observed.

Discussion

The purpose of this work is to determine whether sentential influences on speech interpretation are the result of top-down influences on acoustic-phonetic representation, or of post-perceptual processes that weigh the output of feedforward sentential and acoustic-phonetic analyses. The current results demonstrate that a number of brain regions associated with higher linguistic representations influence activation in the left pSTG, a region strongly implicated in acoustic-phonetic processing, in trials in which sentential context influences

the categorization of perceptually ambiguous phonemes. This result strongly supports the predictions of an interactive model.

Given the largely unsystematic mapping between supralexical semantic or syntactic representation and speech sounds, it is not surprising that the SMG, a brain region associated with lexical representation, is among those that influence the pSTG. The idea that the SMG is involved in lexical representation is supported by converging evidence from aphasiology and BOLD imaging. Patients with damage to the left SMG often exhibit reproduction conduction aphasia, a syndrome in which word comprehension is spared, but produce frequent lexical phonological paraphasias and show word repetition deficits that are sensitive to lexical frequency (Goldrick & Rapp, 2007). Neuroimaging studies further demonstrate that SMG BOLD responses are modulated by lexical properties, including phonological neighborhood density (Prabhakaran, Blumstein, Myers, Hutchison, & Britton, 2006), and phonological competitor environment in speech perception (Righi, Blumstein, Mertus, & Worden, 2009) and production (Peramunage, Blumstein, Myers, Goldrick, & Baese-Berk, 2011). Gow's dual lexicon model (2012) suggests that the SMG is the home of wordform representations that mediate the mapping between acoustic-phonetic and articulatory representations in speech production and perception. This finding is supported by a series of results that show that SMG influence on pSTG activation accompanies behavioral evidence of lexical influences on speech interpretation (Gow & Nied, 2014; Gow & Olson, under review; Gow et al., 2008).

It is somewhat surprising that the SMG is the primary lexical influence on acoustic-phonetic processing in this paradigm. Both Hickok and Poeppel's dual stream model of speech processing (2004; 2007), and Gow's (2012) dual lexicon model argue that the pMTG, not the SMG, mediates the mapping between acoustic-phonetic and both syntactic and semantic levels of representation. One possibility is that the current results reflect simple lexical top-down effects on speech perception that are unrelated to sentential bias manipulations. As noted above, the SMG has been shown to play this role in other work, specifically in tasks that involve explicit phonemic categorization of phonetically ambiguous segments (Gow & Nied, 2014; Gow et al., 2008). Moreover, like these other studies, our analyses compared the processing of phonetically ambiguous items in the experimental condition, with unambiguous items in the control condition. Under such a reading, it is still unclear why a purely lexical mechanism would consistently favor the sententially consistent interpretation when both consistent and anomalous lexical interpretations of the input were available.

The dual lexicon model suggests an explanation. In this model, wordform representations in the two lexica serve as hidden nodes, mediating mappings between acoustic phonetic representation and articulation (SMG) and semantic or syntactic representation (pMTG). Gow (2012) argues that different types of form representation are computationally optimal for mediating these two types of mapping. In our study, the behavioral task requires explicit phoneme categorization. In this case, voice onset time information may map more systematically to articulatory representation than to semantic or syntactic representations. We suggest that sentential context influenced the selection of wordforms at the pMTG, and that the presumably strong mapping between wordform representations in the pMTG and SMG in turn allowed the selection of SMG representations that clearly specified

representations of voicing cues. This interpretation is consistent with the finding that MTG influences on SMG precede most SMG influences on STG activation. In tasks that require semantic interpretation, but not explicit phoneme categorization, SMG mediation may not be required. We found this pattern in a study in which lexical factors influenced the perception of coarticulated speech in a sentence-picture matching task (Gow & Segawa, 2009).

Several areas commonly associated with a broader semantic network (Binder et al., 2009; Price, 2012) also directly influenced left pSTG activation in these trials. These included left pars triangularis, left anterior MTG, and several regions in the left SFG. The left inferior frontal gyrus is frequently activated in studies of sentential influences on speech categorization (Davis et al., 2011; Guediche et al., 2013; Obleser & Kotz, 2010; Obleser et al., 2007; Sohoglu et al., 2012; Zekveld et al., 2006). Studies intended to isolate semantic processes frequently implicate the left inferior frontal gyrus, but show more reliable activation in the pars orbitalis than the pars triangularis (Binder et al., 2009; Bookheimer, 2002; Fiez, 1997). It is nevertheless unclear what role the inferior frontal gyrus plays in this work because this region is widely implicated in many task relevant processes including phonological and syntactic processing and the executive processes related to working memory (c.f. Badre & Wagner, 2007; Hagoort, 2005; Nixon, Lazarova, Hodinott-Hill, Gough, & Passingham, 2004; Price, 2012; Vigneau et al., 2006). One possible interpretation is that pars triangularis influences on pSTG activation reflect non-semantic working memory processes. In addition to its role in speech processing, the pSTG plays a role in auditory short-term memory. Patient studies show that damage to this region impairs both speech processing and working memory performance (Leff et al., 2009). BOLD imaging results further suggest that this role is governed by executive control processes localized in the left inferior frontal gyrus (Cohen et al., 1997). Here, interactions between Pars triangularis and pSTG may reflect processes involved in calling up earlier sentential context to guide perceptual processing, or increased processing effort required to form canonical phonological working memory representations of phonetically ambiguous input.

