THE NEUROSCIENTIST®

Reviews at the Interface of Basic and Clinical Neurosciences

Ⓢ SAGE Publications

Indexed in MEDLINE

Spatiotemporal Dynamics of Word Processing in the Human Cortex

KSENIJA MARINKOVIĆ
Athinoula A. Martinos Center for Biomedical Imaging
Massachusetts General Hospital
Harvard Medical School

Understanding language relies on concurrent activation of multiple areas within a distributed neural network. Hemodynamic measures (fMRI and PET) indicate their location, and electromagnetic measures (magnetoencephalography and electroencephalography) reveal the timing of brain activity during language processing. Their combination can show the spatiotemporal characteristics (where and when) of the underlying neural network. Activity to written and spoken words starts in sensory-specific areas and progresses anteriorly via respective ventral ("what") processing streams toward the simultaneously active supramodal regions. The process of understanding a word in its current context peaks about 400 ms after word onset. It is carried out mainly through interactions of the temporal and inferior prefrontal areas on the left during word reading and bilateral temporo-prefrontal areas during speech processing. Neurophysiological evidence suggests that lexical access, semantic associations, and contextual integration may be simultaneous as the brain uses available information in a concurrent manner, with the final goal of rapidly comprehending verbal input. Because the same areas may participate in multiple stages of semantic or syntactic processing, it is crucial to consider both spatial and temporal aspects of their interactions to appreciate how the brain understands words. NEUROSCIENTIST 10(2):142–152, 2004. DOI: 10.1177/1073858403261018

Language is essential to our communication with others and to our conceptualization of the world in general. It is largely through language that we share our uniqueness and our ideas and that we express ourselves as individuals while crafting social relationships and conforming to the intricate web of our social milieu. Through words, we acquire a multitude of information; we articulate our thoughts, memories, and feelings; we empathize with others; we play with words and delight in mirth when sharing jokes. Because language is so fundamental yet so complex, because it interfaces with so many of our cognitive facilities, its underlying brain networks ought to be extensive and interconnected with neural systems supporting other capacities.

The earliest glimpses into this complex neural organization of language came from lesion evidence and from psycholinguistic experiments, providing a foundation for the classical language models. More recently, great advances in imaging technology have given strong momentum to the field, resulting in an upsurge in the number of studies investigating the neural basis of language.

Lesion-based "classical" models of visual language (Geschwind 1965) suggest the importance of the areas surrounding the Sylvian fissure, predominantly on the left. In this view, reading proceeds in a serial fashion starting in the visual cortex, followed by angular gyrus and Wernicke's area (access to word form and phonological conversion) and Broca's area (access to motor code). Recent neuroimaging evidence has confirmed the importance of the perisylvian region, but in addition, it has suggested other brain areas that contribute to language processing and has challenged the idea of serial processing (Mesulam 1998). Neuroimaging studies using PET and, more recently, fMRI, confirm the view that language is supported by distributed and interactive brain areas predominantly on the left (Raichle 1996; Fiez and Petersen 1998; Buckner and others 2000; Cabeza and Nyberg 2000).

Methodological Synopsis

PET and fMRI are powerful techniques able to reveal functional changes in the brain during performance of a cognitive or other task. They rely on hemodynamic changes because they measure blood-related parameters such as blood flow, blood oxygenation, and glucose

metabolism. Consequently, they measure the electrical neuronal activity only indirectly, via the accompanying hemodynamic changes. For example, when a brain region is activated by a particular task, its metabolic demands are met by increased delivery of blood and oxygen, giving rise to the fMRI signal. The exact nature of the neuronal events inducing these vascular changes is not yet understood, but their coupling is under intense investigation (Devor and others 2003). Because these vascular changes take place over seconds, a time scale much longer than the millisecond speed of neural processes underlying thought, the hemodynamic methods cannot accurately reflect the timing of the brain events. However, the spatial resolution of these methods, particularly the fMRI, is excellent and is at millimeter levels with high-field magnets. Based on their high anatomical precision, these methods can unambiguously show where the activation changes are occurring in the brain (Fig. 1).

To study the temporal characteristics ("when") of language processing, however, electromagnetic techniques offer online insight into the neuronal activity as it unfolds in real time. Electroencephalography (EEG) measures electric potentials generated by synaptic currents in the cortical layer of the brain through electrodes attached to the scalp. To relate EEG changes to the discrete events in the environment, such as words, event-related potentials (ERPs) are obtained by averaging EEG epochs time-locked to word onset. Similarly, magnetoencephalography (MEG) measures magnetic fields generated by synaptic currents through sensors in a device that resembles a large hair dryer. These methods, especially the ERPs, have been used extensively in studying language processing with millisecond precision (Halgren 1990; Osterhout and Holcomb 1995; Helenius and others 1998; Kutas and Federmeier 2000). Even though such studies have contributed immensely to our understanding of the temporal stages, or "when" of these processes, they have difficulties in unambiguously localizing "where" they are generated (Fig. 1).

Realistic models of the neurophysiology of language strive to describe the functional organization of the brain networks subserving language comprehension, as well as their anatomical distribution, roles, and hierarchical interdependence. In other words, they need to reveal the attributes of the brain regions implicated in language with respect to "what" (linguistic functions), "where" (neural regions subserving those functions), and "when" (timing of their respective contributions). Recent efforts have used a multimodal approach to integrate the respective advantages of complementary neuroimaging methods. Thus, the fMRI can be used to determine where the task-related changes are occurring, and the MEG or EEG can elucidate the timing, or when, of those changes (George and others 1995; Dale and Halgren 2001). Such integrated spatiotemporal information can reveal the dynamics of the neural circuits underlying language processing as it is occurring in the brain (see Box 1).
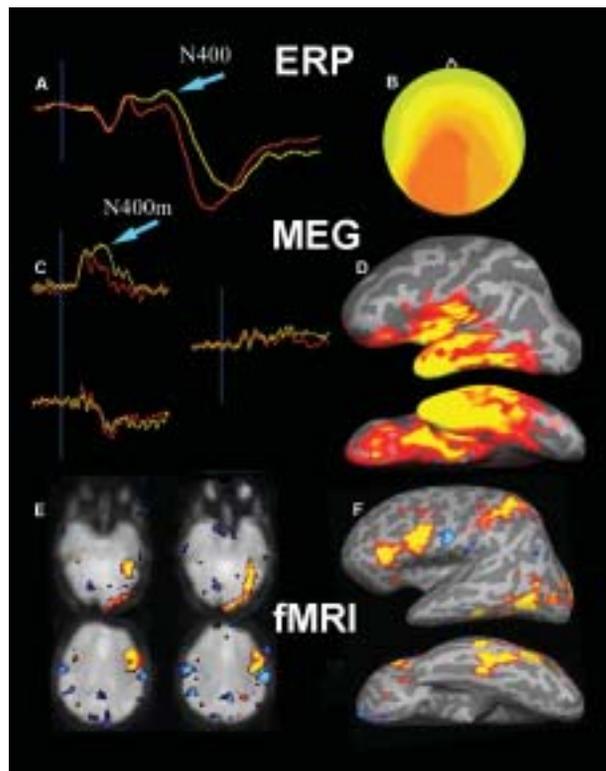
Fig. 1. Electromagnetic and hemodynamic methods. This figure illustrates responses to words measured with event-related potentials (ERPs), magnetoencephalography (MEG), and fMRI. Sample waveforms from one ERP channel (*A*) and three complementary sensors at one MEG location (*C*) reflect neural activity in real time. The N400 deflection and its magnetic equivalent, N400m, thought to index semantic integration, are marked with blue arrows. These electromagnetic methods, however, cannot unambiguously localize the generators of the activity measured on the scalp. Topographic estimates of the ERP signal are illustrated in *B* and the anatomically constrained MEG method in *D* (see Box 1 for explanation). The bottom row shows an example of the fMRI activity to words as seen in axial slices (note that the left side of the brain is on the right) (*E*) or on the cortical surface that is inflated for better visibility (*F*). The fMRI has excellent spatial resolution, revealing activations in the left inferotemporal and left prefrontal regions to words, but cannot accurately reflect the timing of their engagement (fMRI data from Oscar-Berman, with permission).

