Psycholinguistics Electrified II (1994-2005).

Marta Kutas[1], Cyma K. Van Petten[2], Robert Kluender[3]

To appear in M.A. Gernsbacher & M. Traxler (Eds.), *Handbook of Psycholinguistics, 2nd edition.* New York: Elsevier Press.

1Departments of Cognitive Science and Neurosciences, University of California, San Diego, La Jolla, California
2Department of Psychology, University of Arizona, Tucson, Arizona
3Department of Linguistics, University of California, San Diego, La Jolla, California

Acknowledgements: M. Kutas (a Lady Davis fellow at Hebrew University) was supported from grants HD22614 and AG08313. We thank E. DeOchoa and T. Urbach for assistance.

Correspondence to:
Dr. Marta Kutas
until Fall, 2006:
Department of Psychology
Hebrew University, Mt. Scopus
Jerusalem, Israel
mkutas@mscc.huji.ac.il
972-(0)2-588-0154

from September, 2006:
Department of Cognitive Science
University of California, San Diego
La Jolla, California 92093-0515
mkutas@ucsd.edu
858-534-7450

1. INTRODUCTION

Since the first edition of the *Handbook of Psycholinguistics* in 1994, investigations of language processing via electromagnetic recordings have proliferated beyond the possibility of coverage in any single chapter. Our aim here is to offer a sampling of the more seminal, influential, and controversial event-related brain potential (ERP) studies within the psychology of language, focusing on the last decade. Out of necessity, we restrict the review to studies of healthy young adults as this segment of the population is the typical baseline against which to assess results from infants, children, middle-aged and older adults, and individuals with neurological or psychiatric disorders. Length limitations also forced us to skip studies of speech perception and production, and those bearing on the automaticity of semantic processing, topics we plan to address in some future venue.

In 1994, there were only two dominant noninvasive techniques to offer insight about the functional organization of language from its brain bases: the behavior of brain-damaged patients (neuropsychology), and event-related brain potentials (ERPs). Positron emission tomographic and magnetoencephalographic (MEG) measures were just beginning to contribute to our understanding. Over the ensuing decade-plus, these have been joined by functional magnetic resonance imaging, transcranial magnetic stimulation, event related spectral changes in the electroencephalogram, and noninvasive optical imaging (see Gratton & Fabiani, 2001; Gratton, Fabiani, Elbert, & Rockstroh, 2003 for review of the last and newest technique). As outlined below, three of these methods are closely related in their neural and physical bases: ERPs, event-related frequency changes in the electroencephalogram, and magnetoencephalography. After that brief review of the neural bases of these methods, we devote a modicum of attention to the latter two methods and chiefly focus on ERP studies of language processing. The remainder of the review is then devoted to four major domains of language processing: visual word recognition, basic semantic processing, higher-level semantic processing, and syntax and morphology.

2. ELECTROMAGNETIC MEASURES OF BRAIN ACTIVITY

2.1 Neural activity and the electrotroencephalogram

Interactions between neurons are the essence of brain activity. These interactions consist of current flow – the movement of charged ions – across cell membranes, such that the direction and magnitude of current flow in one neuron depends on the neurons it communicates with. A recording electrode close to a neuron can detect one sort of rapid change in voltage (or potential) caused by rapid changes in current flow: the action potential that causes neurotransmitter release in the vicinity of another neuron. Placing an electrode close to a single neuron is too invasive for use in healthy humans. After neurotransmitter is released and bound by other neurons, the result is a change in current flow

across the membranes of those other post-synaptic cells. These small changes in current flow can sum in different ways depending upon the number, location and timing of active synapses on the target neuron, as well as on the inward or outward directions of current flow. The target neuron may fire its own action potentials, reduce its firing rate, or show no change in firing rate but become more or less responsive to future inputs. An immediate change in activity due to synaptic input may also be accompanied by changes in gene transcription resulting in long lasting structural modifications of the neuron. Whatever the short or long-term outcome, current flow around synapses *is* the currency of neural communication.

Like action potentials, the small changes in voltage around active synapses can be recorded by nearby electrodes. The summed activity of many synapses on many neighboring neurons (called a *field* potential) can also be recorded by a pair of electrodes – one placed directly in neural tissue and one some distance away. Perhaps surprisingly, summated synaptic potentials can also be recorded outside the head, noninvasively from electrodes placed on the scalp; this record of fluctuating voltage across time is the *electroencephalogram (EEG).* The amplitude of the EEG is considerably smaller than invasively-recorded field potentials because the skull is a strong electrical insulator. Like field potentials, the amplitude and polarity of the EEG depends on the number and amplitude of the contributing synaptic potentials, on whether current is flowing into or out of cells (i.e., movement of positive or negative ions, excitatory or inhibitory synaptic potentials), and on the geometric relationship between the synapses and electrode (i.e., current flow toward versus away from the electrode, or both toward and away, which will lead to cancellation of the opposing signals; Nunez, 1981). Finally, any record of electrical potential (voltage) consists of the difference between two locations (like the positive and negative poles of a battery), so that the polarity and spatial distribution of the EEG across the head depends on what pairs of sites are chosen. Most typically, a single location or pair of locations that are somewhat more insulated from brain activity – such as that provided by the thick mastoid bones behind the ears for mastoid or earlobe sites, or the air-filled sinuses for the nosetip are used for *reference* for each scalp site, although other references are possible. Given the low electrical conductivity of the skull, electrical potentials recorded from the scalp must reflect the activity of large numbers of neurons, estimated 1000 to 10,000 for the smallest signals recorded. Cortical pyramidal cells are likely to dominate the EEG signal, because they are the largest and most numerous cell type, and because their dendritic processes are spatially parallel to their neighbors; such an organization leads to summation of the small electrical fields generated by each active synapse.

2.2 Event-related brain potentials (ERPs)

The main emphasis of this chapter, however, is on electrical or magnetic brain activity that is synchronized to some external event (i.e., an event related brain potential or ERP). At the scalp an ERP (5-10uV) is substantially smaller in amplitude than the background EEG (50-100 uV) and is, therefore, generally extracted by computer averaging. This involves recording ERPs to repeated

presentations of conceptually, if not physically, similar stimuli. Voltage fluctuations generated by neurons, unrelated (or at least not phase-locked) to the processing of the stimuli of interest will be random with respect to stimulus onset time and thus cancel each other, leaving a record of event-related activity. The number of stimuli needed for a reliable average is a function of the amplitude of the ERP component and the question under study: the smaller the component, the more trials that are needed to extract it from the spontaneous EEG (“noise”).

The major statistical assumption in averaging is that the signal is indeed time locked to the averaging trigger whereas the “noise” is not. For the early "sensory" portion of the ERP, the time-locking assumption is well-supported. In the case of later portions of the ERP which are instead elicited by higher-level "cognitive" analyses of the stimulus, the latency of the signal may not be invariant with regard to stimulus onset on a trial by trial basis, but there are techniques to correct for misalignment (see, e.g., Handy, 2005).

2.2.1 Peaks and components

The ERP waveform of voltage plotted against post-stimulus time consists of a series of positive and negative peaks; these are typically compared to a pre-stimulus baseline, that is a short (100-200 ms) record of activity (or preferably inactivity) immediately preceding each experimental stimulus, although other baselines are possible. Voltages are thus only negative or positive with respect to the baseline.

The ERP peaks are typically labeled according to their polarity (negative [N] or positive [P]) and latency in milliseconds relative to stimulus onset (e.g., N100, P230, P300). Occasionally, peaks are designated by their polarity and ordinal position in the waveform (e.g., N1, P1, N2). Sometimes, the labels denote a functional description (e.g., mismatch negativity or MMN) or refer to its presumed neural generator (e.g., auditory brainstem response) or its most reliable scalp location (e.g., LAN or left anterior negativity). The mix of descriptive and functional labels brings us to the distinction between an ERP peak, readily observed by the eye, and the more abstract concept of a "component" (see Allison, Wood & McCarthy, 1986; Donchin, Ritter & McCallum, 1978).

The underlying notion of a ‘component’ is clear: the processing of any external stimulus occurs over time, so that different parts of the nervous system with different functions are likely involved at different time points. The ERP is a record of this neural processing, so that different temporal intervals of the waveform are likely to reflect different anatomical locations and different functional processes, although any particular interval may reflect more than one brain region/functional process. One set of factors, visible in a single ERP waveform, bears some, usually unknown, relationship to the anatomy of the underlying neural generators: polarity, latency from stimulus onset, and relative amplitude across scalp locations (i.e., scalp distribution). Most commonly, the ERP waveform is reduced to a series of peak or mean amplitude measurements relative to a pre-stimulus baseline. There also exist a number of algorithms for decomposing the ERP waveform into some weighted average of subcomponents, which may provide truer reflections of the neural components

responsible for particular component psychological processes. Two such decomposition techniques are Independent Component Analysis (Delorme & Makeig, 2004; Makeig, Bell, Jung, & Sejnowski, 1996) and spatial or spatio-temporal Principal Component Analysis (Beauducel & Debener, 2003; Spencer, Dien & Donchin, 2001).

The second set of factors critical for identifying some portion of the ERP as a unitary component involves comparisons between two or more experimental conditions to determine which manipulations influenced a particular temporal region of the waveform. Susceptibility to some experimental manipulation is essential for component identification, making "peak" or for that matter ICA or PCA component nonsynonymous with "component". The functional characterization offered by psychologists and the neural characterization that might be offered by a physiologist are thus all part of the definition of an ERP component, under ideal circumstances. However, circumstances are rarely ideal. A functional characterization is most easily carried out via experiments involving large numbers of healthy human subjects, whereas a neural characterization typically requires converging evidence from animal models, neurological patients undergoing invasive clinical procedures, and scalp recordings from patients with defined brain damage (Arezzo, Vaughan, Kraut, Steinschneider & Legatt, 1986; Buchwald & Squires, 1982; Halgren, 1990; Knight, Scabini, Woods & Clayworth, 1989; McCarthy, Wood, Williamson & Spencer, 1989; Pineda, Swick & Foote, 1991). In this review, we focus on functional characterizations, and discuss psycholinguistic manipulations implemented by varying the stimuli and/or the instructions to the participants, referring to anatomical generators when known.

As a general rule, the amplitudes, latencies and scalp distributions of the earlier ERP components (with latencies <100 ms) are highly reproducible across sessions within an individual (Halliday, 1982). Moreover, systematic variations in the physical parameters of the evoking stimulus (e.g., intensity, frequency, duration) lead to predictable changes in these early components reflecting the altered activation of sensory pathways. Hence, the earlier evoked components are considered to be "exogenous" or stimulus bound; they are generally relatively impervious to an individual’s state of alertness or attentiveness. This invariance in the face of changing psychological states makes them an excellent diagnostic tool for certain sensory and neurological disorders (Chiappa, 1983).

For psycholinguistic purposes, the more informative brain waves are the so-called endogenous components, which may precede or follow a triggering event by hundreds of milliseconds. An "event" in this case refers to a stimulus, a response, a voluntary movement, or a cognitive operation for which an external timing marker can be specified. The relative (although not total) insensitivity of endogenous components to variations in the physical stimulus parameters contrasts with their exquisite responsivity to task demands, instructions, and subjects' intentions, decisions, expectancies, strategies, mental set, and so on. In other words, endogenous ERP components are not "evoked" by a stimulus but are elicited by the perceptual and cognitive operations that are engendered by that stimulus. The same physical stimulus may or may not be followed by a particular endogenous component depending on how the subject chooses to

process it. The term "late" component is often used interchangeably with “endogenous” component because most of these potentials occur with a latency beyond 100 ms, although some earlier potentials can be modulated by cognitive processes.

2.3 Magnetoencephalography and event-related magnetic fields

Current flow in the brain produces small magnetic fields in addition to the voltage fields recorded as EEG. Epochs of the magnetoencephalogram (MEG) following stimulus presentation can be averaged to derive the event-related magnetic field. Although both the raw MEG and the event-related fields resemble their electrical counterparts in many ways, some physical differences make the anatomical origins of the magnetic signals easier to localize (while preserving the same temporal resolution as electrical signals; for review see Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993). One reason is that although the skull is a very good electrical insulator and thus imposes a spatial blurring between the brain and the scalp, bone is magnetically transparent. The magnetic fields recorded just outside the head are also more strongly influenced by the geometrical orientation of intracranial current flow. The latter fact is a mixed blessing. On the one hand, the convoluted shape of the cortex ensures that current flow in different sulci and gyri will have distinctive orientations; this is useful in modeling the location of tissue responsible for a magnetic field (particularly when combined with structural magnetic resonance scans showing the gyral/sulcal pattern of each subject). On the other hand, only current flow that is at least somewhat tangential to the surface of the head will produce a detectable magnetic field. Thus it is primarily cortical activity in sulci, rather than in gyri (where the pyramidal cells are oriented perpendicular to the skull) that can be detected. This is only a minor limitation as it is estimated that two-thirds of the cortical sheet lies in sulci (Armstrong, Schleicher, Omran, Curtis, & Zilles, 1995). Finally, MEG is less sensitive to cortical sources located far away from the scalp, because the magnetic signal shows a steeper decline with increasing distance between neural source and external sensor. Overall, MEG presents perhaps the best combination of spatial and temporal resolution of noninvasive methods in common use. However, MEG studies are not very common, because the recording devices (*SQUID,* superconducting quantum inference device) are expensive and, to date, not as widely supported by routine clinical applications as magnetic resonance scanners are.

3. VISUAL WORD RECOGNITION

Any reader must first classify visual inputs as linguistic (e.g., letters, words) rather than as non-linguistic objects and then further categorize letter strings as pronounceable or not, meaningful or not, etc. The spatiotemporal dynamics of visual processing has been investigated via scalp and intracranial ERPs and MEG. While the exact timing details vary somewhat because of methodological and analytic differences, results overall suggest that visual

responses become increasingly selective for classes of visual stimuli over time, with an especially critical role for left inferior occipito-temporal areas in visual word processing.

3.1 Intracranial data: orthographic and nonorthographic stimuli

Allison and colleagues recorded evoked potentials directly from the cortical surface to a variety of visual stimui -- sinusoidal gratings, pictures of faces, word and nonword letter strings, number strings, and animate (butterflies) and inanimate (cars) objects, etc, -- in a large number of patients with intractable seizures (Allison, McCarthy, Nobre, Puce & Belger, 1994; Allison, Puce, Spencer & McCarthy, 1999). The earliest activity in the visual cortex (V1 and V2) was evidenced by N100 and P100 components which are sensitive to luminance, luminance contrast, and stimulus size but not to stimulus category. Other relatively early responses in posterior visual cortex were sensitive to sinuosoidal gratings (P120-N180-P260). Approximately 20-30 ms later, category-specific activations were observed in more ventral areas (e.g., P150-N200-N290-N700). For instance, several different cortical patches within extrastriate cortex generated surface negativities with peak latencies around 200 ms (N200), but different patches were specific to faces, objects, or letter strings (Allison et al. 1994; Nobre et al. 1994). Letter-string specific areas in the posterior fusiform gyrus responded equally to words and nonwords, whereas the anterior fusiform gyrus was sensitive to properties of letter strings (Nobre et al., 1994). On the surface of the anterior fusiform gyrus, bilaterally, a P400 component was specific to real words; a potential of the opposite polarity (N400) was observed just superior to this region, indicating locally-generated activity (McCarthy et al. 1995). Overall, depth recordings reflect the segregation of the ventral object recognition system into functionally discrete regions.

3.2 Scalp-recorded and MEG data: orthographic and nonorthographic stimuli

Schendan, Ganis, and Kutas (1998) compared ERPs to object-like and word-like stimuli as shown in Figure 1. Regardless of assigned task, a negative peak at around 95 ms (N100) over midline occipital sites was smaller for single object-like stimuli than for any variety of “string” stimuli. This distinction was quickly followed (~10 ms later) by a differentiation between strings of real letters (words and pseudowords) versus those of non-letter characters (icon strings, pseudo-font strings). The first sign of specialized processing of “linguistic” stimuli in the scalp record thus appeared around 105 ms, perhaps reflecting the experience-based tuning of the visual system to rapidly detect physical stimuli with real letter properties. Around 200 ms, word ERPs were distinguishable from those to random letter strings. This ERP difference alone, however, does not warrant the conclusion that the brain has identified one type of stimulus as a word and the other as not a word, given that the stimuli also differ in amount of prior exposure (recency, or frequency of constituent letters, bigrams, entire strings, etc.).