The anterior MTG and SFG ROIs may play a more directly semantic role in this task. Patterson and colleagues (Patterson, Nestor, & Rogers, 2007) review evidence from imaging and pathology and suggest that the anterior MTG serves as an amodal, non-lexical semantic hub for distributed modality and property-specific components of a broadly distributed semantic network. Hickok and Poeppel (2007) identify an overlapping but slightly posterior anterior MTG region as playing a role in a combinatorial network involved in both semantic and syntactic processing. Both interpretations suggest that this region is crucial to sentential analysis, but they carefully delineate it as a non-lexical level. Similarly, Binder and colleagues (Binder et al., 2009) identify a region including our SFG ROIs as portions of a left dorsomedial prefrontal cortex region that is distinct from dorsolateral prefrontal cortex, and that plays a role in goal-directed semantic retrieval. Given the largely unsystematic mapping between semantic or syntactic representation and speech sounds, it is puzzling how these regions would act directly on pSTG without lexical mediation. One possibility is that the ventral lexicon associated with pMTG extends more anteriorly than previously suspected, with anterior and posterior regions devoted to different semantic categories (Warrington & McCarthy, 1987). The role of SFG may also be open to reinterpretation. Lesion studies suggest that the SFG plays a role in working memory (du Boisgueheneuc et

al., 2006). If this is the case, its role may be related to the working memory role attributed to left pars triangularis above.

The current results suggest that sentential context influences acoustic-phonetic processes. In general, interaction potentially provides a network of mutual constraints that should support robust speech processing. It is not immediately clear though why interaction needs to reach the level of perceptual analysis. Within a partially feedforward system, the output of acoustic-phonetic processing could be wordforms that are largely consistent with two words (e.g. DUSK and TUSK). Sentence level constraints could then select the better alternative without interacting with acoustic-phonetic representation. Several observers have noted that top-down “filling-in” appears to serve no purpose in a system that has already “figured out” the correct interpretation of a perceptual ambiguity (Dennett, 1992; Fodor, 1983; Norris et al., 2000). The interactions between pSTG and frontal regions associated with working memory suggest a possible purpose – error correction. Error correction is fundamental to computer hardware and software design because it makes processing more robust by minimizing ambiguity that otherwise accumulates through any imprecision in the copying and routing of output of computations (Shannon, 1949). There is clearly potential for this kind of representational drift in the context of a distributed neural language processing system that employs parallel processing of acoustic-phonetic, phonological, lexical, semantic and syntactic representations and interacts with similarly distributed perceptual, working memory and attention systems.

Our results also suggest that interactions between frontal regions and the pSTG reflect working memory operations. If sentential influences on lexical selection did not influence acoustic-phonetic representation in pSTG, the representation in working memory could diverge from lexical representation. One can easily imagine similar representational drift causing divergences in other levels of linguistic representation, leading to representations that cannot be bound into a coherent whole. Therefore, to the extent that parallel processes and representations interact, there appears to be a computational value to interactive processes that mutually constrain and coordinate all levels of representation.

In summary, using effective connectivity analyses of MRI-constrained MEG and EEG data, we have demonstrated the existence of many top-down influences on acoustic-phonetic representation in left pSTG activation that accompany sentence level semantic and syntactic influences on the interpretation of phonetically ambiguous speech sounds. These include influences mediated by hypothesized wordform representations in left SMG, and indirect influences (via SMG) by wordform representations mediating the mapping between sound and syntax and semantics in left pMTG. These results support an interactive functional architecture for spoken sentence processing.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to thank Cristin Sullivan for conducting the behavioral pilot study and writing the stimuli, Conrad Nied for digital manipulation of stimuli, subject testing and the development of our automated processing stream, Nao Suzuki, and Reid Vancelette for their assistance in running the experiment, Jennifer Michaud, Mark Vangel and Juliane Venturelli Silva Lima for statistical advice and review, and David Caplan, Tom Sgouros, Bob McMurray and two anonymous reviewers for their thoughtful feedback on earlier versions of the manuscript. This work was supported by the National Institute of Deafness and Communicative Disorders (R01 DC003108) and benefited from support from the MIND Institute and the NCR Regional Resource Grant (41RR14075) for the development of technology and analysis tools at the Athinoula A. Martinos Center for Biomedical Imaging.

References

- Abada SH, Baum SR, Titone D. The effects of contextual strength on phonetic identification in younger and older listeners. *Experimental aging research*. 2008; 34(3):232–250. [PubMed: 18568981]
- Badre D, Wagner AD. Left ventrolateral prefrontal cortex and the cognitive control of memory. *Neuropsychologia*. 2007; 45(13):2883–2901. [PubMed: 17675110]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, B (Methodological)*. 1995; 57(1):289–300.
- Binder JR, Desai RH, Graves WW, Conant LL. Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*. 2009; 19(12):2767–2796. [PubMed: 19329570]
- Bookheimer S. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual review of neuroscience*. 2002; 25:151–188.
- Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. *J Acoust Soc Am*. 1988; 84(1):101–114. [PubMed: 3411038]
- Borsky S, Shapiro LP, Tuller B. The temporal unfolding of local acoustic information and sentence context. *J Psycholinguist Res*. 2000; 29(2):155–168. [PubMed: 10709181]
- Borsky S, Tuller B, Shapiro LP. “How to milk a coat:” The effects of semantic and acoustic information on phoneme categorization. *Journal of the Acoustic Society of America*. 1998; 103(5): 2670–2676.
- Chang EF, Rieger JW, Johnson K, Berger MS, Barbaro NM, Knight RT. Categorical speech representation in human superior temporal gyrus. *Nature neuroscience*. 2010; 13(11):1428–1432. [PubMed: 20890293]
- Cho RJ, Campbell MJ. Transcription, genomes, function. *Trends in genetics : TIG*. 2000; 16(9):409–415. [PubMed: 10973070]
- Cohen JD, Perlstein WM, Braver TS, Nystrom LE, Noll DC, Jonides J, Smith EE. Temporal dynamics of brain activation during a working memory task. *Nature*. 1997; 386(6625):604–608. [PubMed: 9121583]
- Connine CM. Constraints on interactive processes in auditory word recognition: The role of sentence context. *Journal of Memory and Language*. 1987; 26(2):527–538.
- Connine CM, Blasko DG, Hall M. Effects of subsequent context in auditory word recognition: Temporal and linguistic constraints. *Journal of Memory and Language*. 1991; 30:234–250.
- Connine CM, Clifton C Jr. Interactive use of lexical information in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*. 1987; 13(3):291–299. [PubMed: 2953858]
- Dale AM, Sereno MI. Improved localization of cortical activity by combining EEG and MEG with MRI cortical reconstruction: A linear approach. *Journal of cognitive neuroscience*. 1993; 5(2): 162–176. [PubMed: 23972151]
- Davis MH, Ford MA, Kherif F, Johnsrude IS. Does semantic context benefit speech understanding through “top-down” processes? Evidence from time-resolved sparse fMRI. *Journal of cognitive neuroscience*. 2011; 23(12):3914–3932. [PubMed: 21745006]