A Word's Voyage

The Ventral Modality-Specific Streams in Processing Spoken or Written Words

Ventral or "what is it" processing pathways have been described for both visual (Ungerleider and Mishkin 1982) and auditory (Rauschecker and Tian 2000) sensory modalities, based on lesion evidence, as well as the strong anatomical connections underlying the two streams in primates. Originating in their respective primary sensory areas, they extend anteriorly into the temporal cortex and the inferior prefrontal regions (Wilson and others 1993). Even though these pathways process

Box 1: Anatomically Constrained MEG

Magnetoencephalography (MEG) signals are recorded from the brain while the subject sits with his or her head inside the helmet-shaped lower end of the device containing the sensors. Electroencephalography (EEG) can be recorded concurrently. MEG and EEG directly reflect the activity of synaptic currents with a millisecond precision. However, because many different generator configurations inside the brain can yield an identical magnetic field pattern outside of the head, their spatial configuration cannot be uniquely determined. Estimating a solution requires making certain assumptions about the signal sources (Hämäläinen and others 1993). Intracranial recordings in humans and other evidence indicate that language tasks engage multiple brain regions in parallel (Halgren and others 1994; Buckner and others 2000), indicating a distributed model for the estimation. The anatomically constrained MEG uses anatomical MRI information about each subject's brain. It relies on the assumption that the synaptic potentials generating the MEG or EEG signal arise in the cortex (Dale and Sereno 1993; Dale and others 2000). Thus, the estimates are constrained to the cortical ribbon, which is usually inflated for better visibility (Fischl and others 1999). The resulting series of dynamic statistical parametric maps is similar to the maps generated for fMRI or PET data, except that they unfold in time with excellent temporal resolution in the form of "brain movies." Because of the intrinsic uncertainty of these estimates, firm inferences about the underlying neural architectonics are not justified. However, using fMRI in the same subjects and with the same task (Dale and others 2000) can further inform the inverse solution and provide independent validation of the estimated sources. The excellent spatial resolution of the fMRI complements the temporal sensitivity of the MEG and affords integrated insight into the brain networks subserving language ("where") and the timing ("when") of the involved neural components (Dale and Halgren 2001).
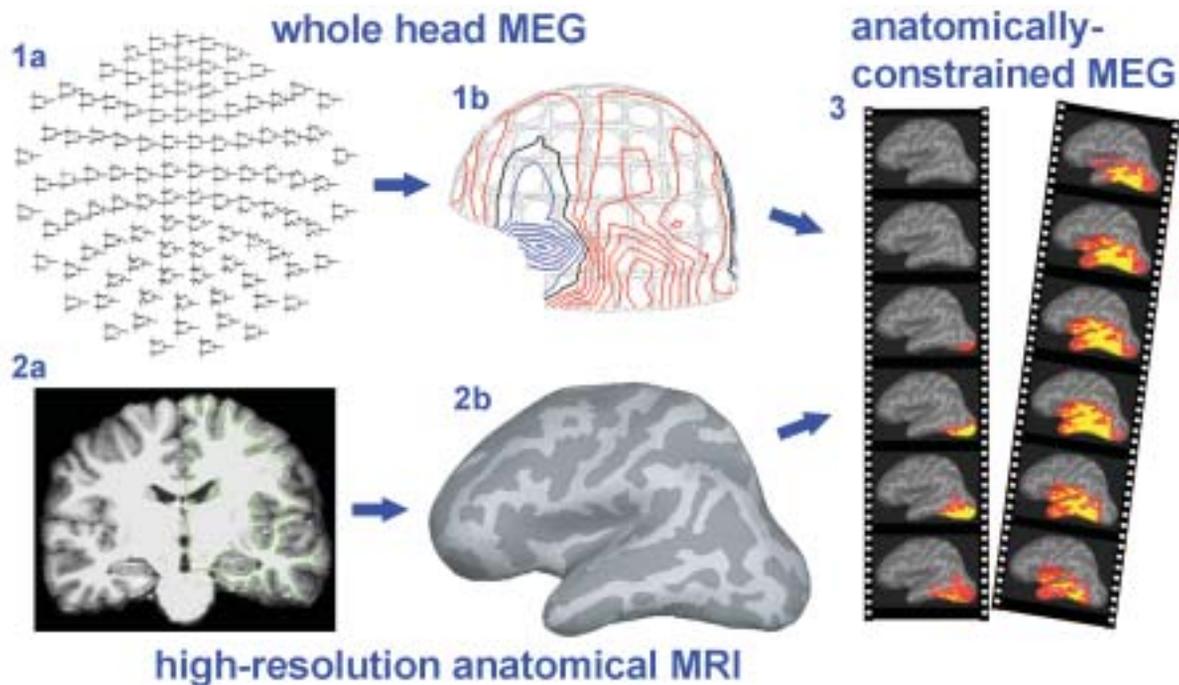
Fig. 2. The basis of the anatomically constrained magnetoencephalography (MEG) analysis method. MEG signals are recorded with a whole-head device and presented as waveforms (*1a*) or magnetic fields (*1b*) on the surface of the head. Based on high-resolution anatomical MRI (*2a*), cortical surface for each subject is reconstructed (*2b*) and used to estimate signal generators. The activity is estimated as it unfolds in time, resulting in brain movies (*3*). Because most of the cortex is hidden in folds, the reconstructed surface is inflated for better visibility of the estimated activity. Dark gray denotes the folds and the light gray the crowns of the cortical gyri.