Insert Figure 1 About Here

Bentin, Mouchetant-Rostaing, Giard, Echallier & Pernier (1999) examined ERPs to orthographic (words, pseudowords, consonant letter strings) and nonorthographic stimuli (alphanumeric strings, strings of forms) in various oddball tasks. Comparisons were between the frequent, non-target stimuli across different tasks, designed to induce different levels of analysis (visual, phonological/phonetic, phonological/lexical, and semantic). The earliest electrophysiological distinctions arose in the visual size-judgment task in which the occipito-temporal N170 distinguished orthographic from non-orthographic stimuli: the N170 was reliably larger over the left than the right hemisphere for orthographic stimuli and marginally reversed for non-orthographic stimuli. This scalp potential resembles the intracranial N200 component elicited by all sorts of visual stimuli, albeit with different non-overlapping distributions within posterior fusiform gyrus for orthographic (words, pseudowords, nonwords) versus nonorthographic (faces) stimuli (Allison et al., 1994); the N200 for words is more left-lateralized while that for faces is bilateral or right-lateralized.

The next reliable ERP difference – an N320 larger over left hemisphere sites - distinguished pronounceable (words, pseudowords) letter strings from unpronounceable consonant strings. Simon et al. (2004) found that this component was modulated by word frequency and eliminated by massive repetition. Soon thereafter, an N350, similar to the N320 albeit with a wider scalp distribution including temporo-parietal areas, distinguished phonologically legal from phonologically illegal letter strings. Finally, an N450 similar to the N350 but extending to fronto-central areas, distinguished words from pseudowords and pseudowords from consonant strings in a task aimed at inducing semantic processing by asking participants to respond to all abstract (as opposed to concrete) words and pseudowords.

Similar results have been obtained in MEG studies comparing letter strings of various lengths, to symbol strings of equivalent lengths to letter-like symbols (rotated letters) embedded in varying amounts of Gaussian noise. The earliest responses around 100 ms in midline occipital cortex are modulated by visual noise, vary with string length, and more generally increase with visual complexity; this so-called Type I response has been linked to low level visual analyses such as extraction of nonspecific image properties – perhaps contrast borders (Tarkiainen, Helenius, Hansen, Cornelissen, & Salmelin, 1999; Tarkiainen, Cornelissen, & Salmelin, 2002). The first distinction between the processing of letter or letter-like strings versus symbol strings occurs around 150 ms over left inferior occipitotemporal regions (M170), with greater activity for letter strings (Tarkiainen et al., 1999). Like the electrical N170 and intracranially-recorded N200, the M170 does not distinguish among words, pseudowords, and consonant strings (Salmelin et al 1996). Coincident activity over right occipital-temporal regions is modulated by string length but shows no preference for letter strings. Subsequent to the M170, various word stimuli elicit an M250, sensitive to phonotactic probability (Pylkkänen, Stringfellow, & Marantz, 2002); and an M350,

sensitive to lexical frequency (Embick, Hackl, Schaeffer, Kelepir, & Marantz, 2001). The latencies of both the M250 and M350 vary with phonotactic probability but not neighborhood density (Stockall et al., 2004), although neighborhood effects interact with probability and density around the region of the M350.

In scalp recordings, orthographically legal, pronounceable pseudowords elicit ERPs that are qualitatively similar to words for several hundred milliseconds (up to ~450 ms), although ERP amplitudes typically differ. Unlike orthographically illegal, unpronounceable nonwords, but similar to real words, pronounceable pseudowords elicit an N400. With no surrounding context, the amplitude of the N400 to pseudowords may be about the same as that to real words with low usage frequency (this has not been carefully investigated). The N400 is not believed to be identical to the M350. The brain thus seems to deal with pseudowords, which although potentially meaningful, have no particular, learned meaning, no differently than real words for a considerable period after words have been differentiated from nonwords, suggesting that this early distinction may reflect differential amounts of prior exposure. Indeed, at about the same time that pseudowords are differentiated from real words, both frequency and repetition effects are observed for real words, suggesting that amount of prior exposure is one critical factor. All written words, for example, elicit a negativity (220-400 ms) over left anterior scalp (lexical processing negativity or LPN) whose peak latency varies with the eliciting word's frequency of usage (King & Kutas 1998).

4. SEMANTIC PROCESSING AND THE N400 COMPONENT

Of the large number of ERP components sensitive to language processes, the N400 is the best-used to date. The label "N400" refers to a negative-going voltage in the ERP peaking in amplitude around 400 ms after stimulus onset. This component was first noted by Kutas and Hillyard (1980a, 1980b) in a comparison of sentence-final words that formed predictable completions and those that were semantically improbable or incongruent (left column, 2). While predictable endings elicited a broad positive waveform from 200 to 600 ms, the incongruent words elicited a large negative wave in this time range. It is important to note that the N400 semantic context effect -- the difference between the two conditions – extends over some period of time. Labeling experimental effects by the latencies of their peaks is conventional in ERP research, but ERP activity always has a temporal duration (as do single neuron responses). An onset latency ~200 ms is typical of semantic context effects for visual words; context effects on spoken words typically begin somewhat earlier (as early as 50 ms in natural speech, because of coarticulatory information from the previous word, or 150 ms when the eliciting words are recorded separately and spliced into the speech stream).

Insert Figure 2 About Here

Across the 1980s, it became clear that the *N400 sentence congruity effect* was only one indication of a much broader sensitivity to semantic context, and there was nothing special about anomalous completions. First, even congruent sentence completions elicited N400s whose amplitude was directly (inversely) proportional to the goodness-of-fit between the sentence frame and the eliciting word, as indexed by offline cloze probability (Kutas & Hillyard, 1984; Kutas, Lindamood, & Hillyard, 1984). Second, context effects were observed in word pairs: the second words of semantically related word pairs elicited smaller N400s than those of unrelated pairs (right column, Figure 2; Bentin, McCarthy, & Wood, 1985; Boddy, 1981). Third, examination of the ERPs elicited by the intermediate words of sentences presented one at a time in serial order showed large N400s for the first open-class words, which became progressively smaller as the sentence (specifically semantic) context built up and constrained subsequent words (Van Petten & Kutas, 1990, 1991; Van Petten, 1989, 1993). This sentence-position effect on N400 amplitude was observed only in isolated sentences for which readers had no prior inkling of the sentence topic and not for sentences in discourse, which did not introduce completely new topics (Van Petten, 1995). Also in the 1980s, it was shown that spoken words and signs in American Sign Language elicited N400s that were qualitatively similar to their visual counterparts, and also reduced in amplitude by supportive semantic context (Kutas, Neville, & Holcomb, 1987; Holcomb & Neville, 1990, 1991; Neville, Mills, & Lawson, 1992). The correct characterization of the N400 context effect is thus not that anomalous or unrelated words elicit unusual brain responses, but rather that a large negativity between 200-500 ms or so (N400) is the default response, and that its amplitude is reduced to the degree that context aids in the interpretation of a potentially meaningful stimulus. Finally, as detailed in section *4.2* below, the amplitude in the N400 region of the ERP is sensitive not only to the context surrounding a word, but also to the lexical characteristics of the eliciting words themselves.

N400-like potentials are also evident in response to other meaningful stimuli – line drawings, photos, and environmental sounds – and also reduced in amplitude when these nonverbal stimuli are preceded by conceptually related stimuli (Ganis, Kutas, & Sereno, 1996; Holcomb & McPherson, 1994; Plante, Van Petten, & Senkfor, 2000; Van Petten & Rheinfelder, 1995). These closely resemble the verbal N400 in waveshape and timing, but have slightly different spatial distributions across the scalp. The conceptual context effect for pictorial materials has shown a more anterior maximum than the analogous effect for printed words, and the effect for meaningful nonlinguistic sounds has a small lateral asymmetry that is opposite that for spoken words (left-greater-than-right for sounds, right-greater-than-left for words. These data suggest that verbal and nonverbal N400s reflect similar cortical computations occurring in different, but overlapping, populations of neurons. Overall, the extant data suggest that N400 amplitude is a general index of the ease or difficulty of retrieving stored

conceptual knowledge associated with a word (or other meaningful stimuli), which is dependent on both the stored representation itself, and the retrieval cues provided by the preceding context.

4.1 Neural bases of the N400

There is no single, completely direct path from scalp-recorded ERPs to certain knowledge of their neural generators, but instead multiple methods that contribute to this knowledge. These include scalp recordings from patients with brain damage in known locations, intracranially-recorded ERPs in patients with electrodes implanted prior to surgery for the relief of epilepsy, and recordings of evoked magnetic fields. Application of these three methods implicate the left temporal lobe as the largest source of the scalp N400, with a substantial but lesser contribution from the right temporal lobe (see Van Petten & Luka, in press).

In split-brain patients, stimuli presented to hemispheres that also have productive speech capability elicit N400 context effects, whereas stimuli presented to a mute hemisphere do not. Because speech production is strongly dominated by the left hemisphere in most neurologically intact individuals, this suggests that the N400 is more dependent on the left than right hemisphere processes (Kutas, Hillyard, & Gazzaniga, 1988). Large amplitude reductions and delayed latencies of the N400 semantic context effect are observed in patients after strokes in the left temporal lobe or temporoparietal junction – broadly, the same regions leading to an aphasic syndrome marked by a semantic comprehension deficit (Friederici, Hahne, & von Cramon, 1998; Hagoort, Brown, & Swaab, 1996; Swaab, Brown, & Hagoort, 1997). Indeed, there is a close correspondence between magnitudes of N400 effects and standardized comprehension test scores in these aphasic patients (Kojima & Kaga, 2003; Marchand, D'Arcy, & Connolly, 2002). Lesions in the perisylvian region of the right hemisphere lead to smaller-than-normal N400 context effects, albeit with more modest reductions than following similar left hemisphere lesions (Kotz & Friederici, 2003). In contrast to the severe impact of left temporal and inferior parietal damage on the scalp-recorded N400, patients with damage restricted to the frontal lobe have normal N400 effects, and may even show N400 effects to syntactic violations that elicit different components in normal controls (Friederici, von Cramon, & Kotz, 1999; Hagoort, Wassenaar, & Brown, 2003a; Swick, Kutas, & Knight, 1998; see also Swick, 2004). Thus, although the frontal lobe is critical for many aspects of language processing, it makes little direct contribution to the N400.

Electrodes placed directly on the cortical surface, or within the depths of the cortex are used in patients being evaluated for possible surgical relief of seizures resistant to drug treatment. The potentials recorded from these electrodes have the same neurophysiological basis in synaptic activity as scalp-recorded ERPs, but can show large amplitude gradients within a distance of a few centimeters. A potential recorded in the anterior medial part of the temporal lobe (anterior to the hippocampus, in the vicinity of the collateral sulcus dividing the fusiform gyrus from the parahippocampal gyrus) has the same timecourse as

the scalp-recorded N400, and is sensitive to the same experimental manipulations (Nobre et al.,1994; Nobre & McCarthy, 1995). Other research groups have reported that what appears to be the same ERP component, in the same location, is reduced by repetition of words and line drawings (Elger et al., 1997; Fernández et al., 2001; Guillem, N'Kaoua, Rougier, & Claverie, 1996; Smith, Stapleton, & Halgren, 1986). The anterior medial part of the temporal lobe is the core brain region affected in a neurodegenerative disease known as *semantic dementia*, in which patients suffer a progressive loss of semantic knowledge with relative preservation of phonology, syntax, and recent episodic memory (Mummery, Patterson, Price, Ashburner, Frackowiak, & Hodges, 2000; Patterson & Hodges, 2000). The convergence between the intracranial recordings and the neuropsychological data are a very strong indication that this brain region is critical for access to semantic memory, and almost certainly contributes to the scalp-recorded N400 activity.

MEG studies using dipole models all suggest sources of the N400m in the left superior and/or middle temporal gyri, with a more individually variable source in the homologous right hemisphere region (Halgren et al., 2002; Helenius, Salmelin, Service, & Connolly, 1998; Helenius et al, 2002; Kwon et al., 2005; Simos, Basile, & Papanicolaou, 1997). Halgren et al. (2002) applied a distributed source modeling method showing the spatial extent of cortical activity, and found that most of the left temporal lobe (including inferior and anterior regions) was more active for incongruent than congruent sentence completions, with additional activity in the right anterior temporal lobe.

Overall, the neuropsychological, intracranial, and MEG results converge to suggest that both temporal lobes are responsible for the scalp-recorded N400 component, but that the left hemisphere makes a larger contribution than the right. Comparisons between language modalities (spoken, written, signed), between literal and nonliteral language, and between conceptual relationships expressed by words versus nonverbal stimuli await further research.

4.2 Lexical factors

4.2.1 Words, Pseudowords, Nonwords

When letter strings are presented in lists or pairs, words that are unrepeated, semantically unrelated to previous words, and low in frequency elicit very large N400s, as do orthographically legal, pronounceable nonwords (*pseudowords).*[1] By contrast unpronounceable nonwords elicit little or no N400 activity (Anderson & Holcomb, 1995; Bentin et al., 1985; Chwilla, Brown & Hagoort, 1995; Holcomb & Neville, 1990; Rugg & Nagy, 1987; Smith & Halgren, 1987; Ziegler, Besson, Jacobs & Carr, 1997). Likewise, illegal nonwords do not yield reliable incidental repetition effects (Rugg & Nagy, 1987). By contrast, immediately repeated pseudowords do elicit an ERP repetition effect smaller than that observed for real words (Doyle, Rugg, & Wells, 1996), although the contribution of N400 and a distinct ERP component related to episodic memory retrieval to the pseudoword repetition effect have not been disentangled (see Olichney et al., 2000). However, pseudowords derived from real words seem to show characteristic ERP concreteness effects (Kounios & Holcomb 1994),

suggesting that pseudowords that closely resemble real words may contact semantic memory.

Recently, Deacon et al. examined both repetition of words (TRAIN-TRAIN) and pseudowords (WOLM-WOLM), as well as semantically related pairs of words (TULIP-ROSE) and pseudowords derived from related words (PLYNT-TLEE from PLANT-TREE). The task was a delayed decision on a probe following each pair (Deacon, Dynowska, Ritter, & Grose-Fifer, 2004). ERPs to both words and pseudowords were modulated by repetition, as in previous studies. More interestingly, the "related" pseudowords elicited smaller N400s than unrelated pseudowords (derived from a pair of unrelated words). This finding is similar but different from a behavioral study appearing about the same time: Perea and Lupker (2003) reported that masked pseudowords created by letter transposition (JUGDE) led to faster lexical decision times for related words (COURT), but that letter-replaced pseudowords (more like Deacon's derived pseudowords, e.g., JUDPE) did not. The differential results be due to the different dependent measures, or differences between masked and unmasked contexts. Forster and Hector (2002) also reported slower reaction times to reject (unmasked) derived pseudowords (e.g. TURPLE) during a semantic categorization task. At first blush, all of these studies suggest that pseudoword processing activates semantic representations of (at least some) orthographically similar words. However, Deacon et al. drew a subtly different conclusion, based on an experiment in which repetition effects were also observed for pseudowords not obviously derived from real words (*nonderived pseudowords*). Deacon et al. argued that N400s to pseudowords -- and repetition effects in general -- are unlikely to be due to semantic activation *per se*, because nonderived pseudowords are not likely to activate word representations in the mental lexicon. Instead, they concluded that "N400 appears to be generated by orthographic/phonological analysis and is attenuated by the top-down feedback of semantic information to the orthographic/phonological level" (page 60).

We are more inclined to view N400 activity as arising from the semantic system itself, but it is nontrivial to distinguish this account from Deacon et al.'s (2004) feedback account on the basis of existing data. However, we believe that the starting assumption of their argument – that nonderived but legal pseudowords do not activate any semantic representations – is speculative. Recent behavioral work suggests that the word recognition system may be remarkably tolerant of mismatches between the actual input and real words. In her dissertation work, Guerrera (2004) observed that scrambled masked primes still produced repetition effects in lexical decision (e.g., SIEDAWLK or SDIWELAK both speeded RTs to target SIDEWALK). This work did not address semantic activation, but suggests caution in assuming that a letter string is too distant from a real word to contact at least some aspect of its representation in memory.