- Dennett, DC. 'Filling-in' versus finding out. In: Pick, HL.; Van den Broek, P.; Knill, DC., editors. *Cognition, Conception, and Methodological Issues*. Washington, D.C.: American Psychological Association; 1992. p. 33-49.
- Diaz MT, Swaab TY. Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. *Brain research*. 2007; 1146:85–100. [PubMed: 16952338]
- Dronkers NF, Wilkins DP, Van Valin RD Jr, Redfern BB, Jaeger JJ. Lesion analysis of the brain areas involved in language comprehension. *Cognition*. 2004; 92(1–2):145–177. [PubMed: 15037129]
- du Boisgueheneuc F, Levy R, Volle E, Seassau M, Duffau H, Kinkingnehun S, Dubois B. Functions of the left superior frontal gyrus in humans: a lesion study. *Brain : a journal of neurology*. 2006; 129(Pt 12):3315–3328. [PubMed: 16984899]
- Fiez JA. Phonology, semantics, and the role of the left inferior prefrontal cortex. *Human brain mapping*. 1997; 5(2):79–83. [PubMed: 10096412]
- Fischl B. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*. 2004; 14(1):11–22. [PubMed: 14654453]
- Fischl B, Sereno MI, Tootell RBH, Dale AM. High resolution intersubject averaging and a coordinate system for the cortical surface. *Human brain mapping*. 1999; 8(4):272–284. [PubMed: 10619420]
- Fodor, JA. *The modularity of mind*. Cambridge, MA: MIT Press; 1983.
- Gagnepain P, Henson RN, Davis MH. Temporal predictive codes for spoken words in auditory cortex. *Current biology : CB*. 2012; 22(7):615–621. [PubMed: 22425155]
- Geweke J. Measurement of linear dependence and feedback between multiple time series. *Journal of the American Statistical Association*. 1982; 77(378):304–313.
- Goldrick M, Rapp B. Lexical and post-lexical phonological representations in spoken production. *Cognition*. 2007; 102(2):219–260. [PubMed: 16483561]
- Gow DW. The cortical organization of lexical knowledge: A dual lexicon model of spoken language processing. *Brain and language*. 2012; 121(3):273–288. [PubMed: 22498237]
- Gow DW, Caplan DN. New levels of language processing complexity and organization revealed by granger causation. *Frontiers in psychology*. 2012; 3:506. [PubMed: 23293611]
- Gow DW, Gordon PC. Coming to terms with stress: effects of stress location in sentence processing. *Journal of Psycholinguistic Research*. 1993; 22(6):545–578. [PubMed: 8295163]
- Gow DW, Nied AC. Rules from words: Phonotactic biases in speech perception. *PLoS one*. 2014; 9(1): 1–12.
- Gow DW, Olson BB. Lexical mediation of phonotactic frequency effects on spoken word recognition: A Granger causality analysis of MRI-constrained MEG/EEG data. *Journal of Memory and Language*. (under review).
- Gow DW, Segawa JA. Articulatory mediation of speech perception: a causal analysis of multi-modal imaging data. *Cognition*. 2009; 110(2):222–236. [PubMed: 19110238]
- Gow DW, Segawa JA, Ahlfors SP, Lin F-H. Lexical influences on speech perception: A Granger causality analysis of MEG and EEG source estimates. *NeuroImage*. 2008; 43(3):614–623. [PubMed: 18703146]
- Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 1969; 37(3):424–438.
- Groppe DM, Choi M, Huang T, Schilz J, Topkins B, Urbach TP, Kutas M. The phonemic restoration effect reveals pre-N400 effect of supportive sentence context in speech perception. *Brain research*. 2010; 1361:54–66. [PubMed: 20831863]
- Guediche S, Salvata C, Blumstein SE. Temporal cortex reflects effects of sentence context on phonetic processing. *Journal of cognitive neuroscience*. 2013; 25(5):706–718. [PubMed: 23281778]
- Hagoort P. On Broca, brain, and binding: a new framework. *Trends in cognitive sciences*. 2005; 9(9): 416–423. [PubMed: 16054419]
- Hämäläinen M, Sarvas J. Realistic conductivity geometry model of the human head for interpretation of neuromagnetic data. *IEEE transactions on Biomedical Engineering, BME-36*. 1989
- Hannagan T, Magnuson JS, Grainger J. Spoken word recognition without a TRACE. *Frontiers in psychology*. 2013; 4