information in a largely serial manner, there are feedback connections that affect early stages of processing in a "top-down" manner, as well as interactions between the two streams (Bullier and others 1996). The overall picture that emerges from studies using a multimodal approach and other evidence indicates that words initially activate regions of the ventral processing stream in a sequential manner. Activity starts in sensory-specific areas and progresses anteriorly toward the apparently supramodal (sensory-nonspecific) temporal and prefrontal regions, forming networks that underlie semantic and contextual integration. Figure 3 illustrates such a
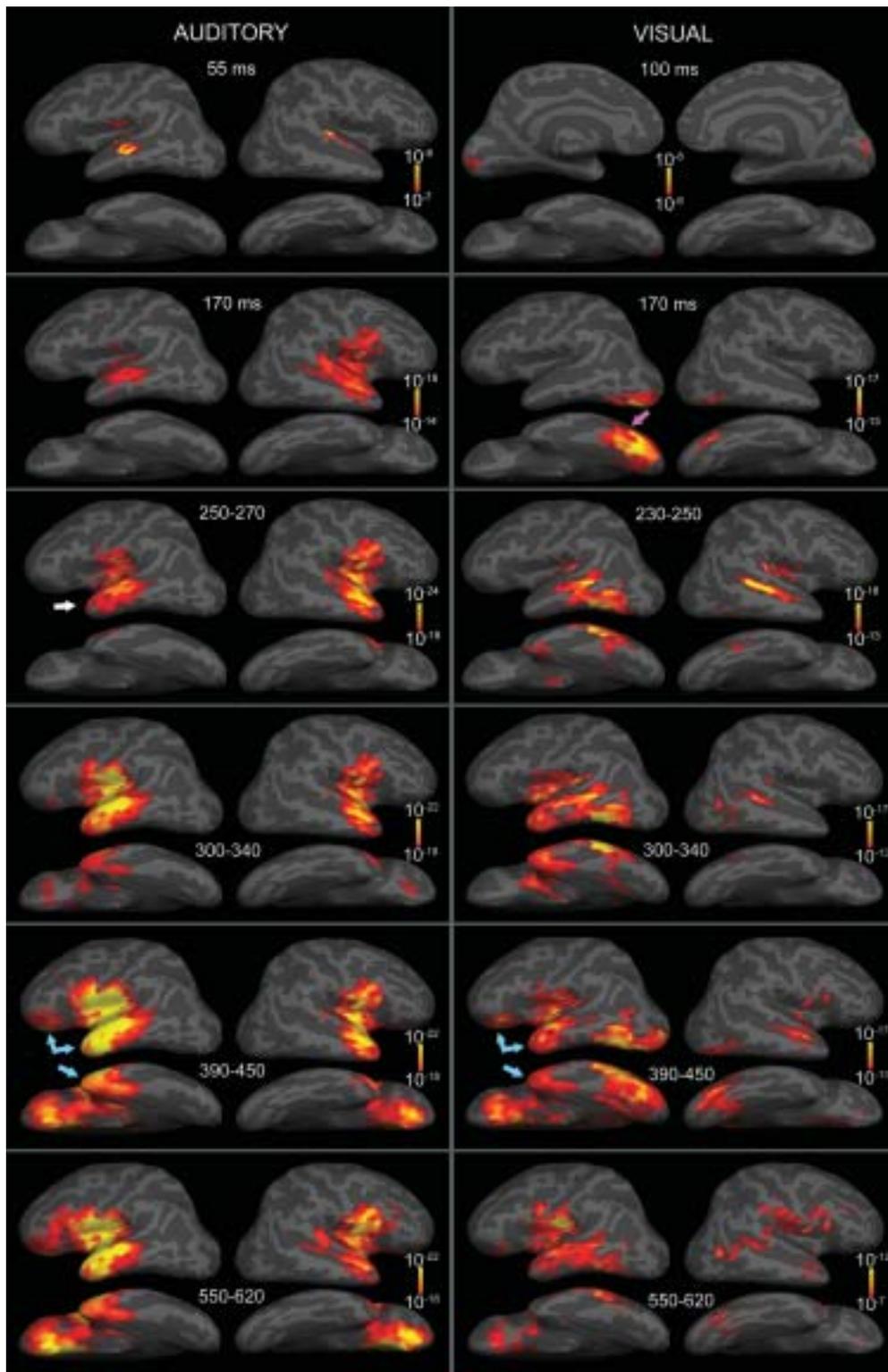
Fig. 3. Group average anatomically constrained magnetoencephalography estimated activity to spoken and written words. Subjects were presented with either spoken or written words denoting animals or objects and were asked to estimate their sizes. A comparison of the group average activation to spoken (auditory modality) or written words (visual modality) obtained in the same group of subjects is shown. Snapshots of "brain movies" at selected latencies illustrate how the activity starts in sensory-specific areas and spreads anteriorly via ventral ("what") processing streams toward the highly overlapping, apparently supramodal temporal and prefrontal regions. The process of understanding a word peaks at about 400 ms after word onset (known as N400) and results from interactive contributions of these areas. Whereas processing of written words was left lateralized, understanding spoken words engaged bilateral regions with left-dominant prefrontal activity (adapted from Marinkovic and others 2003, with permission).

progression to spoken and written words in real time, as estimated with the anatomically constrained MEG (aMEG) (Marinkovic and others 2003). As expected, the earliest activity can be seen in the respective sensory areas—the superior temporal region to spoken words at ~55 ms and the occipital area to written words at ~100 ms. In both cases, the activity proceeds in the anterior direction along the respective ventral streams. Very similar overall activation patterns have been reported with fMRI (Booth and others 2002).

Reading a Word

The activity spreads forward from the occipital area and peaks at ~170 ms in the left ventral temporo-occipital area. This corresponds to word-selective focal peaks observed in the left inferotemporal cortex at a similar latency with intracranial recordings (Halgren and others 1994; Nobre and others 1994), MEG (Dhond and others 2001; Tarkiainen and others 2002), and current source-estimated ERPs (Curran and others 1993) during linguistic tasks. Left inferotemporal area has been termed *visual word form area* because of its presumed relative specialization for prelexical processing of visual word-like stimuli (McCandliss and others 2003). However, this idea has been challenged because this region participates in a variety of tasks not involving word processing (Price and Devlin 2003). Its relative specificity has to be viewed within the context of other proposed material-specific areas in the ventral visual stream such as face-specific processing (Kanwisher and others 1997). These proposed material-specific areas in the ventral stream may encode certain visual characteristics and project them to distributed higher order association areas for further processing of their semantic, emotional, mnemonic, and other dimensions (Klopp and others 2000).

Hearing a Word

The activity to spoken words spreads anterolaterally from the primary auditory region to encompass the lateral superior temporal area and the temporal pole (Fig. 3). This has been termed the *ventral* or "what" auditory processing stream (Rauschecker and Tian 2000), as an analog to the visual domain. Spoken words are acoustically complex signals that unfold in time and in the context of ongoing speech. They are processed initially in the auditory cortex by general acoustic processors, followed by voice-specific processing in the superior temporal area bilaterally (Cabeza and Nyberg 2000) and speech-selective areas in the superior temporal sulcus of the left hemisphere (Scott and others 2000). Thus, speech recognition relies on modality-specific auditory regions at an early stage with increased reliance on the ventral stream specialized for processing speech, followed by assistance from left-dominant supramodal temporo-prefrontal areas that may facilitate word recognition in a top-down manner (see below and Box 2).