4.2.2 Vocabulary class

Kutas and Hillyard (1983) first noted that *open-class* or "content" words (nouns, verbs, most adjectives, -ly adverbs) elicited different ERPs than *closed-class* or "function" words (pronouns, articles, conjunctions, prepositions, etc) in

sentences. Subsequent experiments have attempted to determine which aspects of the vocabulary distinction -- word length, frequency of usage, repetition, contextual constraint, abstractness of meaning, referentiality, syntactic role, etc -- are responsible for these differences.

Closed class words in sentences typically elicit smaller N400s than open class words. Van Petten & Kutas (1991) suggested that this may reflect the converging influences of higher frequency of usage, higher repetition rate, and greater predictability of closed class items within sentences. And, indeed, each of these factors has been found to modulate N400. When closed class words are contextually unexpected in a sentence context, they too can elicit sizeable N400s (King & Kutas, 1995). However, the semantic content of the eliciting words may also be relevant. "Wh-words" such as "who" or "what" elicit larger N400s than do complementizers such as "that" (Kluender and Kutas, 1993a). Likewise, in the context of a sentence ("I wonder whether the candidate was annoyed THAT/WHEN…"), WHEN was found to be associated with a larger N400 than THAT (McKinnon & Osterhout 1996). These effects most likely reflect differences in referential specificity, the richness of the information retrieved from semantic memory. In line with this suggestion, Münte et al. (2001) found that N400 amplitudes to (German) closed class items in a lexical decision task were smaller than those to open class items even after frequency matching.

Another reliable difference between open-class and closed-class words is a late ramp-shaped negativity over frontal scalp sites, called the *N400-700,* which is larger for closed-class items (Van Petten & Kutas 1991; Neville et al., 1992; Osterhout, Bersick & McKinnon, 1997; King & Kutas 1998; Brown, Hagoort, & Keurs ter, 1999; Münte et al. 2001). Van Petten and Kutas first observed variability in this component when comparing closed class words in random word strings, syntactically legal but semantically anomalous sentences, and congruent sentences. The frontal N400-700 proved sensitive to both sentence type and word position. Early in a sentence, the N400-700 was essentially absent in all three conditions, but grew in amplitude over the course of congruent sentences. As the N400-700 also had not been observed when open- and closed-class words appeared in a lexical decision task (Garnsey, 1985), we suggested that it might be a member of a family of slow negative-going potentials – known as the Contingent Negative Variation (CNV) – typically seen between a warning and an imperative stimulus that an individual actively anticipates or prepares for (McCallum & Curry, 1993). In the case of sentences, closed class items presumably serve as a syntactic signal (warning) that a new head of a constituent is imminent. Brown et al. (1999) also identified this negativity to closed-class items as a CNV but speculated that it warns the reader that the next word is likely to be meaningful. Either of these functional interpretations is consistent with the finding that the N400-700 is significantly smaller in patients with Broca's aphasia (Keurs ter, Brown, Hagoort & Stegeman, 1999). Münte et al. (2001), however, observed N400-700 effects not only during sentence reading task but also with word lists (lexical decision task), although only for very high frequency closed class items (mostly determiners). They thus proposed a modified version of our hypothesis limited to determiners. As they noted, testing

this hypothesis would require comparing ERPs to various types of closed class items (determiners, conjunctions, prepositions), matched on critical variables such as length, frequency, and sentence position.

These ERP differences in the N400, N400-700, and LPN have been used to argue both for and against a qualitative and neural distinction for open versus closed class words. People generally agree that the N400 difference is quantitative rather than qualitative (although see Neville et al., 1992).
The N400-700 is a somewhat more plausible candidate for a vocabulary-class marker in that open-class words never elicit N400-700s as large as closed-class. However, this component also reflects the contextual milieu in which closed-class words appear and may reflect the typical functional role of such words in parsing, rather than pure representational differences between the word classes.

Finally, recent results indicate that the ERP is also sensitive to divisions within open class words such as that between nouns and verbs, albeit differently in word lists and sentences (Federmeier, Segal, Lombroso, & Kutas, 2000; Rösler, Streb, & Hahn, 2001; Khader, Scherag, Streb & Rösler, 2003).

4.2.3 Word frequency

Word frequency refers to an individual's life history of encounters with a particular word (estimated from normative frequency counts). High frequency words tend to elicit reliably smaller N400s than low frequency words (Allen, Badecker, and Osterhout, 2003; Barber, Vergara & Carreiras, 2004; Van Petten, 1993; Van Petten & Kutas, 1990, 1991a,b). This frequency effect is qualified by interactions with both repetition (within an experiment) and within-sentence semantic constraints (Van Petten, Kutas, Kluender, Mitchener, & McIsaac, 1991). When words are repeated in lists, or when entire sentences are repeated, the N400 frequency effect disappears upon second presentation (Smith & Halgren, 1987; Rugg, 1990; Besson, Kutas, & Van Petten, 1992). In other words, low frequency words show a disproportionate repetition effect. Within sentences, semantic factors are also capable of abolishing the N400 frequency effect. Early in a sentence, low frequency words elicit significantly larger N400s than high frequency; this frequency effect is progressively attenuated during the course of a sentence. This ordinal word position by frequency interaction is due to semantic constraints since it is not present in either random word strings or in syntactically legal but semantically anomalous sentences; in both, the N400 effect remains unabated throughout. The N400 word frequency effect is also unaffected by grammaticality. Allen et al. (2003) found that high frequency verbs (WORK) elicited smaller N400s than low frequency verbs (SWAY) whether they were grammatical or ungrammatical when they appeared within a sentence (e.g., "The man will WORK/SWAY/WORKED/SWAYED on the platform.") that participants judged for acceptability. The N400s to low frequency ungrammatical verbs were indistinguishable from those elicited by low frequency grammatical verbs, at least for regular verbs, whereas the subsequent P600 component was larger for ungrammatical than grammatical verbs but was unaffected by lexical frequency.

4.2.4 Concrete versus abstract words

Within the category of nouns, those depicting a tangible object (often pictureable) have often been associated with a larger negativity in the N400 region than less imageable nouns. This *concreteness effect* tends to be more pronounced over frontal than parietal scalp, unlike the more centro-parietal distribution of the N400 semantic context effect for written words. The concreteness effect is larger when word processing goes beyond surface level features and when contextual constraints are weak.

Paller, McIsaac & Kutas (1987) first noted greater negativity between 300-900 ms to concrete than abstract words during a concrete/abstract judgment task, although Smith and Halgren (1987) did not see a similar effect during. lexical decision. Kounios and Holcomb (1994) demonstrated the importance of task parameters within a repetition priming paradigm. They found larger (somewhat frontal) N400s to concrete than abstract words, but larger differences during an abstract/concrete judgment than in lexical decision. Assigned task also modulates the concreteness effect in sentences, which is larger in tasks requiring semantic analysis and mental imagery (although these two effects had slightly different distributions), and absent in a letter search task (West & Holcomb, 2000). As Kounios and Holcomb (1994) had also observed different ERP repetition effects for these two word types, they argued that their topographical results were most consistent with the dual-coding theory's structural account of concreteness effects: namely, that concrete words had a processing advantage relative to abstract words because they not only activated a verbal semantic memory store, like abstract words, but also a nonverbal image-based semantic memory store (Paivio, 1991).

The concreteness effect also has been examined in sentences designed to bias the final word to be either concrete or abstract, as participants rendered sense-nonsense judgments (Holcomb, Kounios, Anderson & West, 1999). When congruent, abstract and concrete words elicited indistinguishable ERPs. When anomalous, concrete words elicited somewhat more negative potentials at frontocentral sites than abstract words (in addition to a centroparietal N400 elicited by both concrete and abstract words). The results of a follow-up experiment with less predictable congruent endings (lower cloze) demonstrated that an N400 concreteness effect, although more widespread across the scalp, also could be elicited by congruent endings. These findings led to the "context-extended dual coding" hypothesis which maintains the notion of two different memory stores – a verbal one and an imagistic one – but offers similar effects of context for concrete and abstract words in the verbal system together with greater context effects for concrete words in the image system. The hypothesis also stipulates that concreteness effects can be overridden by supportive contexts under the assumption that contextual information is available prior to concreteness information. For auditory word pairs, Swaab, Baynes and Knight (2002) showed that a single related word is not sufficient context to override the concreteness effect. Although the ERPs to highly imageable words showed greater negativity than those to less imageable words, the concreteness effect was equivalent for related and unrelated pairs.

Two features of the concreteness effect warrant further discussion: (1)

although it begins coincident with the N400, the concreteness effect can last well beyond the typical N400 context effect (sometimes to1000 ms post-stimulus onset); (2) not just the amplitude but also the scalp topography of the concreteness effect varies with task variables. West and Holcomb (2000) noted that whereas the concreteness effect in their semantic task had a frontal extent it was still evident at posterior sites, much like that observed for sentences with low constraint. By contrast, the concreteness effect in their mental imagery task was progressively more frontal with time and like the concreteness effect for anomalous sentence endings was absent posteriorly. They thus suggested that the concreteness effect may sometimes comprise an N400 reflecting semantic processing plus an N700 reflecting image-based processing.

4.2.5 Orthographic Neighborhood

In reading, words with more orthographic neighbors (other words than can be formed by changing one letter) elicit larger N400s than words with fewer neighbors, although words with more neighbors elicit faster lexical decision times (Holcomb, Grainger, & O'Rourke, 2002). They attributed this latter effect to greater global semantic activation when a word from a dense neighborhood is encountered, because this includes partial activation of numerous other words that form near-matches (activation that must ultimately be suppressed in order to zero in on the meaning of the current word).

4.3 Using N400 latency to track the timing of semantic processing

One of the appeals of the ERP as a dependent measure is its exquisite temporal resolution, which can be used to clarify the relative timing and sequence of distinct processes in comprehension and production. Many components of the ERP show a fairly broad range of latencies that can be readily linked to the onset or completion of different aspects of stimulus analysis (see Kutas, McCarthy, & Donchin, 1977 for stimulus evaluation and decision making using the P3b, and Schmitt, Münte, & Kutas, 2000 for studies of speech production using the no-go N2 and Lateralized Readiness Potential components). In contrast, N400 latency for visual words is generally quite stable in the face of experimental manipulations, particularly as compared to the lawful variation in amplitude across conditions in the same subjects. There are only a handful of cases of reliable differences in the latency of N400 context effects in the visual modality. It appears that strong semantic relationships in word pairs may elicit slightly earlier effects than weaker relationships (Kutas & Van Petten, 1994, Figure 7; Luka & Van Petten, in preparation). In lexical ambiguity paradigms, probes related to a contextually inappropriate sense of the ambiguous word are associated with a delayed context effect as compared to contextually appropriate probes (Van Petten & Kutas, 1987; Van Petten, 1995).

A major reason for the stability of N400 latencies in the visual modality may be the constancy in the timing of perceptual processes across psycholinguistically distinct categories of stimuli; perceptual information about the identity of a word is likely to reach the cortical areas involved in meaning construction at much the same time regardless of contextual manipulations.

Moreover, the visual information in a single word is present all at once, and likely analyzed as a single visual pattern. The physical nature of the speech signal is quite different in having an extended temporal duration, so that information about the identity of a word accrues over time. This property of speech makes it more amenable to examinations of the temporal relationships between perceptual and semantic processes.

The earliest ERP studies of auditory word pairs and sentences revealed that the onset (and usually the peak) latency of semantic context effects were well before the acoustic offset of the eliciting words (McCallum, Farmer, & Pocock, 1984; Holcomb & Neville, 1990, 1991). This was not especially surprising given that, even when presented in isolation, most English words can be identified well before their offsets (Grosjean, 1980). More recent studies have explicitly examined the information content of the auditory signal at different points in time, relative to ERPs at those same timepoints.

Connolly and Phillips (1994) were the first to use incongruent sentence completions that shared initial phonemes with the congruent completions of those sentences, as in "The gambler had a streak of bad LUGGAGE". They found that the onset of the difference between congruent and these incongruent completions was delayed, as compared to incongruities with initial phonemes that mismatched the most expected sentence completion. Van Petten et al. pursued this finding by adding a condition, so that sentence completions were either congruent, incongruent with initial overlap, incongruent with final-overlap, or wholly incongruent (e.g., "It was a pleasant surprise to find that the car repair bill was only seventeen DOLLARS/ DOLPHINS/ SCHOLARS / HOSPITALS; Van Petten, Coulson, Rubin, Plante, & Parks, 1999). The rhyming (final-overlap) condition elicited ERPs identical those in the wholly incongruous condition, lending support to the idea that (at least in English) initial phonemes are used to establish a set of lexical candidates, and that "late entry" of candidates based on subsequent phonemes is limited. The sentence congruity effect in the rhyme and wholly-incongruous conditions began ~150 ms after word onset, although the words averaged about 600 ms in duration. As Connolly and Phillips (1994), the sentence congruity effect for words whose initial phonemes overlapped a congruent completion was delayed, beginning only ~400 ms after word onset. More critically, Van Petten et al. established the *isolation points* of the critical words prior to the sentence experiment, via the *gating* method. In this method, listeners are presented with only the first 50 ms of word, or the first 100 ms, etc (in increments of 50 ms), and forced to guess/decide what the word might be. With brief amounts of acoustic input, the number of candidates generated may be as large as the number of participants (and all the candidates might be wrong), but at some point, the large majority of participants correctly specify the actual word. The signal duration when sufficient acoustic information is present to pick out one word, and eliminate alternatives with similar onsets (i.e., to distinguish CAPTAIN from CAPTIVE and CAPSULE and CAPTION) is the isolation point, which ranged from 100 to 700 ms for the critical words in this experiment. When ERPs were time-locked to the isolation points of the critical words, the difference between congruent and incongruent words with shared initial

phonemes began at the isolation point, because this was when listeners could first determine that the sentence completion was not, in fact, congruent (see Figure 3). More interestingly however, the sentence congruity effect for words whose initial phonemes were inconsistent with a congruent completion (e.g., SCHOLARS in the sentence about a car repair bill) began some 200 ms before the isolation point. This result indicates that listeners initiate semantic processing, including integration with a sentence context, with only partial perceptual information about word identity.

Insert Figure 3 About Here

Van den Brink and colleagues also compared auditory sentence completions that were congruent, wholly incongruent, or incongruent but sharing initial phonemes with a congruent completion (van den Brink, Brown, & Hagoort, 2001; van den Brink & Hagoort, 2004). They also observed substantially delayed sentence congruity effects when the acoustic onsets of the incongruent words were consistent with a potentially congruent completion. The results from three laboratories are thus remarkably consistent in showing that the timing of ERP sentence congruity effects closely track the auditory input, although somewhat different descriptions of these results have been offered. Van Petten et al (1999) describe their results in terms of continuous semantic processing, with the onset of the sentence congruity effect depending on how long the acoustic input remains compatible with an acceptable sentence completion. Connolly and colleagues instead describe their results in terms of an early ERP component sensitive to the match or mismatch between expected and incoming words at a phonological level (Phonological Mismatch Negativity or PMMN) versus a later one reflecting semantic processing (N400). Van den Brink and colleagues likewise describe their results in terms of two distinct phenomena: an early "N200" reflecting whether or not "assessment of form-based activated lexical candidates reveals the presence of a candidate that fits the semantic and syntactic constraints of the preceding sentence" versus an N400 that "indexes difficulty in lexical integration" (page 1079). These disputes may or may not be of interest to the non-ERP researcher, given the consistency of the result that semantic context effects begin before word identification has run to completion.

5. HIGHER-LEVEL SEMANTICS: SENTENCES AND DISCOURSE

In the previous section, we described cases in which N400 amplitude or latency reflects interactions between lexical characteristics of the eliciting word and sentence context. In this section, we describe interactions between different levels of semantic context: single word, sentence-level, and discourse-level, and then examine exactly what sorts of predictions readers/listeners derive from sentential or discourse contexts. The extant results suggest that on-line

comprehension involves an interplay between the current functional organization of semantic memory and more local contextual constraints.

5.1 Single word versus sentential semantic context

Several experiments have directly compared the ERP effects of lexical and sentential contexts. Kutas (1993) contrasted sentence-final words varying in cloze probability to strongly related, moderately related, and unrelated word pairs extracted from the same sentences. ERPs to the sentence completions were more positive overall than those elicited by the second words of the pairs (presumably due to sentence wrap-up effects), but the context effects were qualitatively quite similar, although somewhat larger and earlier in sentences. We also compared intermediate words of normal congruent sentences to "syntactic prose" (meaningless but grammatically correct sentences) as a measure of sentence-level context, and related to unrelated word pairs embedded in the syntactic prose as a measure of lexical context. These two varieties of semantic context produced qualitatively similar N400 effects, with the same onset latency, although the sentential effect lasted longer. When both lexical association and sentence-level congruity could aid in constructing a message-level interpretation of a sentence, both forms of context influenced N400 amplitude in an additive fashion (Van Petten, 1993; Van Petten, Weckerly, McIsaac & Kutas, 1997).