- Hickok G, Poeppel D. Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition*. 2004; 92(1–2):67–99. [PubMed: 15037127]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8:393–402. [PubMed: 17431404]
- Kalikow DN, Stevens KN, Elliot LL. Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustic Society of America*. 1977; 61(5):1335–1351.
- Kutas M, Hillyard SA. Brain potentials reflect word expectancy and semantic association during reading. *Nature*. 1984; 307:161–163. [PubMed: 6690995]
- Leff AP, Schofield TM, Crinion JT, Seghier ML, Grogan A, Green DW, Price CJ. The left superior temporal gyrus is a shared substrate for auditory short-term memory and speech comprehension: evidence from 210 patients with stroke. *Brain : a journal of neurology*. 2009; 132(Pt 12):3401–3410. [PubMed: 19892765]
- Marslen-Wilson W, Tyler LK. The temporal structure of spoken language understanding. *Cognition*. 1980; 8(1):1–71. [PubMed: 7363578]
- Massaro DW. Testing between the TRACE model and the fuzzy logical model of speech perception. *Cognitive Psychology*. 1989; 21(3):398–421. [PubMed: 2758786]
- McClelland JL, Elman JL. The TRACE model of speech perception. *Cognitive Psychology*. 1986; 18(1):1–86. [PubMed: 3753912]
- Mesgarani N, Cheung C, Johnson K, Chang EF. Phonetic feature encoding in human superior temporal gyrus. *Science*. 2014; 343(6174):1006–1010. [PubMed: 24482117]
- Milde T, Leistriz L, Astolfi L, Miltner WH, Weiss T, Babiloni F, Witte H. A new Kalman filter approach for the estimation of high-dimensional time-variant multivariate AR models and its application in analysis of laser-evoked brain potentials. *NeuroImage*. 2010; 50(3):960–969. [PubMed: 20060483]
- Miller GA, Heise GA, Lichten W. The intelligibility of speech as a function of the context of test materials. *Journal of Experimental Psychology*. 1951; 41(5):329–335. [PubMed: 14861384]
- Miller GA, Isard S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*. 1963; 2(3):217–228.
- Miller JL, Green KP, Schermer T. A distinction between the effects of sentential speaking rate and semantic congruity on word identification. *Perception and Psychophysics*. 1984; 36(4):329–337. [PubMed: 6522228]
- Nixon P, Lazarova J, Hodinott-Hill I, Gough P, Passingham R. The inferior frontal gyrus and phonological processing: an investigation using rTMS. *Journal of cognitive neuroscience*. 2004; 16(2):289–300. [PubMed: 15068598]
- Norris D, McQueen JM, Cutler A. Merging information in speech recognition: feedback is never necessary. *The Behavioral and brain sciences*. 2000; 23(3):299–325. discussion 325–270. [PubMed: 11301575]
- Obleser J, Eisner F. Pre-lexical abstraction of speech in the auditory cortex. *Trends in cognitive sciences*. 2009; 13(1):14–19. [PubMed: 19070534]
- Obleser J, Kotz SA. Expectancy constraints in degraded speech modulate the language comprehension network. *Cereb Cortex*. 2010; 20(3):633–640. [PubMed: 19561061]
- Obleser J, Wise RJS, Dresner AM, Scott SK. Functional Integration across Brain Regions Improves Speech Perception under Adverse Listening Conditions. *Journal of Neuroscience*. 2007; 27(9):2283–2289. [PubMed: 17329425]
- Oostendorp TF, Oosterom A. Source parameter estimation in inhomogeneous volume conductors of arbitrary shape. *IEEE transactions on biomedical engineering*. 1989; 36:382–391. [PubMed: 2921073]
- Patterson K, Nestor PJ, Rogers TT. Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature reviews. Neuroscience*. 2007; 8(12):976–987. [PubMed: 18026167]
- Peramunage D, Blumstein SE, Myers EB, Goldrick M, Baese-Berk M. Phonological neighborhood effects in spoken word production: an fMRI study. *Journal of cognitive neuroscience*. 2011; 23(3):593–603. [PubMed: 20350185]

- Prabhakaran R, Blumstein SE, Myers EB, Hutchison E, Britton B. An event-related fMRI investigation of phonological-lexical competition. *Neuropsychologia*. 2006; 44(12):2209–2221. [PubMed: 16842827]
- Price CJ. A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*. 2012; 62(2):816–847. [PubMed: 22584224]
- Righi G, Blumstein SE, Mertus J, Worden MS. Neural systems underlying lexical competition: An eye tracking and fMRI study. *Journal of cognitive neuroscience*. 2009; 22(2):213–224. [PubMed: 19301991]
- Shannon, CE. A mathematical theory of communication. Urbana, IL: University of Illinois Press; 1949.
- Sharon D, Hämäläinen MS, Tootell RB, Halgren E, Belliveau JW. The advantage of combining MEG and EEG: comparison to fMRI in focally stimulated visual cortex. *NeuroImage*. 2007; 36(4):1225–1235. [PubMed: 17532230]
- Sivonen P, Maess B, Lattner S, Friederici AD. Phonemic restoration in a sentence context: evidence from early and late ERP effects. *Brain research*. 2006; 1121(1):177–189. [PubMed: 17027933]
- Snijders TM, Petersson KM, Hagoort P. Effective connectivity of cortical and subcortical regions during unification of sentence structure. *NeuroImage*. 2010; 52(4):1633–1644. [PubMed: 20493954]
- Sohoglu E, Peelle JE, Carlyon RP, Davis MH. Predictive top-down integration of prior knowledge during speech perception. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2012; 32(25):8443–8453. [PubMed: 22723684]
- van den Brink D, Brown CM, Hagoort P. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of cognitive neuroscience*. 2001; 13(7):967–985. [PubMed: 11595099]
- Vigneau M, Beaucousin V, Herve PY, Duffau H, Crivello F, Houde O, Tzourio-Mazoyer N. Meta-analyzing left hemisphere language areas: phonology, semantics, and sentence processing. *NeuroImage*. 2006; 30(4):1414–1432. [PubMed: 16413796]
- Warrington EK, McCarthy RA. Categories of knowledge. Further fractionations and an attempted integration. *Brain : a journal of neurology*. 1987; 110(Pt 5):1273–1296. [PubMed: 3676701]
- Zekveld AA, Heslenfeld DJ, Festen JM, Schoonhoven R. Top-down and bottom-up processes in speech comprehension. *NeuroImage*. 2006; 32(4):1826–1836. [PubMed: 16781167]