Supramodal or Modality-General Networks Underlying Access to Meaning

N400—Semantic Integration

At ~230 ms after stimulus onset, a transitional phase ensues as the modality-specific streams access the supramodal networks for semantic access and contextual integration. Using ERP methodology, language studies have described a scalp-recorded negativity peaking at ~400 ms (termed N400), which is thought to index access to meaning. A larger N400 is evoked by sentence-terminal words that do not fit the overall meaning of a sentence (e.g., I like my tea with nails) (Kutas and Hillyard 1980). Natural speech perception is a complex process as it requires parsing and integration of sounds, assembling of word sequences, and syntactic processing. Likewise, reading requires analysis of the visual word form, followed by integration on the lexical, semantic, syntactic, and discourse level. In an attempt to reduce some of this complexity, many studies have focused on studying linguistic processing at the level of a single word comprehension. The N400 amplitude is attenuated to individually presented words that are easier to process because they are repeated, are semantically primed (preceded by a related word: bread, butter), or have higher frequency of occurrence (Kutas and Federmeier 2000). Because its amplitude decreases with ease of semantic processing and integration, the N400 is commonly conceptualized as reflecting attempts to access and integrate a semantic representation into a contextual stream (Halgren 1990; Osterhout and Holcomb 1995; Hagoort, Indefrey, and others 1999). This process is not limited to spoken or written language, as similar N400 effects obtain for other stimuli that convey meaning such as American Sign Language, environmental sounds, pictures (Kutas and Federmeier 2000), or even stimuli that potentially convey meaning such as pseudowords (word-like, pronounceable but meaningless letter strings, for example, "pontel") (Halgren 1990).

N400—Generators in the Brain

Understanding the neural underpinning of N400 would get us closer to the crucial issue of understanding how the brain derives meaning out of seemingly arbitrary series of sounds or visual patterns. But where is the N400 coming from? The scalp ERPs do not have the spatial resolution to reveal the brain areas that contribute to the N400. To find out more precisely where the neural generators of the N400 are located in the brain, we need to get as close as we can to them. We need to get an "insider story" of the "when" and "where" of the language function. In special cases, it is possible to record the intracranial ERPs from the electrodes implanted in human brains during language tasks (Halgren and others 1994; Nobre and others 1994; Elger and others 1997; Marinkovic and others 2000). These recordings can unambiguously ascertain the brain regions that generate

synaptic currents, as they are sensitive to the locally generated macropotentials (Fig. 4). However, such recordings can only be done in selected patients who are implanted for clinical reasons of seizure monitoring in presurgical evaluation of epilepsy. Even though recordings can only be obtained from limited areas in the brain due to clinical constraints, a consistent picture emerges when the results are pooled across many patients (Halgren and others 1994) (Fig. 5). Main generators of the N400 to individually presented words are located in the ventral and anterior temporal lobe and in the inferior prefrontal cortex, in agreement with the aMEG studies (Dale and others 2000; Dhond and others 2001, 2003; Marinkovic and others 2003).

N400—A Generic Process of Constructing Meaning

Both spoken and written words activate overlapping regions in the left hemisphere in the temporal and prefrontal areas (Fig. 6). Furthermore, single words and sentence-terminal words evoke apparently indistinguishable N400 measured with ERP (Kutas and Federmeier 2000) and with aMEG (Halgren and others 2002). Thus, the N400 could reflect a generic process that is elicited by a potentially meaningful stimulus. Temporal, prefrontal, and anterior cingulate regions of a distributed cortical network may provide specialized contributions, with meaning resulting from pooling and a convergence of their respective inputs. Indeed, the N400 is affected by a variety of factors, including those at the lexical level such as frequency, repetition, or semantic associations, as well as those at the sentential and wider discourse levels (Kutas and Federmeier 2000). These contributions may proceed in an interactive and mutually dependent manner during the process of constructing the meaning that fits best in the given context.

When we read a word or when we hear an utterance, we derive its meaning effortlessly and automatically. In fact, we cannot choose *not to* understand a meaningful word that is communicated to us. In that sense, access to meaning may be a generic process whereby phonological, semantic, and syntactic cues are used to integrate the stimulus into the current context (Klein and others 1995). But what about the puzzling observation of a larger N400 to pseudowords than to real words (Halgren 1990)? Similarly, a stronger anterior left inferior prefrontal cortex (aLIPC) activation has been observed to pseudowords than to regular words with fMRI (Clark and Wagner 2003) and PET (Hagoort, Indefrey, and others 1999). If the N400, as subserved by the frontotemporal networks, reflects engagement of the semantic networks, why would meaningless words result in a stronger activation? It is only those fake words that are pronounceable and that conform to orthographic (the way they are written) and phonological (the way they sound) rules of the language that evoke such activation. Actually, we acquire new words continually, so many of the words currently in our vocabulary were initially experienced as pseudowords whose meaning we learned.
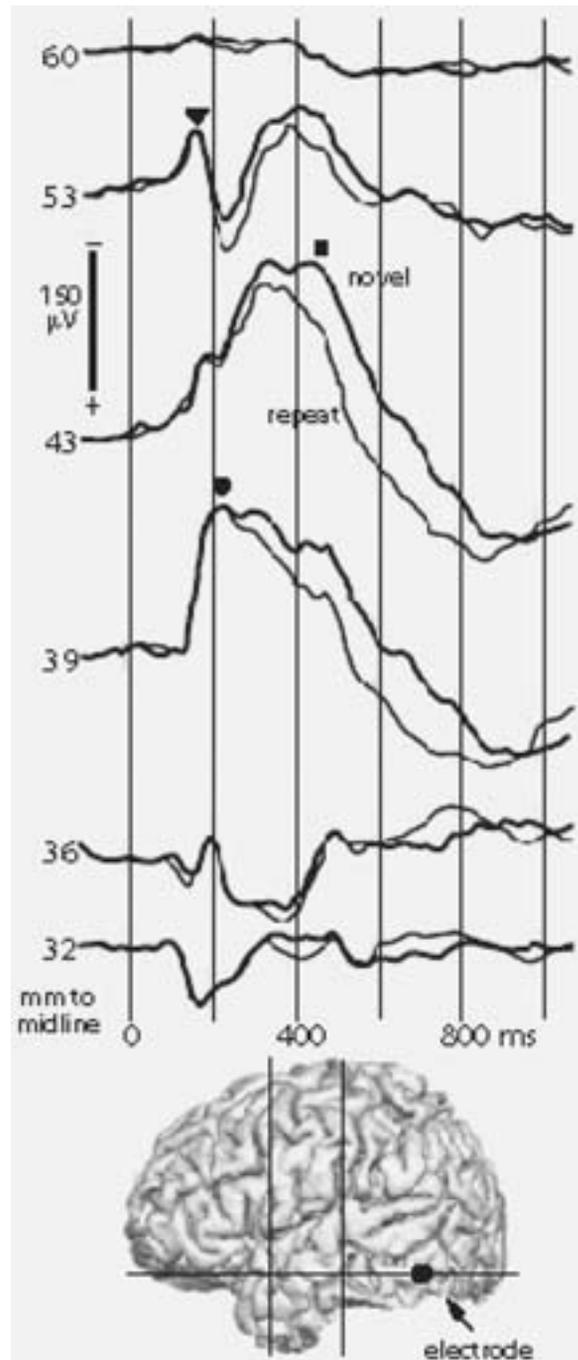
Fig. 4. Intracranial event-related potentials (ERPs) from inferotemporal cortex during a word recognition task. An electrode was implanted in the inferotemporal area with the purpose of directing surgical treatment of epilepsy. Intracranial ERPs can unambiguously identify the timing and location of the brain processes related to a task based on steep potential gradients and inversions. In this case, large and locally generated potentials were evoked during early (170 ms, marked with ▼), transitional (220 ms, marked with ●), and later integration (450 ms, marked with ■) processing stages. This evidence suggests that adjacent or overlapping regions in the inferotemporal area may play distinct roles in different aspects of verbal processing but with different timing and at different processing stages. Consequently, it is important to consider both spatial and temporal information to gain a realistic view of word processing in the brain (adapted from Halgren and others 1994, with permission).
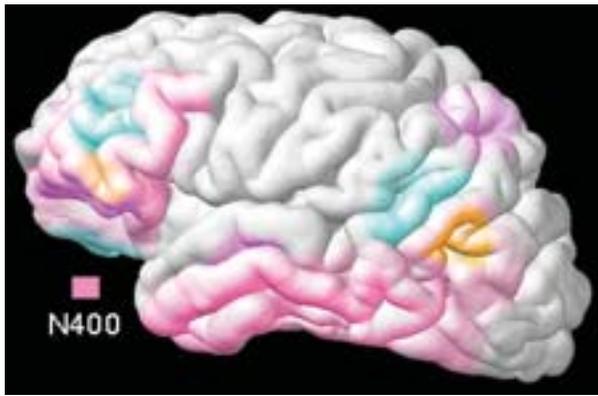
Fig. 5. Time-collapsed intracranial N400. Based on intracranial event-related potential (ERP) recordings across many patients (Halgren and others 1994), the areas contributing to the N400 (pink color) evoked by written words are located along the ventral visual stream, confirming the localization estimates obtained with noninvasive methods such as anatomically constrained magnetoencephalography. With the exception of the temporopolar region, these observations are in general agreement with the fMRI studies of language processing. Other included colors denote areas that generate other intracranial ERP deflections.
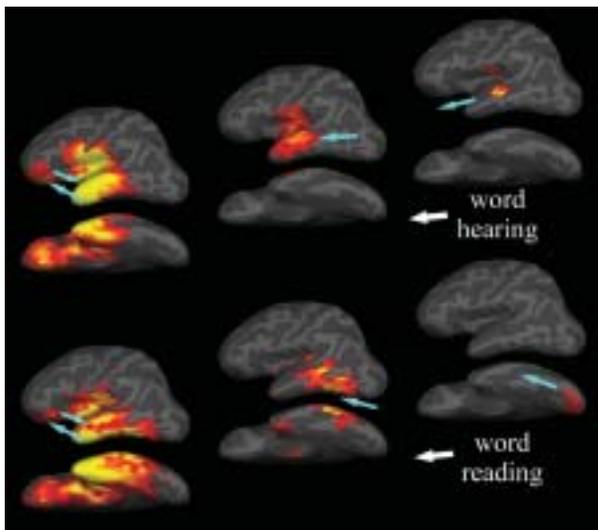
Fig. 6. A two-stage model of processing spoken and written words. Anatomically constrained magnetoencephalography and other evidence indicate that word-evoked activity starts in sensory-specific areas and progresses anteriorly toward the sensory-nonspecific regions primarily in the temporal and prefrontal regions. During the first ~200 ms, material-specific processing takes place in the areas along the ventral processing streams and is then forwarded to distributed supramodal areas for further processing. The brain seems to use all relevant information concurrently in an effort to understand verbal input as rapidly and completely as possible. Sustained interactions among multiple areas allow for the semantic, mnemonic, emotional, and contextual integration of meaning.