More recently, we have pitted lexical and sentential context against one another by constructing sentences in which the final words were lexically associated to an intermediate word, but formed incongruent matches with the broader sentence context (e.g. "The zoo was working to breed the endangered BALD HEAD."). These incongruent completions elicited an N400 as large as unrelated incongruent completions in healthy young adults, and much larger than that elicited by lexically unrelated but congruent completions (Coulson & Van Petten, 2000; Coulson, Federmeier, Van Petten, & Kutas, 2005; Van Petten et al., 1999, in preparation). When lexical association and higher-level congruity are in conflict, the language processor thus favors contextual congruity.

5.1.1 Quantification and negation.

A possible exception to the general principle that "higher-level context overrules lower-level context" comes from studies examining simple subject-predicate statements with quantification or negation in sentence verification tasks. For instance, the final words of true class-inclusion statements ("A robin is a BIRD" or "All/Some apples are FRUITS") elicit N400s equivalent in size to the final words of false statements ("A robin is not a BIRD" or "No apples are FRUITS"; Fischler, Bloom, Childers, Roucos, & Perry, 1983; Kounios & Holcomb, 1992). In these cases with category-exemplar relationships, noting the semantic relationship between the elements of the proposition rather than assessment of truth value seems to be reflected in the N400 (also see section 5.3).

5.2 Sentential versus discourse-level context.

Kutas and Hillyard (1983) first observed an N400 effect of semantic violations within written prose passages, but made no systematic attempt to

separate local from global context. A strong suggestion that the N400 is sensitive to semantic constraints that span sentence boundaries arises from comparing sentence-position effects in isolated sentences to those embedded in coherent text. Although the N400 to open-class words is large for the initial words of isolated sentences and then declines as the sentences progress, there is no equivalent sentence-position effect in connected prose, because even the earliest words of a given sentence continue the topic established earlier (Van Petten, 1995).

St. George and colleagues examined the impact of extra-sentential semantic cues directly, by recording participants' ERPs to all words as they read vague paragraphs that either were or were not preceded by a disambiguating title that made then easier to understand (St. George, Mannes, & Hoffman, 1994 as in Bransford & Johnson, 1972). Although the actual words as well as the local and global context were identical in the two conditions, a comparison of the ERP average of all the words in the untitled paragraphs versus those in the titled paragraphs revealed a smaller N400 for the titled stories. This is unequivocal evidence that the N400 is sensitive to context effects beyond the individual sentence and also reflects global or discourse-level context.

Similarly, van Berkum, Hagoort, and Brown (1999b; van Berkum, Zwitserlood, Hagoort & Brown, 2003) showed that words which elicited N400s of approximately equal amplitude in an isolated (written or spoken) sentence, showed differential N400 activity when they occurred in a discourse context that made one version more plausible than the other. For instance, QUICK and SLOW elicited about the same size N400s in "Jane told her brother that he was exceptionally quick/slow this morning." However, QUICK elicited a much smaller N400 when this sentence was preceded by "By five in the morning, Jane's brother had already showered and had even gotten dressed." The latency and topography of the discourse-level N400 effect are indistinguishable from those observed for various lexical semantic violations within isolated sentences.

The timing of the ERP context effect thus offers no support to models that give temporal precedence to lexical over sentential representations or processes or to sentential level information over discourse level information. Some form of parallel or at least cascaded processing thus must be incorporated into any viable model of language comprehension, unless the priority of word level over higher order information is so short-lived as to be empirically imperceptible, untestable, or theoretically inconsequential. Nor do the N400 data lend any support to language processing accounts that invoke distinct processing mechanisms for the recruitment and/or integration of word versus sentence level, or for intra-sentential (sentence level) versus extra-sentential (discourse level) information during the construction of sentence meaning. This is not to say that there are no functional differences in how different levels of context are initially computed. For instance, Van Petten et al. (1997) demonstrated that readers with smaller working memory capacities are less able to avail themselves of sentential context than high-span readers, but are equally able to utilize single-word contexts. However, after a preceding context has been appropriately interpreted, there appears to be little difference in how single-word, sentence-level, and

discourse-level contexts are applied to the immediate processing of the current word.

5.3 Language-intrinsic semantics versus real-world knowledge

Memory researchers use the term *semantic memory* to refer to a person's store of knowledge independent of the time, place, or manner in which that knowledge was acquired (in contrast to *episodic* memories of single events that occurred in a particular spatial and temporal context). Descriptions of the organization of semantic memory typically include no distinction between word definitions specified in a language and thus known to any speaker with an adequate grasp of vocabulary (e.g., that GOOD would fall somewhere between EXCELLENT and FAIR in a rating scale containing all three words), and facts about the world that may or may not be known to a given native speaker (e.g., Australian but not American speakers of English may know that John Gorton was prime minister between John McEwen and William McMahon). Some linguistic and psycholinguistic theories do, however, draw a distinction between semantic knowledge that is intrinsic to a language and pragmatic knowledge that is independent of the language.

Most electrophysiological studies have used operational definitions of "semantic context", based on what normative samples of participants consider good sentence completions (cloze probability), or related word pairs (production norms or ratings). It seems likely that a large proportion of the items used in comparisons of congruent to incongruent sentence completions utilized "semantic" rather than "pragmatic" knowledge, but most studies have included some mix thereof. Fischler and colleagues were the first to examine whether recently-acquired knowledge modulated N400 amplitude, independent of linguistic knowledge of word definitions (Fischler, Childers, Achariyapaopan, & Perry, 1985). In the first phase of this experiment, participants learned a set of name/occupation pairs ("Matthew is a lawyer"); in the second phase, correct and re-arranged pairs were presented ("Matthew is a lawyer." versus "Matthew is a dentist."). As compared to the true statements, the false items elicited a larger N400. More recently, definitionally incongruent sentence continuations were explicitly compared to those that were incongruent only by virtue of incidental knowledge (e.g., "The Dutch trains are …" was continued with YELLOW (plausible and true), WHITE (plausible but false), or SOUR (incongruent by definition); Hagoort, Hald, Bastiaansen, & Peterson, 2004). The two varieties of inappropriate sentence continuations elicited statistically indistinguishable N400s as compared to the correct sentence continuations, as well as indistinguishable responses in functional magnetic resonance imaging data.[2] Both studies thus present a contrast to the results described above (*section 5.1.1*), in which statements became false due to an inappropriate quantifier or the presence of a negative particle. Altogether, the data suggest that the N400 primarily indexes access to semantic memory, largely independent of the semantic/pragmatic distinction. However, just how general these results are remains to be seen. Negation has not been investigated with materials and tasks that are very natural and no work has yet examined the ERP consequences of systematically

quantifying different sentence constitutents (grammatical subject or object, verb phrase).

5.4 What do contextual constraints specify?

5.4.1 Word-forms versus semantic features (with some remarks on the immediate predictive value of context)

Psycholinguistic descriptions of semantic context effects include several hypothetical mechanisms by which prior context can influence the processing of a current word. One way of categorizing such descriptions is by the timecourse of the proposed mechanism: true *priming* in which the representation of a word is preactivated during the processing of the context and before that word is actually presented (Collins & Loftus, 1975; Morton, 1969), versus various other *integrative* mechanisms that involve interactive processing of context and current word after both have been presented (Neely & Keefe, 1989; Norris, 1986; Ratcliff & McKoon, 1988). Characterizations of semantic context effects can also be divided according to their assumptions about the underlying representation of word meaning, and what is actually constrained by prior context. Some discussions of mechanisms assume that a predictive or priming mechanism necessarily involves anticipation of specific lexical items that might occur next, but other models stipulate that word meanings are comprised of bundles of semantic features shared with other words. In the latter sort of model with distributed representation of word meanings, context effects can be readily simulated by connectionist networks in which encountering one word produces partial activation of other words with shared semantic features (McRae, de Sa, & Seidenberg, 1997; Sharkey, 1989). Modeling semantic context effects that are not based on similarity (thematic relations) – which will include nearly all sentence and discourse effects – is likely to be a more difficult enterprise, although likely feasible (Elman, 2004; Ferretti, McRae, & Hatherell, 2001). A full discussion of these topics is outside the scope of the current review, but we note that questions about when semantic context acts are logically orthogonal to questions about exactly what becomes easier to process with supportive prior context. In empirical work, however, the same results often provide information about both “when” and “what”, so that the two issues are interlaced in the following review.

Both older and newer ERP studies of sentence processing argue very strongly that context can facilitate the processing of words with some appropriate semantic features, even when the specific lexical items could not have been expected or predicted. This was first demonstrated by comparisons between congruent completions of high cloze probability, anomalous completions, and anomalous completions that were semantically related to the congruent words (e.g., “The pizza was too hot to EAT / CRY / DRINK“; Kutas et al., 1984; Kutas & Hillyard, 1984). The related anomalies elicited a larger N400 than the congruent endings, but substantially smaller than the unrelated anomalies. Critically, the difference between the related and unrelated anomalies showed no latency delay, suggesting that words like DRINK were facilitated directly by the sentence

context itself, rather than by some sort of secondary priming between the (un-presented, but predictable) congruent ending and the related anomaly. This result thus argues for a featural semantic representation, and suggests that a sentence context facilitates the processing of words containing at least some features that can be matched to the specifications of the preceding sentence fragment. Moreover, the results suggest that context specifies something about the meaning of upcoming words, but not necessarily a list of candidate word forms *per se*.

More recently, Federmeier and Kutas (1999a,b) refined the "related anomaly" design by constructing contexts that more narrowly constrained the semantic features that would form a good fit. For instance, although both EARRING and NECKLACE are types of jewelry and thus share many semantic features, they also differ in multiple properties such that EARRING is a better completion for the context "I guess his girlfriend really encouraged him to get it pierced. But his father sure blew up when he came home wearing that …". In contrast, NECKLACE is a better completion for the context "She keeps twirling it around and around under her collar. Stephanie seems really happy that Dan gave her that …". The ERPs thus showed much larger N400s for the wrong variety of jewelry (or the wrong team sport, wrong hand tool, etc) than for the congruent word. More critically however, the incongruent words that had high featural overlap with the congruent word elicited smaller N400s than words with low featural overlap (such as MASCARA or LIPSTICK for the examples above). Visual half field studies revealed that this effect was present only for initial presentation to the left hemisphere (Federmeier & Kutas, 1999b).

Although the "related anomaly" experiments suggest that sentence and discourse contexts act to specify the meanings of plausible continuations, and not their physical forms, other results are most compatible with the idea that context also can be used to predict particular words. Some recent experiments have capitalized on situations in which semantic plausibility is linked with a nonsemantic lexical feature. For instance, in both Spanish and Dutch, nouns have grammatical gender. The gender of a noun is largely unpredictable from its meaning, but in grammatically correct sentences, the genders of articles and adjectives must match their nouns. Thus, if readers actively anticipate that "Little Red Riding Hood carried the food for her grandmother in…" A BASKET, specifically, rather than some sort of container generically, the Spanish reader will also predict the feminine article UNA -- to agree with the feminine noun CANASTA (i.e., "Caperucita Roja cargaba la comida para su abuela en una canasta").

Wicha et al. examined ERPs elicited by articles whose gender agreed or disagreed with the most plausible sentence continuation in written Spanish sentences, and in spoken and written sentences that continued with line drawings of objects (Wicha, Bates, Moreno, & Kutas, 2003; Wicha, Moreno, & Kutas, 2003, 2004). Van Berkum and colleagues used a very similar design with spoken Dutch materials, except that the critical words were gender-marked adjectives (e.g., GROOT versus GROTE) that preceded their nouns by several words (van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). In all

cases, articles and adjectives whose gender was inappropriate for the most plausible sentence continuation elicited different ERPs than words whose gender matched. These results are striking because the "inappropriate" articles/adjectives did not introduce any sort of error into the sentences; the observation of ERP differences at the early processing point can only indicate that participants were anticipating specific nouns of specific genders. The nature of the ERP effect differed across experiments, however. In the Spanish experiment using nouns depicted by line drawings, the "wrong gender" articles elicited larger negativities (N400-like activity) than "correct gender" articles (Wicha et al., 2003). In the language-only experiments (both Spanish and Dutch), the gender-mismatching items elicited positive potentials similar to those observed to overt agreement errors ("P600-like" effects, see *section 6*). The differential effects across experiments suggest that participants may have perceived the gender mismatch as an agreement error in the language-only experiments, but a violation of a more meaning-based violation in the mixed-media case.

Although English does not include grammatical gender, DeLong, Urbach and Kutas (2005) exploited a conceptually similar agreement phenomenon to examine the specificity of sentence-based predictions, namely the A/AN alternation based on noun phonology. Sentence fragments were constructed such that the cloze-probability of possible completions (all nouns) ranged from 10% to 90%, and half of those nouns began with a vowel-sound (calling for AN) and half began with a consonant-sound (calling for A). Sentences were visually presented one word at a time, so that ERPs elicited by the articles could be examined contingent on whether they matched a very predictable sentence completion, or a less predictable (but congruent) completion. The articles always agreed with the subsequent nouns, so that no phonological mismatches were encountered during the experiment. The articles elicited N400-activity, whose amplitude was strongly (inversely) correlated with the cloze probability of the article. In other words, if KITE was a favored sentence continuation, then the word AN (as in AN AIRPLANE) elicited a larger N400 than the word A immediately preceding the expected noun. Because the effects were graded according to cloze probability, rather than showing a dichotomous split between more- versus less-favored sentence continuations, DeLong et al. concluded that although readers do predict word-forms, they entertain a range of possibilities that are graded in strength. Like the experiments with grammatical gender, the results strongly suggest that predictions from context may include specific word forms, and that these predictions are made in real time during reading. Much work still needs to reveal exactly what informs these predictions, what neural substrates support predictive processing, and how what people know and how quickly they can access that knowledge influence the routine use and/or efficacy of prediction.

We began this section with a dichotomy: that sentences and discourse may be used to derive expectations for specific lexical items, or for the semantic content (features) of upcoming words. The data, however, seem not to respect this dichotomy. The "related-anomaly" experiments (Federmeier & Kutas, 1999a,b; Kutas & Hillyard, 1984) indicate that words which would never be

predicted on the basis of the preceding context nonetheless undergo more fluent processing if they share semantic features with the most plausible completion. The experiments varying the gender and phonological form of articles clearly indicate that readers and listeners form expectations for specific words, because neither gender markings nor the a/an alternation carry much, if any, semantic information, but are specific to lexical items. We might be tempted to conclude that general expectations about meaning are the norm, and that these become focused onto specific word-forms only when the contextual constraint is extremely strong, such that only a single word exemplifies exactly the meaning specified by the context. Indeed, in the Federmeier and Kutas (1999a) study, the reduction in N400 amplitude to implausible items with high featural overlap was more pronounced in highly constraining sentences. However the graded effects observed in the "A versus AN" experiment would at first glance appear to argue against this (DeLong et al., 2005). Overall, the uneasy match between the initial theoretical dichotomy and the empirical results suggest that a division between word-forms and their meanings may not be the correct description of how the brain processes words, and that some revision of our thinking may be in order.

5.4.2 Immediacy, incrementality, and prediction.

All of the context effects reviewed in this section emerge rather quickly after presentation of information that fit or does not fit with the prior context. In natural speech, van Berkum et al. (2005) showed that the differential response to adjectives whose gender was inconsistent with expectations about a subsequent noun began ~50 ms after presentation of the relevant acoustic input (the inflection at the end of the adjective). This early onset is about the same as when N400 differences between congruent and incongruent sentence completions begin in unedited natural speech (Holcomb & Neville, 1991). With visual presentation, differential activity to articles that do not match anticipated nouns begins somewhat later, around 200-300 ms (DeLong et al., 2005; Wicha et al., 2003, although sooner when time-locked to the distinguishing information, van Berkum et al., 2005). Congruent and incongruent open-class words show differential ERPs beginning around 200 ms after stimulus onset in the visual modality, as do incongruent words differing in their relationship with the congruent words. Combined with the results discussed in section 4.3 showing contextual influences on spoken words before the acoustic information is sufficient to uniquely identify the words (Van Petten et al., 1999), the data point to a major role for top-down contextual influences in reading and listening.