Appendix

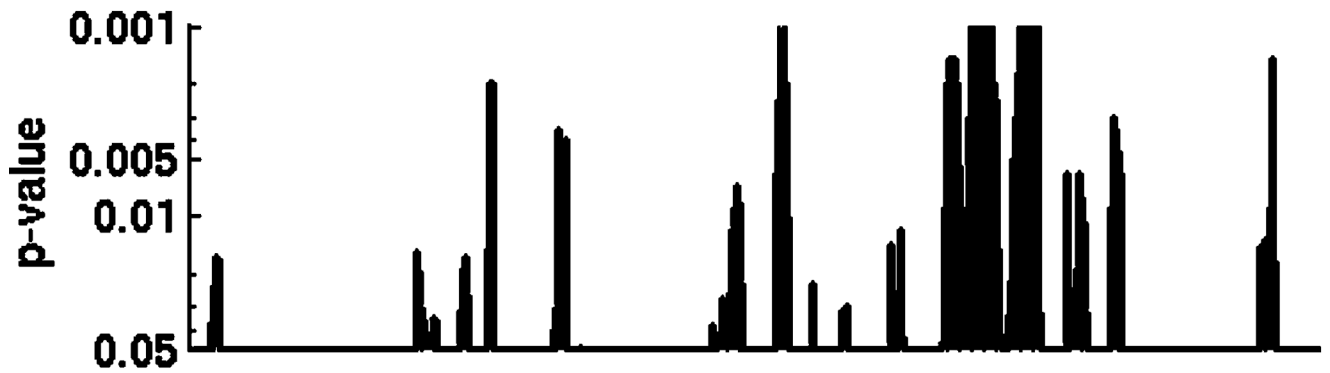
Voiced		Unvoiced	
Sentence	Target	Sentence	Target
The little boy's name is T/DAN.	D	UV rays give you a T/DAN.	T
His smelly cellar is very T/DANK.	D	My fish swims in his T/DANK.	T
Some bears sleep in a T/DEN.	D	Nine plus one will equal T/DEN.	T
Some people who have pneumonia T/DIE.	D	He wore a clip on T/DIE.	T
He will never give a T/DIME.	D	My clock says we have T/DIME.	T
He has chips for the T/DIP.	D	Leave your server a big T/DIP.	T
She says her horoscope is T/DIRE.	D	My bike is missing a T/DIRE.	T
I have a phone bill T/DUE.	D	One plus one always is T/DWO.	T
Employees who fail face clear T/DOOM.	D	We bury kings in a T/DOMB.	T
Relax when your work is T/DONE.	D	His professor's book weighs a T/DON.	T
Rain was almost always pouring T/DOWN.	D	My mom lives in T/DOWN.	T
Our sink has a broken T/DRAIN.	D	You can come here by T/DRAIN.	T

Voiced		Unvoiced	
Sentence	Target	Sentence	Target
Please fix my sinks annoying T/DRIP.	D	Have some fun on your T/DRIP.	T
When he was upset he T/DREW.	D	Your love for me is T/DRUE.	T
When will a hole be T/DUG.	D	She gave her dress a T/DUG.	T
Our beach is near a T/DUNE.	D	Her piano was never in T/DUNE.	T
He calls every woman a T/DAME.	D	I hope his lion is T/DAME.	T
The jungle growth was quite T/DENSE.	D	Her mood was really T/DENSE.	T
She was hoping for a T/DEAL.	D	She likes some shades of T/DEAL	T
There's corn flour in the T/DOUGH.	D	My shoe pinches my big T/DOE.	T
Without watering your flowers will T/DROOP.	D	She joined a brownie T/DROOP.	T
After four beers he was T/DRUNK.	D	I really should pack my T/DRUNK.	T
Wear a raincoat to stay T/DRY.	D	You should give yoga a T/DRY.	T
The moon rises just at T/DUSK.	D	A walrus was missing a T/DUSK.	T
Voiced		Unvoiced	
On summer vacation I get P/BORED.	B	Eat after the drinks are P/BOURED.	P
We collected clams in the P/BAY.	B	The hotel maid earned her P/BAY.	P
The surf's great at that P/BEACH.	B	The girl ate a juicy P/BEACH.	P
It catches worms with its P/BEAK.	B	The climber has reached the P/BEAK.	P
After work she drinks a P/BEER.	B	Edit your essay with a P/BEER.	P
My sister always tries her P/BEST.	B	The sneaky rodent was a P/BEST.	P
You will regret making that P/BET.	B	A dog is a great P/BET.	P
Now is time to sell, not P/BUY.	B	My favorite desert is berry P/BIE.	P
The great Dane was really P/BIG.	B	My friend eats like a P/BIG.	P
Throw your shoes in the P/BIN.	B	Try to knock down the P/BIN.	P
Do not leave any answers P/BLANK.	B	He had to walk the P/BLANK.	P
The Twins have a strong P/BOND.	B	Let's go fishing in the P/BOND.	P
You may now kiss the P/BRIDE.	B	Always do your work with P/BRIDE.	P
There are sheets on the P/BUNKS.	B	Teenage boys may act like P/BUNKS.	P
You forgot to trim that P/BUSH.	B	To close the door just P/BUSH.	P
The cut on my knee P/BLED.	B	The man losing his job P/BLED.	P
The infielder ran towards second P/BASE.	B	The runner maintains a fast P/BACE.	P
He used to throw out his P/BACK.	B	He carried a tent in his P/BACK.	P
This cheese tastes really P/BAD.	B	His desk has a mouse P/BAD.	P
He saw a giant grizzly P/BEAR.	B	He ate a succulent fresh P/BEAR.	P
The suspect was released on P/BAIL.	B	The once tan girl was P/BALE.	P
He favors an assault rifle P/BAN.	B	I washed the dirty frying P/BAN.	P
We had drinks at the P/BAR.	B	The golfers game was below P/BAR.	P
The guard dog started to P/BARK.	B	He took a walk in the P/BARK.	P
She relaxed with a hot P/BATH.	B	They wandered off of the P/BATH.	P