Hence, the increased activations may reflect an *attempt* to reach a semantic and contextual integration, and not the actual retrieval of meaning, as an outcome of such a

process. It is the engagement of this network that is reflected in the N400 and the hemodynamic activation. Different constituent structures provide important modulations to this interactive process during which semantic, mnemonic, emotional, and other aspects are integrated. Their convergence results in the construction of meaning in the appropriate context.

Temporal Lobe Contributions to the Semantic Network

In addition to intracranial recordings and the aMEG localizations, the importance of the anterior temporal lobe in semantic processing is confirmed by the syndrome of semantic dementia. Such patients gradually lose semantic knowledge about the world, and damage in their left polar and inferolateral temporal cortex correlates with their semantic impairment (Mummery and others 2000). However, studies using hemodynamic methods do not give consistent results. Whereas activations in those areas are reliably detected with PET, they are commonly absent in fMRI studies (Schacter and Wagner 1999). Loss of the fMRI signal is specific to areas near air-brain interfaces such as the temporopolar region. Its contribution to semantic processing can be seen reliably by PET only (Devlin and others 2000). The fMRI studies observe activation in the left posterior temporal regions in response to written words and activity of the bilateral temporal regions in response to spoken words, in addition to the left inferior prefrontal area. The posterior portion of the middle temporal gyrus seems particularly sensitive to semantic verbal tasks and is coactivated with the aLIPC during the retrieval of word meaning (Raichle and others 1994; Gold and Buckner 2002). The MEG source modeling approach based on one or very few focal sources has suggested the left posterior temporal (Wernicke's) area as the most likely N400 generator (Simos and others 1997; Helenius and others 1998) during language tasks. Lesion-based evidence also suggests that temporal lobe regions may be relatively specialized for different aspects of semantic memory such as retrieving information related to persons or tools (Damasio and others 1996), but the more recent neuroimaging evidence is equivocal on this issue (Thompson-Schill 2003).

Left Inferior Prefrontal Contributions to the Semantic Network

Impressive effort has been expended in the neuroimaging field to investigate the functional parcellation of the inferior prefrontal regions during language processing. This effort has been frustrated, however, by an imperfect correspondence between the tasks that were employed to engage either the phonological (such as counting syllables) or semantic (such as concrete vs. abstract judgment) aspects of word processing and the brain activation patterns. Nevertheless, neuroimaging evidence suggests that the aLIPC may be predominant in guiding semantic access, whereas the posterior LIPC might contribute

preferentially to phonological tasks (Fiez and Petersen 1998; Poldrack and others 1999; Wagner and others 2001; McDermott and others 2003). Recent evidence suggests, however, that semantic and phonological processes may be subserved by overlapping regions in the inferior prefrontal cortex rather than discrete anatomical regions (Gold and Buckner 2002; Clark and Wagner 2003). An alternative view conceptualizes the aLIPC contributions more broadly as selection among competing alternatives (Thompson-Schill 2003). In this view, the aLIPC would be more activated by a condition associated with more possible alternatives, as compared to a condition with a dominant choice, and would not be limited to semantic attributes. There is evidence of the increased aLIPC engagement during underconstrained conditions, such as in cases of multiple or ambiguous representations (Gold and Buckner 2002). The aLIPC contributions are not limited to verbal stimuli but generalize to other potentially meaningful stimuli. For instance, it has been suggested as the main candidate for the top-down facilitation of visual object recognition (Bar 2003).