While ERP data are no longer alone in suggesting this, we note that they have been offered as evidence for the immediacy and incrementality of (at least) semantic processing during online language comprehension since the early 1980's. To date the ERP experiments have not addressed the full extent of the argument about the nature or completeness of the representations computed immediately on a word by word incremental basis (full or partial, and underspecified). Until quite recently ERP data, however, have stood as part of a small minority arguing not only for immediate semantic analysis and integration into an evolving sentence representation but for contextual neural pre-activation

of upcoming semantic, syntactic, and lexical information.

5.5 Nonliteral language

People use language in different ways for different purposes because it serves various communicative and social functions that go well beyond conveying facts. People don't always mean what they say, or say what they mean directly – and yet typically a reader/listener from the same culture as the speaker has no difficulty understanding that what s/he read or heard was a promise, a threat, a command, an indirect request or that a statement is dripping with irony, funny, or intended to be metaphorical. The psycholinguistic and linguistic literatures are rife with discussions about the extent to which there is a basic distinction between literal and nonliteral language representations and processes. Views span the range from those that argue that the dichotomy between literal and figurative thought or language is a psychological illusion and that a single set of processes is responsible for the processing of both, to the strong claim that figurative language is unusual and special, and as such engages different comprehension processes (Katz, Cacciari, Gibbs & Turner, 1998).

To date there are only a few electrophysiological investigations of nonliteral language processing, specifically of jokes and metaphors. One recurrent theme in these studies is whether the right hemisphere makes a special contribution to the comprehension of nonliteral language. This question has been of interest since early reports that one subtle communicative deficit in patients with damage to the right hemisphere is difficulty understanding nonliteral language (Brownell, Simpson, Bihrle, & Potter, 1990; but see Gagnon, Goulet, Giroux, & Joanette, 2003 for a recent claim that right- and left-hemisphere patients are more similar than different). None of the studies described below included neurological patients, relying instead on less direct means of assessing hemispheric asymmetry: examining the lateral distribution of scalp ERP effects, comparing right- and left-handed participants (on the hypothesis that left-handers have a somewhat more bilateral neural substrate for language, and visual half field presentations.

5.5.1 Jokes

Coulson and Kutas (2001) compared the processing of one-line jokes versus non-joke sentences, with final words matched on cloze probability. Their primary aim was to test a two-stage model of joke comprehension wherein an initial stage of "surprise" registration is followed by a stage of coherence re-establishment. They also were able to assess the psychological reality of *frame-shifting* – a process of activating a new frame from long-term memory in order to reinterpret information already in working memory (Coulson, 2001). Although not specific to jokes, frame shifting is necessary to re-establish coherence when encountering the punch word or line. As in many recent language studies, the specific pattern of results differed depending on contextual constraint (final word cloze above and below 40%) and whether or not individuals "got" the joke. Better joke comprehenders responded to jokes with larger late positivities (500-900 ms),

a sustained negativity over left frontal sites, and -- for those in constraining contexts -- a slightly larger N400 as well. By contrast, in the poorer joke comprehenders, the punch-words elicited an enhanced frontal negativity (300-700 ms). Coulson and Lovett (2004) likewise observed larger late positivities to jokes relative to cloze-equated straight endings, with a laterality influenced by participant handedness and gender. A frontal negativity was seen only in right handers, and a slightly enhanced N400 only in left handers with low verbal skills. The results were not simply explicable in terms of any two-stage theory. However, as the enhanced late positivity to jokes is not dissimilar to those reported for syntactic violations in nonhumorous sentences, it is worth considering the possible commonalities between the two in terms of sentence re-analysis, retrieval and integration of information in working memory, etc.

Coulson and Williams (2005) examined ERPs to similar materials when punch-words or straight endings were presented to one or the other hemifield to ensure that visual information reached one hemisphere slightly before the other. Jokes elicited larger N400s than straight endings only when the sentence-final words went into the right visual field (left hemisphere). With LVF presentation, both jokes and low-cloze straight endings elicited larger N400s than high-cloze non-joke endings, but did not differ from each another. A sustained frontal negativity and a late fronto-central positivity to jokes did not differ with visual field of presentation. Overall, the right hemisphere seems no more stymied by processing a joke as by any other unexpected noun, suggesting that it may be better able to use sentential context to facilitate processing and integration of a punch word. This conclusion is supported by the studies in Coulson and Wu (2005) showing that greater N400 reduction to single words in central vision relevant than irrelevant to an immediately preceding one-line joke as well as a greater reduction when such probe words were presented in the LVF than RVF (right hemisphere).

We can now re-consider whether joke processing differs from that of non-joke sentences. Certainly the data patterns indicate substantial overlap in processing, with the reading of both accompanied by modulations in N400 amplitude. At the same time, there appears to be a difference in the contributions of the two hemispheres to joke and non-joke processing; some aspect (unknown) of joke comprehension appears to be easier for the right hemisphere, as reflected in reduced N400s associated with lateralized presentation of either punch words or joke-relevant probe words following one-liners. Whether the ephemeral sustained negativity over left frontal sites also will prove to distinguish jokes from non-jokes remains to be seen. A similar uncertainty colors the specificity of the late positivities (frontal and/or parietal) that occasionally characterize the ERPs to jokes. What is most clear from these studies is the need to track more than just whether a sentence is a joke or not, including whether participants get it, and stable characteristics of participants such as verbal ability, handedness, familial handedness, and gender. Indeed, this is undoubtedly a valuable lesson for all language studies.

5.5.2 Metaphors

Most current processing models of metaphor comprehension assume that the same operations are involved in literal and metaphorical language comprehension, but that metaphorical language especially taxes certain operations (see Katz et al., 1998). Several sources of behavioral evidence indicate that metaphorical meanings are sometimes available with the same time course as literal meanings and may even compete with each other. Researchers have examined these issues with ERPs as equivalent reaction times don't necessarily translate into equivalent processing demands. Although the specific alternative to the standard view differs across the ERP papers published to date, no electrophysiological study has yet offered any strong evidence for a qualitative difference in the way literal and metaphorical language is processed. The final words of metaphors typically elicit slightly larger N400 amplitudes than equally unexpected (low cloze) words completing literal statements. This suggests that people invoke the same operations, but also do experience more difficulty integrating words with a metaphoric than literal context.

Pynte and colleagues initially established that final words of short metaphoric sentences elicited larger N400s than categorical statements, despite being matched on cloze probability Pynte, Besson, Robichon, & Poli, 1996). Subsequent experiments showed that the ease of processing metaphoric statement, like literal statements, could be modulated by prior context. When presented in isolation, relatively familiar and unfamiliar metaphors elicited equivalent ERPs (e.g., "Those fighters are LIONS." versus "Those apprentices are LIONS."). However, both sets of metaphors benefited from preceding context so that an unfamiliar metaphor with a useful context ("They are not cowardly. Those apprentices are LIONS.") elicited a smaller N400 than a familiar metaphor preceded by a irrelevant context ("They are not naïve. Those fighters are LIONS."), and similarly the familiar metaphors with a useful context were easier to process than unfamiliar metaphors with an irrelevant context. The metaphors-in-context were not compared to a literal condition to determine if the enhanced N400 observed for isolated metaphors disappeared with appropriate context. However, across the multiple experiments, there was no hint of distinct processing stages during metaphor comprehension.

While granting that none of the predictions of the standard view have stood the test of data, Tartter and colleagues raise the possibility that while processing a metaphorical expression comprehenders nonetheless do take note of the anomalous nature of the expression's literal meaning (Tartter, Gomes, Dubrovsky, Molholm, & Stewart, 2002). They suggest this realization may underlie the phenomenological sense of satisfaction experienced when confronting a metaphorical statement. They compared the ERPs to final words completing the same sentence frame either literally, metaphorically, or anomalously (e.g., "The flowers were watered by nature's RAIN / TEARS / LAUGHTER", respectively). Cloze probabilities were higher for the literal endings than the other two conditions (both near-zero). They argue that if context is used to construct a meaningful interpretation of a metaphorical expression without any accompanying appreciation that the expression's literal meaning is anomalous,

then a metaphorical but literally incongruous ending should not elicit an N400. This construal of the N400 as an anomaly detector is problematic given that words that fit but are less expected also elicit sizable N400s; semantic anomalies are neither necessary nor sufficient to elicit N400s. Tartter et al. obtained a three-way amplitude difference in the peak latency range of the N400: anomalous > metaphorical > literal, however, the ERPs to literal completions pulled away from the other two conditions earlier than the differentiation between metaphoric and anomalous completions. This pattern of results suggests (to us) that that semantically anomalous sentence endings were more difficult to process (as reflected in larger and longer N400 congruity effect) than the metaphorical endings which were in turn more difficult to fit with the prior context (as reflected in greater N400 activity) than the literal, congruent endings. The data pattern is also consistent with the view that metaphors are initially processed much the same as semantic anomalies although they are meaningfully resolved in a shorter duration. However, this latter conclusion is somewhat complicated by the difference in cloze probability and frequency between the literal and metaphoric completions.

A significant analytic and empirical step in this area was taken by Coulson and Van Petten (2002) who hypothesized that the same conceptual operations important for understanding metaphors are often also engaged during the comprehension of literal statements These include establishing mappings and recruiting background information, or, more specifically, looking for correspondences in attributes and relations between the target and source domains, setting up the mappings, aligning them, selecting some and suppressing others. By using sentences describing situations where one object was substituted, mistaken for, or used to represent another (the *literal mapping* condition, e.g., “He used cough syrup as an INTOXICANT.”), they created sentences requiring mappings between two objects and the domains in which they commonly occur, albeit with less effort than for a metaphor (e.g., “He knows that power is a strong INTOXICANT.”), but more than for a simple literal statement with fewer or no mappings (e.g., “He knows that whiskey is a strong INTOXICANT.”). ERPs elicited by sentence-final words showed graded N400 activity, with metaphor > literal mapping > literal, although the three conditions were matched in cloze probability. These data indicate that although literal and figurative language may engage qualitatively similar processes (in contrast to the now unpopular “standard view”), increasing the burdens on mapping and conceptual integration can make metaphors more difficult to process. .

Finally, Kazmerski and colleagues examined individual differences in metaphor comprehension, and found that both vocabulary and working memory capacity were important factors as individuals determined whether a metaphoric statement was literally untrue (as compared to false statements without metaphoric interpretations, e.g., “The beaver is a LUMBERJACK.” versus “The rumor was a LUMBERJACK.”). High IQ participants showed greater interference presumably because the figurative meaning was extracted without voluntary effort (Kazmerski, Blasko, & Dessalegn-Banchiamlack, 2003). Lower IQ participants had equivalent N400s for the metaphoric and anomalous statements,

suggesting that they had no additional trouble rejecting metaphorical sentences as untrue. Thus, although individuals with lower IQs clearly understood the metaphors in an offline task, the online evidence provided by the ERP seems to indicate that metaphorical processing is not always obligatory or automatic.

6. MORPHOSYNTACTIC PROCESSING AND RELATED COMPONENTS

This section surveys a number of issues concerning morphological and syntactic processing that have been addressed using ERPs: (1) the encapsulation and/or interaction of semantic and syntactic processes, (2) the influence of other, non-linguistic cognitive variables (such as working memory) on syntactic processing, and (3) the fractionation of syntactic processing into discrete stages. While the jury is still out on most of these issues, a body of evidence has begun to accumulate that allows us to reflect on just how much is known at this point. Invariably, predictions of Fodor's (1983) modularity hypothesis with regard to linguistic representations and processes provide much of the framework for this inquiry.

Before evaluating the evidence, however, it may be useful to invoke a caveat while it is relatively easy, via experimental manipulation of linguistic materials, to obtain differences in the polarity, latency, amplitude, and scalp distribution of brain responses, it is often difficult to ascertain exactly what such differences might reflect functionally.

6.1 Background

As Sections 4 and 5 make clear, the N400 has become well established as a brain index of semantic and pragmatic processing. More recently discovered components related to syntactic and morphological processing have both complicated this picture and raised questions about the extent to which the N400 should be considered an all-purpose index of semantic processing. As early as 1983, Kutas and Hillyard demonstrated that while violations of semantic well-formedness reliably elicited an N400 (but see section 6.2.2), violations of morphosyntactic well-formedness elicited different ERP components. In addition to semantic violations, the study included number agreement discrepancies (e.g. '*she dig*'; '*a balloons*'), as well as both finite and non-finite verb forms in inappropriate sentence contexts ('*to stayed*', '*are consider*'). In contrast to the centro-parietal N400 between 300-500 ms elicited by semantic anomalies, the responses to all three morphosyntactic violations showed *fronto-central* negativities between 300-400 ms and marginally significant parietal *positivities* at 300 ms post onset of words immediately following the violations (Kutas & Hillyard 1983, Figure 4). Although the import of these differences was not entirely clear at the time, Kutas and Hillyard observed that the elicitation of N400s by semantic but not morphosyntactic anomalies pointed to potentially separate underlying neural processing systems.

This state of affairs has largely persisted to the present day: morphosyntactic anomalies of various sorts have typically been associated with

either anterior negative or late positive responses, or with both, but typically not with an N400, or not just an N400. Such morphosyntactically triggered responses exhibit a certain degree of variability with regard to both latency and scalp distribution.

Anterior negative responses are usually either left lateralized or bilaterally distributed (although see Ueno & Kluender, 2003b for right lateralized anterior negativities in Japanese). There are, however, enough reports that this enhanced negativity is most pronounced at left anterior sites that it is commonly referred to as Left Anterior Negativity (LAN). While this response often does not exhibit a clear peak, its latency usually falls between 300 and 500 ms post stimulus onset, and it has also been reported as early as 100 ms. Although this early effect frequently persists into the 300-500 ms latency window and has the same scalp distribution as the later latency LAN, some researchers have proposed a functional distinction between negativities that occur between 100-300 ms post word onset – a so-called early LAN or "ELAN" – and those that occur between 300-500 ms (reserved for the LAN).

The experimental paradigms that elicit an ELAN violate the parser's expectation that the incoming word will be of a particular grammatical category (e.g. a verb rather than a noun following a preposition plus article). Since the early LAN is impervious to the proportion of ill-formed experimental sentences (Hahne & Friederici, 1999), doesn't appear until 13 years of age (Hahne, Eckstein, & Friederici, 2004), is suppressed under degraded visual presentation (Gunter, Friederici, & Hahne, 1999), and is compromised by damage to left anterior regions, as in Broca's aphasia (Friederici, von Cramon, & Kotz, 1999; Kotz, Frisch, von Cramon, & Friederici, 2003), it has been taken by some to index an early automatic process of local phrase structure building, during which word category information is used to assign initial syntactic structure. However, this conclusion remains controversial, as the ELAN has to date been reliably elicited under only a narrow set of conditions involving word category violations of this type.

The functional significance of the LAN (300-500 ms) has been similarly difficult to pin down, as it not only appears (like the ELAN) to word category violations, but also often to agreement violations in word (Hagoort & Brown, 2000) and pseudoword sentences (Münte, Matzke, & Johannes, 1997), and almost always accompanies fully grammatical long-distance dependency (filler-gap) constructions containing no violations (Kluender & Kutas, 1993a; King & Kutas, 1995). The main conundrum here is whether the LAN is a specific response to morphosyntactic illformedness, and/or whether it can (also) be explained in terms of general working memory processes. On the first view, inspired by serial parsing models, the LAN is hypothesized to reflect difficulties in the use of grammatical (as opposed to semantic) information like inflectional morphology (person, number, gender, and case features) used in thematic role assignment. Thus, while the ELAN is hypothesized to index an initial stage of phrase structure building, the LAN itself is hypothesized to index a subsequent processing stage devoted to thematic role assignment. On the second view, inspired by models of verbal working memory, the LAN is hypothesized to index

both a “look forward” function triggered by displaced sentence constituents (e.g. fillers seeking subsequent gaps; Kluender & Kutas, 1993a,b; Kluender & Münte, 1998), as well as a kind of a “look back” function triggered when current, unexpected syntactic information must be reconciled and aligned with preceding information occurring earlier in the sentence, including gaps seeking appropriate fillers (Kluender & Kutas, 1993a,b; King and Kutas, 1995; Ueno & Kluender, 2003a), verbs seeking appropriate subjects (Osterhout & Holcomb, 1992; King & Kutas, 1995; Vos, Gunter, Kolk, & Mulder, 2001), anaphora seeking appropriate antecedents (Coulson, King, & Kutas, 1998; van Berkum, Brown, and Hagoort, 1999a; van Berkum, Brown, Hagoort, & Zwitserlood, 2003), negative polarity items seeking appropriate licensers (Shao & Neville, 1996).