Voiced		Unvoiced	
Sentence	Target	Sentence	Target
The hive was full of angry P/BEES.	B	His favorite vegetables are P/BEAS.	P
The homeless man was forced to P/BEG.	B	He hung his coat on the P/BEG.	P
He never paid his electric P/BILL.	B	He always took a vitamin P/BILL.	P
The children were afraid of the P/BULL.	B	To close the door, you P/BULL.	P
There a dozen in a P/BUNCH.	B	The fighter threw the first P/BUNCH.	P

Experimental

Left Posterior MTG to Left SMG



Left SMG to Left STG

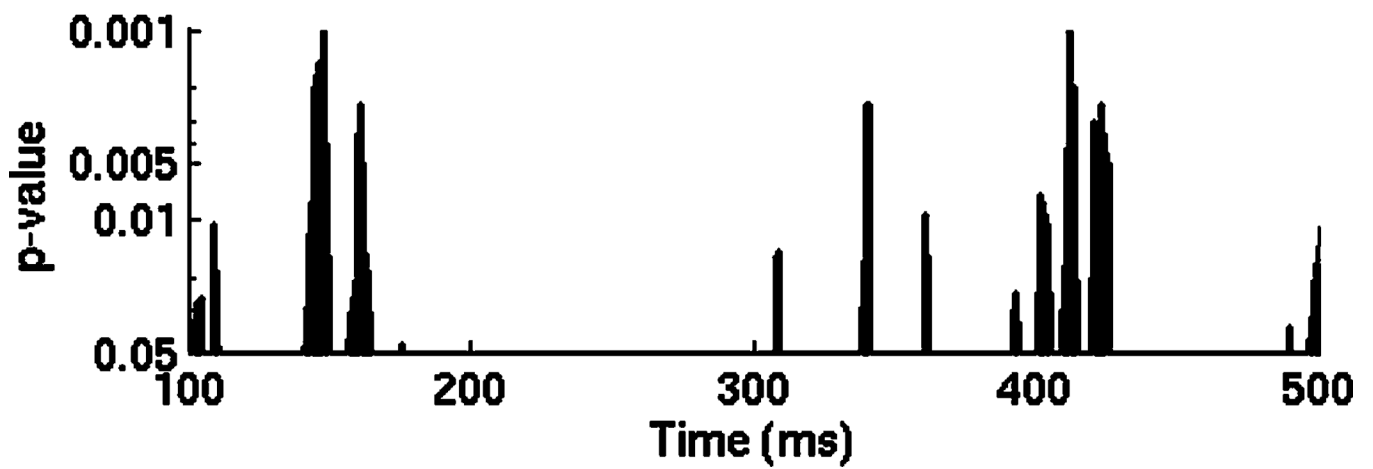


Figure 1.
Overview of the GPS processing stream.

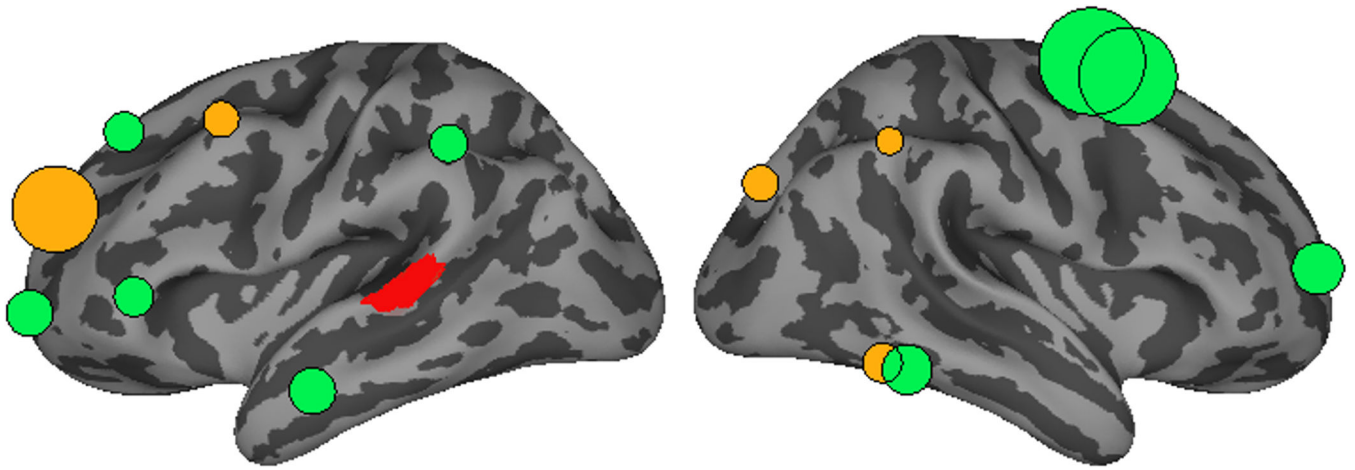


Figure 2. Regions of interest visualized over an inflated averaged cortical surface.