It was argued above that the N400 reflects attempts to access meaning of a stimulus within a given context. Similarly, the aLIPC activation may indicate engagement of the semantic networks during an effort to comprehend a potentially meaningful stimulus. In such a scenario, the aLIPC guides access to relevant knowledge by relying on partial information available at the moment, including semantic as well as nonsemantic attributes. Its major contribution is in facilitating the convergence of semantic access in ambiguous situations. Indeed, fMRI reveals stronger aLIPC activation to words that are only weakly associated (Wagner and others 2001) or to pseudowords (Clark and Wagner 2003). The simultaneous activation of anteroventral temporal with the aLIPC during the N400 may represent a sustained interaction in search of meaning (Dale and others 2000).

Spatiotemporal Dynamics Underlying Understanding Speech

So far, we have primarily considered the neural basis for understanding written words, as they have been studied more extensively. Written words are perceptually more accessible: Letter shapes and word boundaries are perceived more clearly, and word information is available almost instantly in its entirety. On the other hand, spoken words present very different challenges to a listener as they unfold in time. The continuous spoken stream of an utterance is parsed into segments based on the auditory signal properties and is analyzed on perceptual, phonological, semantic, and prosodic levels. The process of deriving meaning from a spoken word, however, does not proceed in a serial fashion but is a result of a continuous interaction between the auditory processors that provide the bottom-up input and other areas at different points in the hierarchy that facilitate recognition in a top-down manner. Spoken words can be identified well before the end of their acoustic signal (Van Petten and others 1999),

suggesting that the semantic search starts operating with only partial input. Indeed, excellent temporal resolution of the ERP and MEG techniques provides evidence for this scenario. The N400 to spoken words peaks only slightly later than the N400 to written words, indicating that the word comprehension precedes or coincides with the end of the word acoustic signal (Van Petten and others 1999; Marinkovic and others 2003).

Neuroimaging studies using fMRI and PET clearly implicate aLIPC in the processing of spoken words, but because of the poor temporal resolution of those techniques, they cannot resolve the timing of its contribution and ascertain its role in the processing hierarchy. One way to probe its contribution to speech recognition is to investigate the effects of the phonological neighborhood density (the number of similar-sounding words) on the aLIPC activation during speech recognition (see Box 2).

Right Inferior Prefrontal Contributions

Whereas most studies show left-lateralized processing of written words, activation of the right inferior prefrontal cortex (RIPC) to spoken words is commonly observed (Buckner and others 2000; Vouloumanos and others 2001, see Fig. 3; Marinkovic and others 2003). Because of the inherent difficulty of understanding spoken words, it has been suggested that the RIPC may be engaged as a supplementary resource, especially when no context is available to prime understanding of their meaning (Friederici and others 2000). The RIPC may contribute to semantic retrieval and can facilitate comprehension through prosody (George and others 1996). There is mounting evidence that the right prefrontal cortex participates in certain aspects of contextual integration. For example, it may contribute to understanding words that have weak semantic associations (Booth and others 2002), which agrees with the finding that patients with lesions in the right hemisphere have trouble understanding jokes or metaphors (Brownell and others 1990). Jokes engage a host of linguistic (semantic, syntactic), mnemonic (working memory and word retrieval), emotional (judging word valence), and higher order integrative processes that allow us to understand their non-literal meaning. Indeed, jokes selectively engage right prefrontal cortex following the N400, during the phase of retrieving the alternate meaning so that the "twist" can be incorporated into the joke context (Marinkovic and others 2001).

Syntactic Processing

Language entails much more complexity than understanding individual words, as they are arranged in sentences and discourse according to syntactic rules. ERP studies show that syntactic violations or ambiguities sometimes elicit an early, often left-lateralized anterior negativity, which can start as early as 150 ms, though commonly between 300 ms and 500 ms after stimulus onset, hypothesized to represent a disrupted initial structural analysis of the incoming words (Friederici 1997).

Box 2: Prefrontal Activity to Spoken Words

The initial segment of a spoken word plays a special role in understanding speech based on the number of lexical competitors. For example, on hearing "pa-" /pa/ as the initial segment of a word, a number of competitors can be invoked such as pace, pay, pain, and so forth. Thus, words that share the initial phoneme with fewer words (low-density neighborhood) are processed faster than words that share the initial segment with many words (high density) because presumably the right "match" is accessed more easily (Vitevitch 2002). We have studied this phenomenon with anatomically constrained magnetoencephalography in a semantic task using spoken words. As illustrated in Figure 7, group average anterior left inferior prefrontal cortex (aLIPC) activation was significantly stronger to high-density words already at 240 ms after the word onset. This result is consistent with an increased need for aLIPC contribution in underconstrained conditions where more completions are possible (Gold and Buckner 2002). An early (~240 ms) aLIPC activation in the auditory modality may represent facilitation of word comprehension by selective top-down influences. Because word meaning cannot be accessed on hearing the first phoneme, aLIPC may

mediate a top-down semantic search based on results of the evolving phonological analysis. This observation supports previous accounts of spoken word recognition (Marslen-Wilson 1987; Hagoort and Brown 2000) whose main idea is that the initial phoneme analysis activates representations of a cohort of possible words. As the sound input unfolds in time, words that continue matching the input remain in the "contest," whereas those that no longer match are eliminated, eventually yielding the best candidate. Continual acoustic input provides the "bottom-up" iterative honing of that list, whereas the higher association areas provide a "top-down" facilitation of this evolving process, resulting in word comprehension and N400. This spatiotemporal profile of activation suggests that the brain uses all resources and input as soon as it becomes available. Most of the network elements are engaged by ~200 ms and thus could continue to exert a top-down influence over subsequent stages of word input and comprehension. The LIPC and left temporal areas have the appropriate connections and cognitive correlates to provide the neural basis for those contributions.
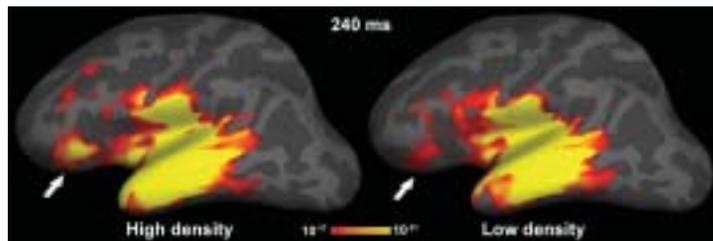
Fig. 7. Early left prefrontal activity to spoken words. Group average anatomically constrained magnetoencephalography to spoken words during a semantic task at 240 ms after acoustic onset. Subjects heard a series of words denoting objects or animals and were asked to judge their sizes. Words that share the initial phoneme with many words (high-density neighborhood) evoke more left prefrontal activation at 240 ms than the words that have fewer competitors. The anterior left inferior prefrontal cortex may provide the top-down facilitation during understanding of spoken words, in accord with other evidence showing its contribution to ambiguous situations.

Alternatively, left-lateralized anterior negativity may reflect working memory load during sentence processing (Kluender and Kutas 1993). Another ERP deflection has been associated with syntactic anomalies or ambiguities: a sustained positivity occurring between 500 ms and 1200 ms after stimulus onset, termed *P600* or *syntactic positive shift* (Hagoort, Brown, and others 1999). The P600 is evoked by a range of changes in sentence structure, including syntactic anomalies (words that violate grammatical structure), syntactic ambiguity (words that clarify ambiguous sentence structure), or sentence complexity (Friederici 1997; Hagoort, Brown, and others 1999). The consensus on the functional role of the P600 has not been reached. It has been hypothesized to index syntax-specific "revision" or "repair" processes that are engaged when the syntactic rules are violated (Friederici 1997), but it has also been suggested to represent a gen-

eral process of reanalysis that is not specific for language (Coulson and others 1998). For example, P600 is elicited by musical chords that do not fit into the musical phrase (Besson and Schon 2001).