Note that these two views of the LAN are not mutually exclusive. Attempts to demarcate these two views have suggested that the LAN elicited by (morpho)syntactic ill-formedness is more reliably left-lateralized than the LAN elicited by long-distance dependencies, but even LANs in response to morphosyntactic violations exhibit bilateral distribution on occasion, with both auditory and visual presentation. What remains consistent across the entire family of LAN and ELAN components, however, is its anterior scalp distribution (see Figure 4).

Late (i.e., later than the ELAN if not LAN and N400) positive ERP components to morphological and syntactic anomalies are typically largest over centro-parietal sites, but can exhibit anterior maxima. This potential (measured between 500-800 ms) is now routinely referred to as the P600, as it often displays maximum amplitude at this latency (see Figure 4), although it can onset as early as 200 ms (following another positive component - the P200) and often appears as a long-lasting positive shift with no clear peak. While the P600 has been observed with a wide variety of violation types, including subject-verb agreement, verb inflection, case inflection, phrase structure, and higher-level syntactic constraints, it is not specific to violations *per se*. Enhanced late positivities have also been observed in syntactically well-formed sentences with a non-preferred structure (e.g., garden path sentences) or with relatively complex syntactic structures, such as those with embedded long distance dependencies.

Insert Figure 4 About Here

As with the LAN component, there are basically two schools of thought with regard to the P600. One is that the P600 is a general purpose response to low probability target events often associated with some form of categorization and/or binary decision (P3b component). Alternatively, among the proposals for limiting the functional significance of the P600 to language contexts are suggestions that it indexes (1) the inability of the parser to assign the preferred structure to the input (Hagoort, Brown, & Groothusen, 1993), (2) a late, controlled (as opposed to automatic) process of syntactic re-analysis or repair once a syntactic error has been detected in a multi-stage parsing model (Friederici,

Hahne, & Mecklinger, 1996), (3) syntactic integration difficulty (Kaan et al., 2000), and (4) any kind of linguistic parsing difficulty (semantic, morphosyntactic, or orthographic; Münte, Heinze, Matzke, Wieringa, & Johannes, 1998). The P600, unlike the early LAN, is sensitive to the proportion of experimental sentences that are syntactically ill-formed (Gunter, Stowe, & Mulder, 1997; Coulson, King, & Kutas, 1998; Hahne & Friederici, 1999), not present for morphosyntactic violations on pseudowords (Münte et al., 1997), and visible in Broca's aphasics (Friederici et al., 1999), but smaller or absent in individuals with damage to the basal ganglia (Frisch, Kotz, von Cramon, & Friederici, 2003; Kotz et al. 2003).

6.2 The encapsulation versus interaction of syntax and semantics

6.2.1 Interactions among ERP components

A number of the earliest studies of syntactic processing contained both standard semantic anomalies, which elicited an N400, as well as various types of morphosyntactic anomalies, which elicited ELANs, LANs, and P600s, although not in the same combinations: Hagoort et al. (1993) reported only a P600, Osterhout & Holcomb (1992) reported both a P600 and a LAN, while Neville et al. (1991) and Friederici, Pfeifer, & Hahne (1993) reported an ELAN in addition to the LAN and P600.

A fair amount of research has since been devoted to determining the extent to which the N400 does or does not interact on line with the purported markers of syntax-related processing (ELAN, LAN, P600). The results of studies crossing semantic and syntactic violations to this end have been rather mixed. Consequently, the answer to this question has been somewhat perplexing, suggesting that the inquiry may need reframing.

Gunter, Stowe, & Mulder (1997) reported that purely morphosyntactic violations elicited P600s of equivalent amplitude to morphosyntactically erroneous words containing an additional semantic anomaly, just as there was no difference in amplitude between the N400s elicited by semantic violations with or without additional morphosyntactic deformations. However, when complexity (in the form of an adverbial adjunct clause separating the main clause verb from its arguments) was added, morphosyntactic violations with and without semantic incongruity elicited ELAN and LAN components in addition to the P600. Moreover, while the P600 varied in amplitude as a function of both semantic congruity and complexity (i.e., it was smaller either when the verb was morphosyntactically and semantically incorrect, or when the sentence containing it had a complex structure), the ELAN and LAN did not.

Osterhout & Nicol (1999) used stimulus materials of the type "One kangaroo at the San Diego Zoo would sometimes SIT / SITTING / WRITE / WRITING all day" in a similar manipulation. They reported that double violations (e.g. WRITING) elicited both an N400 and a P600. Both components showed marginally significant interactions between morphosyntax and semantics, and were not as large as would be predicted by linear summation of the N400 to semantic anomaly and the P600 to morphosyntactic ill-formedness—both facts suggesting some measure of interaction between the two types of processes.

Purely morphosyntactic violations (SITTING) elicited a LAN in one experiment and an N400 in the other.

Using similar manipulations of subject-verb agreement in Finnish, Palohati, Leino, Jokela, Kopra, & Paavilainen (2005) reported that purely morphosyntactic violations elicited an LAN + P600 complex, semantic violations an N400, and the combined violation an LAN/N400 + P600 complex, which, however, was not a mere linear summation of the independent responses to morphosyntactic and semantic anomaly. The combined response showed greater negativity between 300 and 500 ms over the left hemisphere, again indicating an interaction between morphosyntactic and semantic processes. Given the slow presentation rate (one word every 800 ms), however, the results cannot be taken as conclusive.

Instead of manipulating verbal forms, Hagoort (2003) placed morphosyntactic violations of gender and number marking (on definite articles) and semantic violations of plausibility (on following adjectives) in definite noun phrases in both sentence-initial and sentence-final positions. Compared to the control condition (roughly, '$\text{these}_{\text{COMMON.}}$ noisy $\text{YOUNGSTERS}_{\text{COMMON}}$'), in sentence-initial position the semantic violation ('$\text{these}_{\text{COMMON}}$ cloudy $\text{YOUNGSTERS}_{\text{COMMON}}$') elicited a larger N400, the morphosyntactic violation ('$\text{this}_{\text{NEUTER}}$ noisy $\text{YOUNGSTERS}_{\text{COMMON}}$') a larger LAN + P600 complex, and the combined violation ('$\text{this}_{\text{NEUTER}}$ cloudy $\text{YOUNGSTERS}_{\text{COMMON}}$') an even larger N400 than the semantic violation alone, followed by a P600 equivalent to that for morphosyntactic violations. Thus again, the double violation was not a linear sum of the independent responses to morphosyntactic and semantic violations alone, indicating some level of interaction between morphosyntactic and semantic processes. Since Hagoort reported a LAN to morphosyntactic violations, but a more anterior than usual (p. 892) equipotential negativity across the scalp (p. 887) to combined violations in sentence-initial positions, this may have been the same LAN/N400+P600 reported by Palohati et al. (2005). Difference ERPs would help to resolve this issue.

Ainsworth-Darnell, Shulman, & Boland (1998) contrasted subcategorization violations with semantic violations using quartets of sentences like:

> Jill entrusted the recipe to FRIENDS/PLATFORMS before she suddenly disappeared.
>
> Jill entrusted the recipe FRIENDS/PLATFORMS before she suddenly disappeared.

The semantic violation ("to PLATFORMS") elicited an N400, while the syntactic violation("FRIENDS") elicited both a small N400 and the expected P600. However, the combined violation ("PLATFORMS") elicited an N400 that did not differ from that to the semantic violation, and a P600 that did not differ from that to the syntactic violation, suggestingno interaction of syntax with semantics.

In contrast to the above studies, which all relied on visual (sometimes rather slow) presentation, Friederici and colleagues have used spoken materials such as:

Das Brot wurde GEGESSEN
the bread was eaten

Der Vulkan wurde GEGESSEN
the volcano was eaten

Das Eis wurde im GEGESSEN
the ice [cream] was in-the eaten
'The ice cream was eaten in the'

Das Türschloß wurde im GEGESSEN
the door-lock was in-the eaten
'The door lock was eaten in the'

The manipulation in the last two examples is a word category violation because,instead of the expected head noun, the verb ("GEGESSEN", 'eaten') immediately follows the coalesced preposition plus definite article ("im", 'in-the').

Since on a modular view, syntactic processing should precede semantic processing, Hahne & Friederici (2002) predicted that combined violations of syntax and semantics should suppress the N400 elicited by semantic violations alone. Using a 100 ms post-stimulus-onset baseline to compensate for the fact that the critical word was preceded by different lexical items across conditions, the semantic violation elicited an N400, the word category violation an ELAN + P600 complex, and the combined violation a broad anterior negativity plus a broad posterior positivity, both starting early and continuing throughout the epoch, but no N400, as predicted. When participants were instructed to determine whether the sentence made sense or not and to ignore the structural violations, the semantic violation again elicited an N400, the word category violation an ELAN and a very small, unreliable late positivity, and the combined condition a temporally-extended anterior negativity plus a phasic posterior N400—i.e. no P600 whatsoever. Hahne & Friederici concluded that task-dependent variables may suppress or enhance late controlled processes indexed by the N400 and the P600 but do not affect early automatic processes indexed by the ELAN.

An alternative interpretation is that the structure of the word category violation conditions, with the head noun of the aborted prepositional phrase missing, may have discouraged readers from attempting to associate the clause-final verb with the subject noun phrase altogether, effectively blocking N400 modulation—unless participants were explicitly instructed to relate them for the sense judgment. Nevertheless, the task manipulation did not affect the appearance of the ELAN, even though its morphology and distribution in the word category violation and combined violation conditions differed.

Friederici, Gunter, Hahne, & Mauth (2004) used essentially the same paradigm with certain crucial modifications: (1) verbs were used with prefixes that were ambiguous between verbal and nominal readings, with disambiguation in

the suffix (2) nominals beginning with the same prefixes were used in four of the filler conditions to prevent early prediction of word category, and (3) stimulus sentences were extended such that the main clause verb was not sentence-final. They reported an N400 to the semantic violation and a LAN + P600 complex to both the word category violation and the combined violation. In other words, the N400 was again suppressed, although possibly for the same reason as above and/or because of overlap with the subsequent P600, which was larger in response to the double violation than to the word category violation alone, indicating some degree of interaction between syntax and semantics. The elicitation of a late LAN rather than an ELAN (early negativity) in this study was attributed to the fact that the word category information was contained in the disambiguating suffix rather than the prefix of the critical words.

An independent line of research investigating morphosyntactic parsing preferences rather than outright violations *per se* also bears on the question of syntax/semantics interactions: van Berkum and colleagues have shown immediate ERP responses to dispreferred parses triggered solely by preceding referential discourse context with auditory (Brown, van Berkum, & Hagoort, 2000; van Berkum, Brown, Hagoort, & Switzerlood, 2003) and visual presentation (van Berkum, Brown, & Hagoort, 1999a). The experimental paradigm biased preference for or against a relative clause interpretation of a morphosyntactically ambiguous complementizer/relative pronoun ("dat", 'that') by prior introduction of either two unique or two non-unique referents. The use of a definite noun phrase to refer to one of two non-unique referents in the previous discourse elicited a LAN to the head noun. Thereafter, continuations consistent with a complement clause interpretation elicited a P600 when two non-unique referents had been introduced in the preceding discourse context (crucially, even when these were morphosyntactically incompatible with an interpretation of "dat" as a relative pronoun), while continuations consistent with a relative clause interpretation elicited a P600 when two independent and unique discourse referents had been introduced. In other words, discourse context influenced the initial parse of a structural ambiguity.

To sum up, simultaneous violations of syntax and semantics have resulted in the following reported ERP effects: no apparent differences (Ainsworth-Darnell et al., 1998), a larger N400 (Hagoort, 2003), suppression of the N400 (Hahne & Friederici, 2002; Friederici et al., 2004), a slightly smaller N400 and P600 (Osterhout & Nicol, 1999), a smaller P600 (Gunter et al., 1997), a larger P600 (Friederici et al., 2004), and increased anterior negativity of various types (Palohati et al., 2005; Friederici et al., 2004). Bear in mind that there were crucial differences across these studies: some manipulated morphosyntax while others manipulated subcategorization or word category information, some used visual and others auditory presentation. At a minimum, however, there do appear to be interactions of syntax with semantics in on-line processing, both in double violation studies and in manipulations of discourse context to influence morphosyntactic parsing. Moreover, these reported interactions appear to involve the N400, P600 and LAN equally, ERP components that do not emerge any earlier in the ERP record than 200 ms post stimulus onset. Whether or not such

interactions reliably involve or exclude the ELAN is a question that will have to await further research for a definitive answer.

6.2.2 Elicitation of unexpected ERP components

The above investigations of combined (morpho)syntactic and semantic violations were originally designed to test the exact nature of the relationship between the N400 and P600 components, although potential interactions involving other ERP components have also appeared. Beyond this, however, a spate of studies have reported manipulations that surprisingly elicited the opposite component from what one might have otherwise predicted, given the conception of the N400 as an index of semantic processing, and of the LAN, ELAN, and P600 as indices of morphosyntactic processing. These include studies that reported N400s in cases where one might have expected only P600s, and studies that reported P600s effects in contexts where one might have reasonably expected to see N400 amplitude modulations instead.

In the former category are German studies investigating case relations. German, with a relatively free ordering of arguments, relies heavily on case marking rather than on word order to determine thematic relations. Hopf, Bayer, Bader, & Meng (1998) placed case-ambiguous plural nouns at the beginning of German sentences, and verbs that assign either accusative (default) or dative (marked) case to their objects in clause-final position.

> Dirigenten…kann ein Kritiker ruhig UMJUBELN$_{ACC}$/APPLAUDIEREN$_{DAT}$…
> *Conductors can a critic safely celebrate applaud*
> 'Critics can safely celebrate/applaud conductors…'

Dative verbs with case-ambiguous objects ("Dirigenten") elicited a right-lateralized, posterior N400 relative to accusative verbs, and relative to dative verbs with objects in first position clearly marked for dative case ("Musikern", 'musicians-DAT' versus. "Musiker", 'musicians-ACC').

Frisch & Schlesewsky (2001) showed that when two adjacent animate (i.e. semantically reversible) noun phrases were both marked nominative, the second one elicited an N400 + P600 complex, even before the clause-final verb was reached.

> Paul fragt sich, welchen Angler DER JÄGER gelobt hat
> *Paul asks himself which-ACC fisherman the-$_{NOM}$ hunter praised has*
> 'Paul wonders which fisherman the hunter praised'

> *Paul fragt sich, welcher Angler DER JÄGER gelobt hat
> *Paul asks himself which-NOM fisherman the-$_{NOM}$ hunter praised has*
> (uninterpretable with both NPs marked nominative)

This N400 was widely distributed across the scalp, raising the possibility that part of the response may have been a LAN. When these two nominative NPs were semantically asymmetrical-one animate and the other inanimate-the second NP

elicited a P600 but no N400.

> Paul fragt sich, welchen Förster DER ZWEIG gestreift hat
> *Paul asks himself which-ACC forester the-$_{NOM}$ branch brushed has*
> 'Paul wonders which branch brushed the forester'
>
> *Paul fragt sich, welcher Förster DER ZWEIG gestreift hat
> *Paul asks himself which-NOM forester the-$_{NOM}$ branch brushed has*
> (uninterpretable with both NPs marked nominative)

Frisch & Schlesewsky attributed this to the fact that, with the thematic roles of the arguments clearly marked by animacy, the erroneous case marking affected only structure building operations (P600) rather than thematic interpretive processes (N400). Frisch & Schlesewsky (2005) replicated the N400 + P600 complex in response to the second of two animate arguments marked nominative, and showed the same response for two animate arguments in either the dative or the accusative case. Interestingly, the N400 to the second of two accusative NPs was larger than that to the second of two nominative NPs, while the P600 to the second of two dative NPs was larger than that to the second of two nominative NPs.