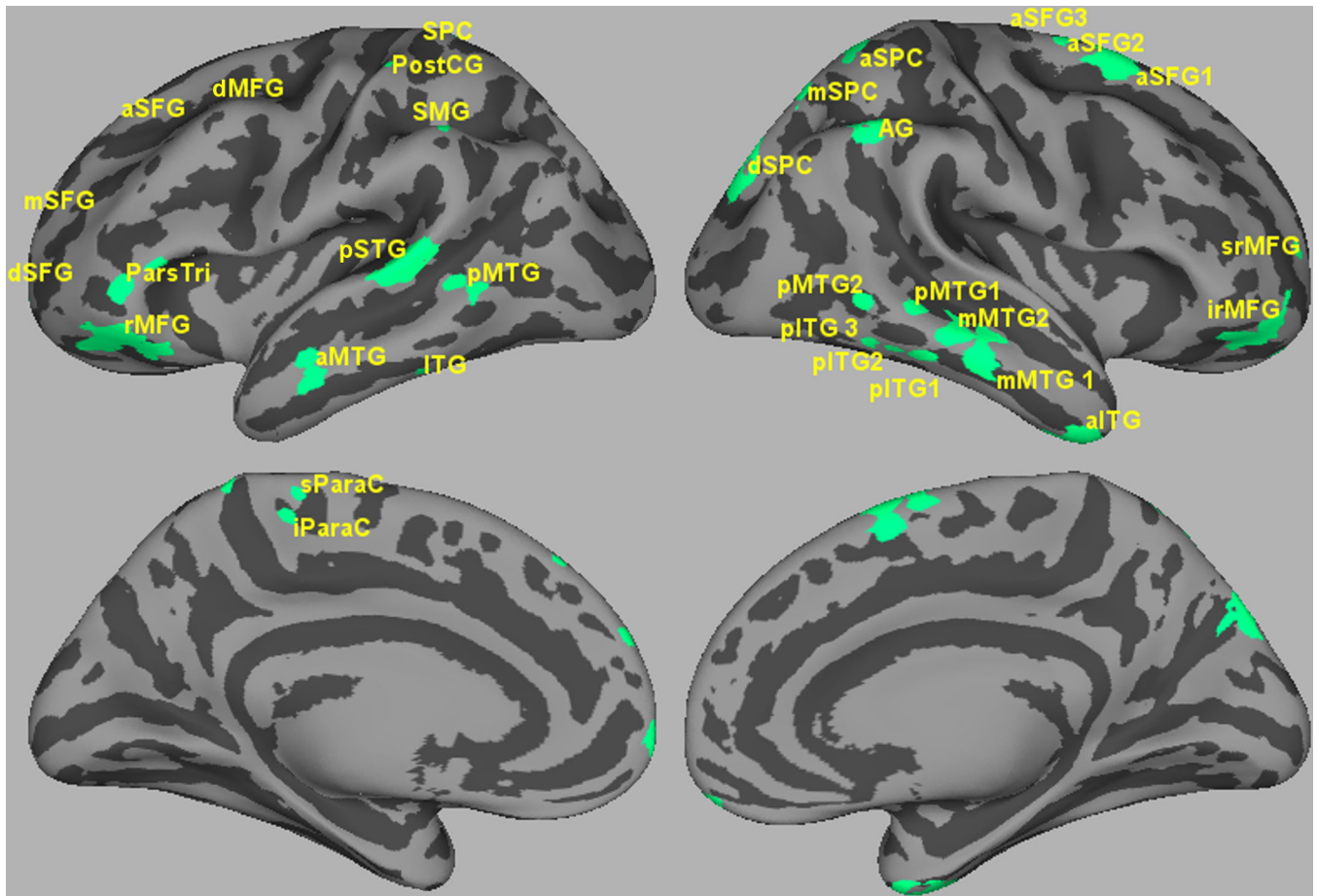


Figure 3.

A comparison of top-down influences on left pSTG activation (shown in red) after the onset (100–500 ms) of experimental versus baseline target-bearing words. The size of circles over ROIs reflects the difference in strength of Granger causation (number of time points that show GCi values with $p < 0.05$). Green circles indicate stronger influence in the experimental condition. Orange indicates stronger influence in the control condition. All differences shown have $p < 0.05$.