Even though the ERP studies suggest that the syntactic and semantic processes may be subserved by distinct generators, a review of the PET and fMRI studies (Kaan and Swaab 2002) indicates that syntactic processing evokes activation in fronto-temporal regions that overlap with semantic or other cognitive functions. The apparent lack of regional specialization for syntax may be indicative of the need to consider both spatial and temporal aspects of processing in the context of distributed networks. Contributing cortical regions may play distinct roles in different aspects of processing but with different timing and at different processing stages (Kuperberg and others 2003). Additional contributions may be provided

by subcortical structures such as basal ganglia (Friederici and others 2003). Alternatively, some key processes in syntactic processing may be occurring in structures such as the basal ganglia that lack the spatial distribution of synaptic elements necessary to produce propagating electromagnetic signals.

Conclusion

After the initial modality-specific processing stage, word processing is subserved by distributed brain regions that are simultaneously active for a protracted period of time. They mainly comprise the temporal and inferior prefrontal areas on the left during word reading and bilateral perisylvian regions during processing speech. This activation culminates in a generic process of word comprehension peaking at about 400 ms (N400). Their relative contributions are modulated by contextual and task-related demands such as difficulty, sensory modality, semantic coherence, priming, and so forth. Neurophysiological evidence suggests that lexical access, semantic associations, and contextual integration are simultaneous and indeed may be inseparable. One plausible interpretation is that the brain uses any information that is available at any given point in time in a concurrent manner, with the final goal of rapidly comprehending the verbal input it was presented with. The fMRI and PET have not yet clearly revealed distinct roles for different areas in supporting different aspects of language. Because the same areas may contribute to multiple stages of processing, the nature of their contributions to language cannot be determined solely from techniques with low temporal resolution. The spatiotemporal dynamics of their participation and their interactions may be elucidated in combination with temporally sensitive methods that can provide the timing aspects of such concerted events.

References

Bar M. 2003. A cortical mechanism for triggering top-down facilitation in visual object recognition. J Cogn Neurosci 15(4):600–9.

Besson M, Schon D. 2001. Comparison between language and music. Ann N Y Acad Sci 930:232–58.

Booth JR, Burman DD, Meyer JR, Gitelman DR, Parrish TB, Mesulam MM. 2002. Modality independence of word comprehension. Hum Brain Mapp 16(4):251–61.

Brownell HH, Simpson TL, Bihrle AM, Potter HH, Gardner H. 1990. Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. Neuropsychologia 28(4):375–83.

Buckner RL, Koutstaal W, Schacter DL, Rosen BR. 2000. Functional MRI evidence for a role of frontal and inferior temporal cortex in amodal components of priming. Brain 123(pt 3):620–40.

Bullier J, Schall JD, Morel A. 1996. Functional streams in occipito-frontal connections in the monkey. Behav Brain Res 76(1–2): 89–97.

Cabeza R, Nyberg L. 2000. Imaging cognition II: an empirical review of 275 PET and fMRI studies. J Cogn Neurosci 12(1):1–47.

Clark D, Wagner AD. 2003. Assembling and encoding word representations: fMRI subsequent memory effects implicate a role for phonological control. Neuropsychologia 41(3):304–17.

Coulson S, King JW, Kutas M. 1998. Expect the unexpected: event-related brain responses to morphosyntactic violations. Language and Cognitive Processes 13:653–72.

Curran T, Tucker DM, Kutas M, Posner MI. 1993. Topography of the N400: brain electrical activity reflecting semantic expectancy. Electroencephalogr Clin Neurophysiol 88(3):188–209.

Dale AM, Halgren E. 2001. Spatiotemporal mapping of brain activity by integration of multiple imaging modalities. Curr Opin Neurobiol 11(2):202–8.

Dale AM, Liu AK, Fischl BR, Buckner RL, Belliveau JW, Lewine JD, and others. 2000. Dynamic statistical parametric mapping: combining fMRI and MEG for high-resolution imaging of cortical activity. Neuron 26(1):55–67.

Dale AM, Sereno MI. 1993. Improved localization of cortical activity by combining EEG and MEG with MRI cortical surface reconstruction: a linear approach. J Cogn Neurosci 5:162–76.

Damasio H, Grabowski TJ, Tranel D, Hichwa RD, Damasio AR. 1996. A neural basis for lexical retrieval. Nature 380(6574):499–505.

Devlin JT, Russell RP, Davis MH, Price CJ, Wilson J, Moss HE, and others. 2000. Susceptibility-induced loss of signal: comparing PET and fMRI on a semantic task. Neuroimage 11(6 pt 1):589–600.

Devor A, Dunn AK, Andermann ML, Ulbert I, Boas DA, Dale AM. 2003. Coupling of total hemoglobin concentration, oxygenation, and neural activity in rat somatosensory cortex. Neuron 39(2):353–9.

Dhond RP, Buckner RL, Dale AM, Marinkovic K, Halgren E. 2001. Sequence of brain activity underlying word-stem completion. J Neurosci 21(10):3564–71.

Dhond RP, Marinkovic K, Dale AM, Witzel T, Halgren E. 2003. Spatiotemporal maps of past-tense verb inflection. Neuroimage 19(1):91–100.

Elger CE, Grunwald T, Lehnertz K, Kutas M, Helmstaedter C, Brockhaus A, and others. 1997. Human temporal lobe potentials in verbal learning and memory processes. Neuropsychologia 35(5):657–67.

Fiez JA, Petersen SE. 1998. Neuroimaging studies of word reading. Proc Natl Acad Sci U S A 95(3):914–21.

Fischl B, Sereno MI, Dale AM. 1999. Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. Neuroimage 9(2):195–207.

Friederici AD. 1997. Neurophysiological aspects of language processing. Clin Neurosci 4(2):64–72.

Friederici AD, Kotz SA. 2003. The brain basis of syntactic processes: functional imaging and lesion studies. Neuroimage 20 Suppl 1:S8–17.

Friederici AD, Meyer M, von Cramon DY. 2000. Auditory language comprehension: an event-related fMRI study on the processing of syntactic and lexical information. Brain Lang 74(2):289–300.

George JS, Aine CJ, Mosher JC, Schmidt DM, Ranken DM, Schlitt HA, and others. 1995. Mapping function in the human brain with magnetoencephalography, anatomical magnetic resonance imaging, and functional magnetic resonance imaging. J Clin Neurophysiol 12(5):406–31.

George MS, Parekh PI, Rosinsky N, Ketter TA, Kimbrell TA, Heilman KM, and others. 1996. Understanding emotional prosody activates right hemisphere regions. Arch Neurol 53(7):665–70.

Geschwind N. 1965. Disconnexion syndromes in animals and man. I. Brain 88(2):237–94.