Bornkessel, McElree, Schlesewsky, & Friederici (2004) capitalized on the flexible ordering of arguments and verb-final word order in embedded clauses in German to create materials in which two case-ambiguous nouns, one singular and one plural, appeared adjacent to each other in an embedded clause, followed by a singular or plural verb. If the verb agreed in number with the first NP, then it had to be the subject and the second NP had to be the object (subject-object order); this is the default ordering of arguments in German. If the verb instead agreed with the second NP, then it had to be the subject of the clause and the first NP had to be the object (object-subject order). While fully grammatical, this is a marked order in German. Neither the thematic roles nor grammatical functions of these two NPs were apparent until the verb. Moreover, since both nouns were case ambiguous, it was possible to use verbs that assign accusative case to their theme objects as well as verbs that assign dative case. Accusative verbs that forced an object-first ordering of arguments elicited a P600 relative to the verbs of subject-first clauses (Friederici & Mecklinger, 1996), while dative verbs that forced an object-first ordering produced an N400. At this time, it is not entirely clear why reanalysis of case relations should elicit an N400 and reanalysis of phrase structural relations a P600. Bornkessel et al. (2004) concluded that case relations must be processed earlier than phrase structural relations.

A series of studies by various researchers during the same time period reported P600s to the verb in experimental sentences such as "For breakfast the eggs would only EAT toast and jam" compared to the same verb in control sentences such as "For breakfast the boys would only EAT toast and jam" (Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). This finding could be attributed to a mere animacy violation, as in Frisch & Schlesewsky (2001), which

would entail that this P600 be interpreted as indexing difficulty with structure-building operations. Kuperberg et al. (2003) essentially adopted this position, and Hoeks, Stowe, & Doedens (2004) likewise interpreted the P600 they observed to sentences containing animacy reversals like 'The javelin has THROWN the athletes' as indexing thematic processing difficulty, specifically "effortful syntactic processing … to obtain a semantically coherent and plausible sentence." (p. 72)

However, Kolk, Chwilla, van Herten, & Oor (2003) and van Herten, Kolk, & Chwilla (2005) also reported P600s to Dutch sentences that reversed expected pragmatic plausibility relations rather than animacy/thematic relations *per se* ('The fox that HUNTED the poachers stalked through the woods'; note that both NP arguments preceded the critical verb in Dutch). Similarly, Kim & Osterhout (2005) provided evidence suggesting that inanimate subject nouns do not lead to a P600 when associated with verbs that fail to encourage a pragmatically plausible combined interpretation (as in "The hungry tabletops were DEVOURING…"). On the other hand, Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb (*in press)* demonstrated that even when the inanimate subject noun and verb were semantically unrelated ("At breakfast the eggs would PLANT…"), the verb still elicited a broadly distributed P600 equivalent and at some sites larger than that to semantically related verbs (EAT). Moreover, van Herten et al (2005) provided evidence that the P600 was not due to a conflict between the expected and the actual inflectional ending on the verb. Finally, Kuperberg, Caplan, Sitnikova, Eddy, & Holcomb (*in press)* showed that the P600 to animacy violations in their stimuli was larger when the inanimate subject was a plausible argument of a transitive verb ("When John arrived at the restaurant, the food would ORDER…") than an implausible argument of an intransitive verb ("When they greeted the Queen of England, the trumpets would CURTSEY…"). While the P600 elicited by morphosyntactic violations ("For breakfast the boys would only EATS…") was larger than that to thematic role animacy violations ("For breakfast the eggs would only EAT…"), the two effects had an otherwise similar latency, morphology, and scalp distribution. As this line of research investigating unexpected N400 and P600 effects is currently very much a work in progress, one can reasonably anticipate additional clarity as more studies emerge in ensuing years.

6.3 Influence of non-linguistic cognitive variables on syntactic processing

6.3.1 Relationship of the P600 to the P3b

The current state of affairs regarding the relationship of the P600 (as a marker of morphosyntactic processing) to the N400 (as a marker of semantic processing)—as outlined in the preceding section—raises a number of general questions about the purported domain-specific nature of the P600 itself. N400s are known to be elicited by non-linguistic sources of semantic or pragmatic information such as line drawings and environmental sounds (see section *4*). Likewise, P600-like potentials have been elicited by spelling errors that leave pronunciation intact (Münte, Heinze et al. 1998), harmonic and melodic violations in music (Besson & Macar, 1987; Janata, 1995; Patel et al., 1998), and violations of geometry (Besson & Macar, 1987), arithmetic rules (Nieddeggen & Rösler,

1999; Nuñez-Peña & Honrubia-Serrano, 2004), and abstract sequences (Lelekov, Dominy & Garcia-Larrea, 2000). In addition to the P600s to thematic violations discussed at the end of the previous section, P600s have also been observed following N400s elicited by standard semantic anomalies (Münte, Heinze, et al., 1998), and following a LAN to semantic violations of hyponymy and negative polarity (Shao & Neville, 1996).

In view of these and other complicating factors, the question has arisen whether the P600 is more judiciously viewed as a manifestation of a domain-general brain response elicted by rare, informative events. More specifically, several studies (Gunter et al., 1997; Coulson et al., 1998; Hahne & Friederici, 1999, but see Osterhout & Hagoort, 1999 for a counterargument) have proposed that the P600 should be considered a member of the P300 family known as the P3b (Donchin & Coles, 1988). Since P3b amplitude is known to be sensitive to the probability of occurrence of relevant stimulus types, each of these studies contrasted the presentation of blocks containing relatively rare (20-25%) versus relatively frequent (75-80%) syntactic violations. In response to frequent syntactic violations, the P600 effect was either reduced or eliminated entirely. By way of contrast, these probability manipulations did not affect preceding LAN or ELAN responses in any of these studies (see also section 6.4).

To further test the relationship between the P600 and the P3b, two studies directly compared ERPs to infrequent auditory oddball tones (a standard paradigm for eliciting a P3b) versus syntactic violations in patients with brain lesions. Frisch et al. (2003) found that in contrast to brain-damaged controls, patients with lesions to the basal ganglia showed no P600s to sentences containing syntactic violations that were equiprobable (50%) with correct sentences, despite a preserved P3b to improbable (20%) high-pitched oddball tones. In Wassenaar et al. (2004), 11 out of 12 normal controls, all five patients with right hemisphere damage, but only six of ten Broca's aphasics with frontal lesions (extent of basal ganglia damage unreported) showed a P600 to equiprobable subject-verb agreement violations, while all but two Broca's aphasics showed a P3b in an auditory tone discrimination task (20% low tones). Both studies thus claimed that the P600 and P3b were independent, that the P600 cannot be a general cognitive response of surprise, context updating, task relevance, salience, or probability, and that it must have a different neural substrate from the P3b.

6.3.2 Role of working memory in syntactic processing

In an early ERP study of English *wh*-questions, Kluender & Kutas (1993a,b) showed in word-by-word comparisons of sentence positions intervening between filler and gap that, relative to *yes/no*-questions, object *wh*-questions consistently elicited greater negativity over left anterior scalp between 300 and 500 ms. Subject *wh*-questions did not show this difference, presumably because the working memory load was no higher than that incurred in *yes/no*-questions. Direct comparisons of subject- versus object-relative clauses in the visual (King & Kutas, 1995) and auditory (Mueller, King, & Kutas, 1997) modalities in English revealed that these transient effects were likely time slices

of longer-lasting, slow anterior negative potentials, left-lateralized with visual (but not auditory) presentation.

In a related paradigm comparing biclausal structures differing only in the first word ("AFTER/BEFORE the scientist submitted the paper, the journal changed its policy"), Münte, Schiltz, and Kutas (1998) showed that BEFORE sentences with reversed chronological order likewise elicited slow negative potentials over left anterior sites relative to AFTER sentences. These differences were larger in participants with higher reading spans, a measure of verbal working memory capacity.

Relative to grammatical subject *wh*-questions, Kluender & Münte (1998) showed anterior negativities to both long-distance object *wh*-questions and to *that*-trace violations. As the negativity to object *wh*-questions consisted of slow frontal potentials (as in King & Kutas 1995) that could also be seen in individual word responses, while that to *that*-trace violations consisted solely of left-lateralized phasic responses, Kluender & Münte hypothesized that there may be a more local LAN related to morphosyntactic processing, and a more global LAN related to processes of verbal working memory.

However, Vos, Gunter, Kolk, & Mulder (2001) crossed morphosyntactic (subject-verb agreement) violations with syntactic complexity (a conjoined versus subject relative clause intervening between subject and verb), additional working memory load (an extraneous monitoring task for one versus three words in the stimulus sentences), and verbal working memory span (based on a listening span test) to show that LAN amplitude was influenced by all of these factors. In other words, these ERP responses suggested that working memory processes do interact with syntactic processing. This evidence was taken as support for the Just & Carpenter (1992) single resource (for both storage and processing) model of verbal working memory, rather than the Waters & Caplan (1996) model of two separate verbal working memory systems for automatic (e.g. syntactic) versus controlled processes.

Subsequent studies of long-distance object dependencies of various sorts both replicated and filled out this picture. Kaan, Harris, Gibson, & Holcomb (2000) failed to replicate slow left-lateralized anterior negative potentials between filler and gap in object dependencies, but did demonstrate P600s to gap-filling at the subcategorizing verb that were partly independent of and yet also partially overlapping with P600s to morphosyntactic anomalies. This P600 was interpreted as an index of syntactic integration costs at the subcategorizing position.

Fiebach, Schlesewsky, & Friederici (2001, 2002) compared indirect *wh*-questions that differed in how many adverbial adjunct prepositional phrases (PPs) occurred between the clause-initial *wh*-filler and the second NP argument, followed by the clause-final verb — one PP (short) or two (long). Only long object *wh*-questions elicited the slow left-lateralized anterior negative potentials reported in earlier studies. In contrast to Münte et al. (1998), however, it was the low (not the high) span readers who produced the larger slow potential effect. Fiebach et al. also replicated in German the P600 to gap-filling in object questions reported by Kaan et al. (2000) for English; however, the P600 was

seen already at the second NP (i.e. the subject NP with nominative case marking in object *wh*-questions) and not at the clause-final verb, indicating that the parser did not wait until the subcategorizer was reached before attempting to fill gaps in this verb-final language. Moreover, unlike the slow left-lateralized anterior negative potential between filler and gap, P600 amplitude to gap-filling did not differ across reading span groups, buttressing Kaan et al.'s (2000) claim that the P600 to gap-filling indexes syntactic integration rather than memory storage costs. While only long object *wh*-questions elicted sustained anterior negativity related to holding a filler in working memory, both long and short object *wh*-questions exhibited P600 effects related to gap-filling, albeit of larger amplitude for long object *wh*-questions.

Felser, Clahsen, & Münte (2003) reported slightly different results for long-distance object dependencies in German. Their stimuli differed from those in Fiebach et al.(2001, 2002) in the following ways. In addition to a *wh*-question condition, they included other types of filler-gap dependencies: a raising condition and two topicalization conditions. However, only object dependencies were included in the design and compared to each other; even the control condition involved short- (rather than long-)distance topicalization of a direct object. This effectively eliminated the possibility of monitoring for slow negativity between filler and gap. Also, the length of the dependency was increased by inserting an adjunct adverbial clause between filler and gap.

A transient LAN was elicited by all three long-distance filler-gap dependencies at the subject NP immediately following the intervening adverbial clause; this response was taken to index retrieval of the filler from working memory following processing of the adverbial clause. However, this LAN did not persist into the following indirect object NP immediately preceding the final verb. Contra Fiebachet al. (2002), there was no P600 at either NP prior to the final verb to indicate early syntactic integration of the filler. Instead, the long-distance topicalization and *wh*-question conditions elicited a phasic LAN relative to the raising construction at the final verb, and the *wh*-question condition also elicited a P600. What remains unclear is (1) why the purported syntactic integration of the filler in the *wh*-question condition would be delayed until the final subcategorizing verb position, and (2) why the filler in the long-distance topicalization condition would not undergo a similar process of integration at this point in the sentence.

Phillips, Kazanina, and Abada (2005) also reported a somewhat divergent set of findings from English. They contrasted long ("The lieutenant knew WHICH ACCOMPLICE the detective hoped that the shrewd witness would RECOGNIZE in the line-up") versus short ("The detective hoped that the lieutenant knew WHICH ACCOMPLICE the shrewd witness would RECOGNIZE in the lineup") embedded *wh*-object dependencies; the control condition consisted of a series of complement declarative clauses ("The lieutenant knew that the detective hoped that the shrewd witness would RECOGNIZE the accomplice in the lineup"). The long *wh*-question condition elicited a sustained anterior negativity beginning with the *wh*-phrase and continuing throughout the dependency, similar to King and Kutas (1995), Kluender & Münte (1998), and Fiebach et al. (2001, 2002). Single-word averages in the long *wh*-condition likewise revealed phasic LAN effects at

the two words immediately following the *wh*-phrase, similar to Kluender & Kutas (1993a,b). Finally, there were P600s at the final verb (RECOGNIZE) in both *wh*-conditions, albeit between 300 and 500 ms in the short *wh*-condition and between 500 and 700 ms in the long *wh*-condition. The occurrence of the P600 at the subcategorizing verb was consistent with Kaan et al. (2000) for English, but the fact that the P600 difference between the long and short *wh*-conditions manifested as a latency rather than an amplitude difference was at odds with the results of Fiebach et al. (2002) for German.

By subjecting the slow negative potential to the long *wh*-condition to both a cumulative (cf. Fiebach et al. 2002) and a non-cumulative analysis (cf. King & Kutas 1995), Phillips et al. (2005) showed that it did not increase in amplitude across the course of the sentence. Rather, it increased only across the intermediate clause containing the bridge verb ("the detective hoped that"); there was no further growth of the slow potential in the most deeply embedded clause ("the shrewd witness would"). The latter was true of the short *wh*-condition as well. Consequently, Phillips et al. claimed that while the sustained negativity likely indexed holding the *wh*-phrase in working memory, it was technically not sensitive to length, and therefore should not be viewed as an index of memory storage costs increasing over the course of a dependency. They attributed this discrepancy in interpretation with that proposed in Fiebach et al. (2002) to the greater length (in number of intervening words) and complexity (in terms of the intervening clause boundary) in their own study. As the length manipulation affected P600 latency rather than amplitude, Phillips et al. characterized P600 latency as an index of filler reactivation, a length-sensitive subprocess of syntactic integration (Gibson, 2000), and P600 amplitude as an index of thematic role assignment and compositional semantic interpretation, integration processes insensitive to length.

ERPs to scrambling in the German *Mittelfeld* ('middle field')—i.e. ordering permutations of subject, indirect object, and direct object, the canonical word order in German—were reported in Rösler et al. (1998). These were generally consistent with effects reported for *wh*-movement manipulations: LAN effects to NPs scrambled leftward—that is, occurring in non-canonical positions—and broad P600-like effects beginning at the final argument of verb-final sentences with non-canonical argument ordering. Ueno & Kluender (2003a) showed related effects in Japanese when object NPs were scrambled leftward into non-canonical sentence positions (OS__V rather than canonical SOV order): slow anterior negative potentials between filler and gap, and both phasic LANs and P600s to gap-filling, which again occurred before the subcategorizing clause-final verb was reached. Both scrambling studies pointed to the same conclusion, namely that the parser actively tries to restore constituents to their canonical underlying positions when faced with non-canonical permutations of word order. This finding may be not unrelated to the sensitivity of the parser to perturbations of the canonical chronological ordering of clausal events (Münte et al. 1998).

Two published studies of verb gapping ("Ron took the planks for the bookcase, and Bill __ the HAMMER with the big head"; Kaan, Wijnen, & Swaab, 2004) reported ELAN-like effects at the word immediately following the gap, a

fronto-central negativity (120-200 ms) in German (Streb, Hennighausen, & Rösler, 2004), and a central-posterior negativity (100-200 ms) in English (Kaan, et al., 2004). Both studies also reported subsequent positivities: a fronto-central positivity (300-500 ms) in English (Kaan et al., 2004), and an apparent (although unmeasured and unreported) widespread midline positivity (400-600 ms) in German (Streb et al., 2004).

Follow-up work will be required to make sense of these commonalities, but based on the ERP evidence available thus far, it can be said with some measure of confidence that the mental processes involved in leftward scrambling appear to be quite similar to those involved in *wh*-movement, while the processes involved in verb gapping appear to be qualitatively different from those active in filler-gap dependencies.