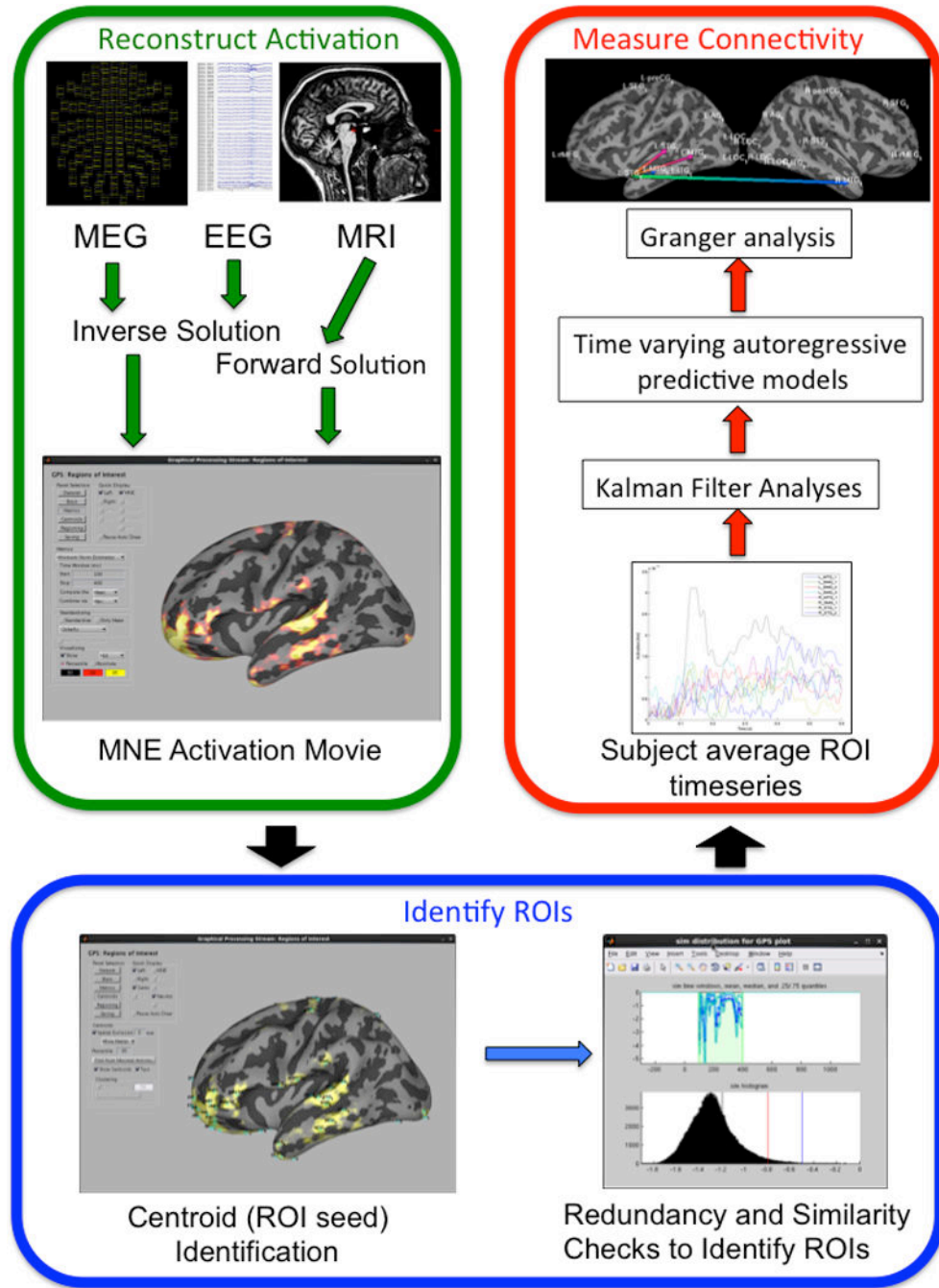


Figure 4. A comparison of the timecourse of Granger causality measures (p-value of each timepoint based on bootstrapping analysis) of left posterior MTG influence on left SMG activation (top) and left SMG influences on left STG activation in the experimental condition. Analyses are time-locked to the onset of the sentence final words bearing phonetically ambiguous onsets.

Table 1

Stimuli and conditions.

Condition	Context	Example	Target	Present	Trials
Exp.	Voiced	<i>The moon rises just at^dusk.</i>	[D]	YES	54
	Unvoiced	<i>A walrus was missing a^dusk.</i>	[T]	YES	54
Baseline	Voiced	<i>The night sky was dark.</i>	[D]	YES	54
	Unvoiced	<i>My car needs a tire.</i>	[T]	YES	54
Distractor		<i>Torn jeans are back in fashion.</i>	[T]	YES	121
		<i>She needs to pay for rent</i>	[B]	NO	337

Table 2

Regions of interest defined over experimental and baseline conditions for the interval of 100–500 ms following the onset of target bearing words. Numbers indicate the rank of mean activation strength in each parcellation unit in units with more than one ROI.

Label	Name	MNI Coordinates (X, Y, Z)		
Left Hemisphere				
ITG	inferior temporal gyrus	-54.9	-42.45	-29.28
MTG	anterior middle frontal gyrus	-63.79	-12.23	-23.19
MTG	posterior middle frontal gyrus	-62.25	-60.1	3.05
ParaC	superior paracentral lobule	-3.86	-25.25	63.77
ParaC	inferior paracentral lobule	-3.61	-31.58	56.91
ParsTri	pars triangularis	-52.18	33.75	1.02
SFG	anterior superior frontal gyrus	-4.96	44.04	49.46
SFG	middle superior frontal gyrus	-5.65	61.75	26.62
SFG	dorsal superior frontal gyrus	-5.67	66.44	1.4
SMG	supramarginal gyrus	-51.72	-45.11	48.93
SPC	superior parietal gyrus	-12.7	-49.46	74.28
STG	superior temporal gyrus	-67.78	-35.11	5.58
MFG	middle frontal gyrus	-33.56	14.91	54.13
postCG	postcentral gyrus	-41.92	-29.77	66.49
rMFG	rostral middle frontal gyrus	-42.91	46.15	-5.48
Right Hemisphere				
AG	angular gyrus	46.35	-51.19	44.28
ITG	anterior inferior temporal gyrus	39.53	-0.76	-46.7
ITG	posterior inferior temporal gyrus 1	57.63	-58.41	-17.38
ITG	posterior inferior temporal gyrus 2	58.96	-45.16	-22.67
ITG	posterior inferior temporal gyrus 3	58.98	-52.33	-22.3
MTG	mid middle frontal gyrus 1	63.08	-25.75	-19.94
MTG	mid middle frontal gyrus 2	69.64	-29.25	-12.3
MTG	posterior middle frontal gyrus 1	66.31	-39.83	-2.29
MTG	posterior middle frontal gyrus 2	56.85	-53.6	-4.6
SFG	anterior superior frontal gyrus 1	21.31	19.24	59.99
SFG	anterior superior frontal gyrus 2	19.25	4.62	66.33
SFG	anterior superior frontal gyrus 3	4.38	9.54	66.82
SPC	anterior superior parietal cortex	16.08	-53.47	60.87
SPC	middle superior parietal cortex	13.4	-72.18	55.27
SPC	dorsal superior parietal cortex	20.68	-86.15	41.58
rMFG	superior rostral middle frontal gyrus	27.72	60.24	10.14
rMFG	inferior rostral middle frontal gyrus	34.23	54.35	-14.04