Gold BT, Buckner RL. 2002. Common prefrontal regions coactivate with dissociable posterior regions during controlled semantic and phonological tasks. Neuron 35(4):803–12.

Hagoort P, Brown CM. 2000. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38(11):1518–30.

Hagoort P, Brown CM, Osterhout L. 1999. The neurocognition of syntactic processing. In: Brown CM, Hagoort P, editors. The neurocognition of language. Oxford: Oxford University Press. p 273–316.

Hagoort P, Indefrey P, Brown C, Herzog H, Steinmetz H, Seitz RJ. 1999. The neural circuitry involved in the reading of German words and pseudowords: a PET study. J Cogn Neurosci 11(4):383–98.

Halgren E. 1990. Insights from evoked potentials into the neuropsychological mechanisms of reading. In: Scheibel AB, Wechsler AF, editors. Neurobiology of higher cognitive function. New York: Guilford. p 103–50.

Halgren E, Baudena P, Heit G, Clarke JM, Marinkovic K. 1994. Spatio-temporal stages in face and word processing. I. Depth-recorded potentials in the human occipital, temporal and parietal lobes [corrected] [published erratum appears in J Physiol Paris 1994;88(2):following 151]. J Physiol Paris 88(1):1–50.

Halgren E, Dhond RP, Christensen N, Van Petten C, Marinkovic K, Lewine JD, and others. 2002. N400-like magnetoencephalography responses modulated by semantic context, word frequency, and lexical class in sentences. Neuroimage 17(3):1101–16.

Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. 1993. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. Reviews of Modern Physics 65(2):413–97.

Helenius P, Salmelin R, Service E, Connolly JF. 1998. Distinct time courses of word and context comprehension in the left temporal cortex. Brain 121(pt 6):1133–42.

Kaan E, Swaab TY. 2002. The brain circuitry of syntactic comprehension. Trends Cogn Sci 6(8):350–6.

Kanwisher N, McDermott J, Chun MM. 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J Neurosci 17(11):4302–11.

Klein D, Milner B, Zatorre RJ, Meyer E, Evans AC. 1995. The neural substrates underlying word generation: a bilingual functional-imaging study. Proc Natl Acad Sci U S A 92(7):2899–903.

Klopp J, Marinkovic K, Chauvel P, Nenov V, Halgren E. 2000. Early widespread cortical distribution of coherent fusiform face selective activity. Hum Brain Mapp 11(4):286–93.

Kluender R, Kutas M. 1993. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. J Cogn Neurosci 5(2):196–214.

Kuperberg GR, Holcomb PJ, Sitnikova T, Greve D, Dale AM, Caplan D. 2003. District patterns of neural modulation during the processing of conceptual and syntactic anomalies. J Cogn Neurosci 15(2):272–93.

Kutas M, Federmeier KD. 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends Cogn Sci 4(12):463–70.

Kutas M, Hillyard SA. 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207(4427):203–5.

Marinkovic K, Dhond RP, Dale AM, Glessner M, Carr V, Halgren E. 2003. Spatiotemporal dynamics of modality-specific and supramodal word processing. Neuron 38(3):487–97.

Marinkovic K, Glessner M, Dale AM, Halgren E. 2001. Humor and incongruity: anatomically-constrained MEG. Prog 742:7.

Marinkovic K, Trebon P, Chauvel P, Halgren E. 2000. Localised face processing by the human prefrontal cortex: face-selective intracerebral potentials and post-lesion deficits. Cognitive Neuropsychology 17:187–99.

Marslen-Wilson WD. 1987. Functional parallelism in spoken word-recognition. Cognition 25(1–2):71–102.

McCandliss BD, Cohen L, Dehaene S. 2003. The visual word form area: expertise for reading in the fusiform gyrus. Trends Cogn Sci 7(7):293–9.

McDermott KB, Petersen SE, Watson JM, Ojemann JG. 2003. A procedure for identifying regions preferentially activated by attention to semantic and phonological relations using functional magnetic resonance imaging. Neuropsychologia 41(3):293–303.

Mesulam MM. 1998. From sensation to cognition. Brain 121(pt 6):1013–52.

Mummery CJ, Patterson K, Price CJ, Ashburner J, Frackowiak RS, Hodges JR. 2000. A voxel-based morphometry study of semantic dementia: relationship between temporal lobe atrophy and semantic memory. Ann Neurol 47(1):36–45.

Nobre AC, Allison T, McCarthy G. 1994. Word recognition in the human inferior temporal lobe. Nature 372(6503):260–3.

Osterhout L, Holcomb P. 1995. Event-related potentials and language comprehension. In: Rugg MD, Coles MGH, editors. Electrophysiology of mind: event-related brain potentials and cognition. Oxford: Oxford University Press.

Poldrack RA, Wagner AD, Prull MW, Desmond JE, Glover GH, Gabrieli JD. 1999. Functional specialization for semantic and phonological processing in the left inferior prefrontal cortex. Neuroimage 10(1):15–35.

Price CJ, Devlin JT. 2003. The myth of the visual word form area. Neuroimage 19(3):473–81.

Raichle ME. 1996. What words are telling us about the brain. Cold Spring Harb Symp Quant Biol 61:9–14.

Raichle ME, Fiez JA, Videen TO, MacLeod AM, Pardo JV, Fox PT, and others. 1994. Practice-related changes in human brain functional anatomy during nonmotor learning. Cereb Cortex 4(1):8–26.

Rauschecker JP, Tian B. 2000. Mechanisms and streams for processing of "what" and "where" in auditory cortex. Proc Natl Acad Sci U S A 97(22):11800–6.

Schacter DL, Wagner AD. 1999. Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. Hippocampus 9(1):7–24.

Scott SK, Blank CC, Rosen S, Wise RJ. 2000. Identification of a pathway for intelligible speech in the left temporal lobe. Brain 123(pt 12):2400–6.

Simos PG, Basile LF, Papanicolaou AC. 1997. Source localization of the N400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. Brain Res 762(1–2):29–39.

Tarkiainen A, Cornelissen PL, Salmelin R. 2002. Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. Brain 125(pt 5):1125–36.

Thompson-Schill SL. 2003. Neuroimaging studies of semantic memory: inferring "how" from "where." Neuropsychologia 41(3):280–92.

Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. Analysis of visual behavior. Cambridge (MA): MIT Press. p 549–86.

Van Petten C, Coulson S, Rubin S, Plante E, Parks M. 1999. Time course of word identification and semantic integration in spoken language. J Exp Psychol Learn Mem Cognit 25:394–417.

Vitevitch MS. 2002. Influence of onset density on spoken-word recognition. J Exp Psychol Hum Percept Perform 28(2):270–78.

Vouloumanos A, Kiehl KA, Werker JF, Liddle PF. 2001. Detection of sounds in the auditory stream: event-related fMRI evidence for differential activation to speech and nonspeech. J Cogn Neurosci 13(7):994–1005.

Wagner AD, Pare-Blagoev EJ, Clark J, Poldrack RA. 2001. Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. Neuron 31(2):329–38.

Wilson FA, Scalaidhe SP, Goldman-Rakic PS. 1993. Dissociation of object and spatial processing domains in primate prefrontal cortex. Science 260(5116):1955–8.