6.4 The fractionation of syntactic processing

The modularity hypothesis makes strong claims with regard to the encapsulation of language from other cognitive modules, and of syntax from semantics within the language module. As we have seen, ERP studies on aggregate point to a certain amount of cross talk among these domains. There appears to be some influence of non-linguistic factors (working memory and statistical probability) on ERP measures of human sentence processing, and purported ERP indices of syntactic and semantic processes also seem to interact. Within the syntactic processing module proper, very strong claims have likewise been made about the dissociation of a variety of syntax-related ERP effects. These are for the most part motivated by an ideological commitment to serial parsing models. This section reviews and evaluates some of these claims.

6.4.1 ELAN

Much discussion has focused on the existence of the early left anterior negativity or ELAN, first reported by Neville et al. (1991) using visual presentation and by Friederici et al. (1993) using auditory presentation. Hagoort, Wassenaar, & Brown (2003b) attempted to replicate the word category violation effect in Dutch using visual presentation with minimal pair sentences ('The lumberjack dodged the vain PROPELLER/PROPELLED on Tuesday'). Note that since both conditions contained a semantic violation, this was in some sense a semantic versus semantic plus morphosyntactic violation manipulation. The end result was that an anomalous noun was presented in one condition and an anomalous verb in the other— but, as the authors pointed out, with zero cloze probability in both cases, eliminating another potential confound. The avoidance of the missing head noun confound in the Neville et al. (1991; "What did the scientist criticize Max's OF…?') and in the Friederici et al. (1993; 'The ice cream was in-the EATEN') word category violation stimulus paradigms was another advantage. The word category violations in this study ('The lumberjack dodged the vain PROPELLED…') elicited a LAN + P600 complex, but no ELAN, perhaps because the stimuli were presented visually.

Rossi, Gugler, Hahne, & Friederici (2005) borrowed the double violation paradigm used in studies of syntax-semantics interactions to attempt a

dissociation of early ERP effects related to word category information (ELAN) and to morphosyntax (LAN) using auditory materials. Because the stimuli were constructed in active rather than passive voice, they translate more or less directly into English (results indicated in parentheses): 'The boy in kindergarten SINGS/SANG a song', 'The boy in SINGS/SANG a song' (ELAN, LAN, P600), 'The boy in kindergarten SING/SANGS a song' (LAN, P600), and 'The boy in SING/SANGS a song' (ELAN, LAN, P600). Because of the differences in the words immediately preceding the verb across conditions, a post-stimulus baseline of 100 ms was used. Both the word category violation and the combined violation produced continuous negativity between 100 and 600 ms, whereas the subject-verb agreement violation by itself showed a significant LAN difference with the control condition only between 450 and 650 ms. Rossi et al. (2005) interpreted this as an indication that the processing of word category information takes primacy over the processing of other types of syntactic information like morphosyntax.

6.4.2 Is there more than one P600?

In Mecklinger, Schriefers, Steinhauer, & Friederici (1995), fast comprehenders (i.e. participants who responded to comprehension questions with RTs < 800 ms) produced an enhanced early positivity beginning in the P200 peak in response to the sentence-final auxiliary in object relative clauses containing only case-ambiguous feminine nouns. The use of case-ambiguous nouns delayed resolution of the subject versus object relative clause reading until the clause-final verb complex. Mecklinger et al. suggested that this early positivity (P345) indexed a rapid revision of the parse—while leaving hierarchical phrase structure intact—toward the dispreferred object relative clause interpretation. The same comparison of relative clauses was contrasted with an analogous comparison of SOV vs. OSV complement clauses in Friederici, Mecklinger, Spencer, Steinhauer, & Donchin (2000). Object complement clauses elicited only a late positivity (500-900 ms), whereas object relative clauses elicited both an early (300-500 ms) and a late positivity (contra Mecklinger et al., 1995). , This was attributed to the wider variety of materials used.

Hagoort, Brown, & Osterhout (1999) claimed that the revision of syntactic ambiguity toward dispreferred continuations elicits more frontally distributed P600 effects, while the repair of ungrammatical sequences elicits more posterior P600 effects. To test this, Friederici, Hahne, & Saddy (2002) crossed grammaticality (subject-verb agreement violations) with complexity (topicalization of a simple noun phrase versus a more complex verb plus noun phrase complex). The grammaticality manipulation yielded an N400 + posterior P600 (500-1100 ms), while the complexity manipulation produced an earlier frontal positivity (500-700 ms) as well as a widespread later positivity (800-1100 ms). As the earlier frontal positivity in the complexity manipulation involved neither ambiguity nor a need for revision, it was interpreted as an index of structural complexity more broadly construed..

This picture was complicated by the results of Kaan et al. (2000), who reported a posterior (rather than anterior) P600, not in response to either an

outright violation or a dispreferred parse, but to a more complex structure. Kaan and Swaab (2003) investigated relative clause attachment ambiguity in order to shed further light on this issue. They compared preferred ("the cake beside the pizzas that WERE brought"), dispreferred ("the cakes beside the pizza that WERE brought"), and ungrammatical ("the cake beside the pizza that WERE brought") continuations of sentences. Relative to preferred continuations, both dispreferred and ungrammatical continuations elicited positivity with the same posterior distribution. When these stimulus materials with two attachment sites for the relative clause were compared to a simpler grammatical structure with only one attachment site for the relative clause, all of the more complex continuations elicited greater positivity at frontal electrodes, regardless of whether they were preferred, dispreferred, or ungrammatical. Taken together, these results cast doubt on the claim that dispreferred continuations elicit frontal positive differences in the ERP record, while ungrammatical continuations elicit positivities with posterior maxima (Hagoort et al. 1999). However, a discrepancy remains between the frontal positivity elicited by structural complexity in Friederici et al. (2002) on the one hand, and the posterior positivity elicited by structural complexity in Kaan et al. (2000) on the other.

Further studies have raised additional questions with regard to the proper functional characterization of the P600. For example, Frisch, Schlesewsky, Saddy, & Alpermann (2002) showed a P600 not only at the point of syntactic disambiguation (in a sentence ambiguous between SOV and OSV word order), but already at the introduction of syntactic ambiguity into the parse itself (a case-ambiguous feminine noun occurring as the first argument in a sentence). This suggested that more than one word order alternative was under consideration during the ambiguous region, in line with predictions of parallel rather than serial parsing models, where the simplest structural alternative is always preferred initially. Bornkessel, Schlesewsky, & Friederici (2002) showed that the clause-final verbs of sentences in which phrase structure and case marking were held constant, but thematic relations varied based on the choice of predicate (e.g. a psychological predicate assigning the roles of theme and experiencer to the subject and object vs. a standard transitive predicate assigning the roles of agent and patient) elicited a late positivity between 200 and 600 ms when the case marking was unambiguous (i.e. using masculine nouns in German). NPs ambiguously marked for case (i.e. proper names and feminine nouns in German) elicited no such difference. Bornkessel et al. argued that late positivities should therefore be redefined as indexing more general hierarchical rather than purely syntactic information.

6.5 Morphosyntactic processing: conclusions

This section started out with relatively unequivocal and self-assured statements about the dissociation of semantic and syntactic processes as indexed by the N400 and the LAN/P600, respectively, as first suggested by Kutas & Hillyard (1983). We have now in some sense come full circle, however, as studies have begun to blur the lines of this once pleasantly simple picture. As outlined in section *6.2.2,* it now seems incontestable that at least certain

manipulations of case relations in German elicit N400s, and that certain manipulations of pragmatic plausibility at the lexical level elicit P600s.

As for the questions with which this section started out, the outlines of some of the answers have begun to emerge. First, with regard to the encapsulation versus interaction of syntactic and semantic processes, there is a fair amount of evidence that syntactic and semantic information do interact to some degree, although not with 100% predictability, and not always in the same way. Second, as to the influence of other, non-linguistic cognitive variables on syntactic processing, there seems to be a consensus that working memory does play a major role in syntactic processes (albeit to varying degrees in different individuals), but is not itself syntax-specific. Likewise, it seems undeniable that P600-like effects have been elicited across a number of related cognitive domains. While lesion studies have demonstrated that the P3b to auditory tone discrimination tasks can be preserved when the P600 to syntactic violations is compromised, it seems unrealistic to expect complete overlap of brain representation for auditory tone discrimination and syntactic processing of language when even the auditory and visual P3b generators are not identical.

Perhaps the least amount of consensus is found regarding the fractionation of syntactic processing into discrete stages. While several neural models of language processing are currently available that may ultimately prove useful in having helped to shape and clarify our thinking about language processes in the brain (Friederici, 2002; Hagoort, 2003b), it is at present not clear whether our ability to theorize may have already outstripped our existing empirical base. Clearly, we still do not understand completely what the N400 indexes, let alone the exact nature of the more recently discovered language-related ERP components. By continuing to rely on the same types of known experimental paradigms, we may thus be putting the cart before the horse, and consequently not make as much progress as one might wish for. Because if we rush to load up the cart by assigning premature functional significance to differences that are relatively easy to come by, we may never get around to feeding the horse that's supposed to be pulling the wagon in the first place.

One possible way of getting around to this is addressed in Frisch, Hahne, & Friederici (2004):

> "…one has to keep in mind that our paradigm is somewhat 'artificial' in that sentences with violations like the ones we used rarely occur in normal processing. This argument…applies to all paradigms testing ungrammatical structures…seeing that violations seem to produce especially clear changes in the electrical activity of the brain." (p. 215)

This is an important acknowledgement that should not be completely or complacently ignored. As a research strategy, continuing to pursue the study of violation types may not necessarily answer as many questions as it raises. Reversing this trend requires a commitment to taking the road less traveled and sticking to it in small, systematic steps—an approach to which the results and ensuing inferences of language ERP studies sans violations can already attest.

FOOTNOTES

1. This is the standard result in adults fluent in a language differs from that of children acquiring their first language. Words known to 13-20 month-olds elicit larger N400s than both unknown real words and pseudowords (Mills, Coffey-Corina, & Neville, 1997; Mills, et al., 2004). When infants learn associations between pseudowords and novel objects, N400 amplitude increases (Mills, Plunkett, Prat, & Schafer, 2005). These developmental results indicate that the neural systems for word recognition and retrieval of meaning function somewhat differently in the immature state, before many words have been acquired and their meanings organized in an adult manner. Infant N400 effects resemble adult N400s over posterior sites, but are accompanied by large frontal negativities atypical of adult comprehension studies. Friedrich and Friederici (2004) similarly observed that unrelated picture-word pairs elicit more frontal N400 context effects in 19 month-olds than in adults. Overall, frontal cortex may be more critically involved in word comprehension during first-language learning than adulthood.

2. Hagoort et al. (2004) did, however, observe that compared to the correct sentence continuations, the world-knowledge violations elicited increased power in the gamma (30-70 Hz) frequency band, during roughly the same latency range as the N400, whereas the semantic violations did not (see Makeig, Debener, Onton, & Delorme, 2004 for the value of mining information from different EEG frequency bands after stimulus presentation).

REFERENCES

Ainsworth-Darnell, K., Shulman, H. G., and Boland, J. E. (1998). Dissociating brain responses to syntactic and semantic anomalies: Evidence from event-related potentials. *Journal of Memory and Language, 38*, 112-130.

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes, 18*, 405-430.

Allison, T., McCarthy, G., Nobre, A. C., Puce, A., & Belger, A. (1994). Human extrastriate visual cortex and the perception of faces, words, numbers, and colors. *Cerebral Cortex, 5,* 544-554.

Allison, T., Puce, A., Spencer, D. D., & McCarthy, G. (1999). Electrophysiological studies of human face perception. I: Potentials generated in occipitotemporal cortex by face and non-face stimuli. *Cerbral Cortex, 9,* 415-430.

Allison, T., Wood, C.C., & McCarthy, G. (1986). The central nervous system. In M.G.H. Coles, E. Donchin, & S.W. Porges, *Psychophysiology: Systems, Processes, and Applications* (pp. 5-25), New York, N.Y..: Guilford Press.

Anderson, J., & Holcomb, P. (1995). Auditory and visual semantic priming using different stimulus onset asynchronies: An event related brain potential study. *Psychophysiology, 32*, 177-190.

Arezzo, J. C., Vaughan, H. G., Kraut, M. A., Steinschneider, M., & Legatt, A. D. (1986). Intercranial generators of event-related potentials in the monkey. In R. Q. Cracco & I. Bodis-Wollner (Eds.), *Evoked potentials: Frontiers of clinical neuroscience* (pp. 174-189). New York: Alan R. Liss, Inc.

Armstrong, E., Schleicher, A., Omran, H., Curtis, M., & Zilles, K. (1995). The ontongeny of human gyrification. *Cerebral Cortex, 5,* 56-63.

Barber, H., Vergara, M., & Carreiras, M. (2004). Syllable-frequency effects in visual word recognition: Evidence from ERPs. *Neuroreport15*, 545-548.

Beauducel, A., & Debener, S. (2003). Misallocation of variance in event-related potentials: simulation studies on the effects of test power, topography, and baseline-to-peak versus principal component quantifications. *Journal of Neuroscience Methods, 124,* 103-112.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials associated with semantic priming. *Electroencephalography and Clinical Neurophysiology, 60*, 343-355.

Bentin, S., Mouchetant-Rostaing, Y., Giard, M. H., Echallier, J. F., & Pernier, J. (1999). ERP manifestations of processing printed words at different psycholinguistic levels: Time course and scalp distribution. *Journal of Cognitive Neuroscience, 11,* 235- 260.

Besson, M., & Macar, F. (1987). An event-related potential analysis of incongruity in music and other non-linguistic contexts. *Psychophysiology, 24*, 14-25.

Besson, M., Fischler, I., Boaz, T., & Raney, G. (1992). Effects of automatic associative activation on explicit and implicit memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18,* 89-105.

Boddy, J. (1981). Evoked potentials and the dynamics of language processing. *Biological Psychology, 13,* 125-140.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A.. D. (2004).

Multidimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language, 51,* 495-522.
Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Beyond syntax: Languagerelated positivities reflect the revision of hierarchies. *NeuroReport, 13*, 361-364.
Bransford, J. D., & Johnson, M. K. (1972). Contextual prerequisites for understanding: Some investigations of comprehension and recall. *Journal of Verbal Learning and Verbal Behavior, 11,* 717-726.
Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience, 11*, 261-281.
Brown C.M., van Berkum J.J.A., & Hagoort, P. (2000). Discourse before gender: An event-related brain potential study on the interplay of semantic and syntactic information during spoken language understanding. *Journal of Psycholinguistic Research*, 29, 53-68.
Brownell, H., Simpson, T., Bihrle, A., & Potter, H. (1990). Appreciation of metaphoric alternative word meanings by left and right brain-damaged patients. *Neuropsychologia, 28,* 375-383.
Buchwald, J. S. & Squires, N. S. (1982). Endogenous auditory potentials in the cat: A P300 model. In C. Woody (Ed.), *Conditioning: Representation of involved neural function* (pp. 503-515). New York: Raven Press.
Chiappa, K. H. (1983). *Evoked potentials in clinical medicine.* New York: Raven Press.
Chwilla, D. J., Brown, C. M., & Hagoort, P. (1995). The N400 as a function of the level of processing. *Psychophysiology, 32*, 274-285.
Collins, A., & Loftus, E. (1975). A spreading activation theory of semantic processing. *Psychological Review, 82,* 407-428.
Connolly, J. F., & Phillips, N. A. (1994). Event-related potential components reflect phonological and semantic processing of the terminal words of spoken sentences. *Journal of Cognitive Neuroscience, 6,* 256-266.
Coulson, S. (2001). *Semantic Leaps. Frame-shifting and conceptual blending in meaning construction*. Cambridge, MA: Cambridge University Press.
Coulson, S., Federmeier, K., Van Petten, C., & Kutas, M. (2005). Right hemisphere sensitivity to word and sentence level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 129-147.
Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes, 13*, 21-58.
Coulson, S.,& Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders, *Neuroscience Letters, 316*, 71-74.
Coulson, S., & Lovett, C. (2004). Handedness, hemispheric asymmetries, and joke comprehension. *Cognitive Brain Research, 19*, 275-288.
Coulson, S., & Van Petten, C. (2000). Lexical priming and sentence processing. Presentation at Thirteenth CUNY sentence processing conference, San Diego, March 2000.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory and Cognition, 30*, 958-968.
Coulson, S., & Williams, R. F. (2005). Hemispheric asymmetries and joke comprehension. *Neuropsychologia, 43*, 128-141.
Coulson, S., & Wu, Y-C. (2005). Right hemisphere activation of joke-related information: An event-related brain potential study. *Journal of Cognitive Neuroscience, 17*, 494-506
Deacon, D., Dynowska, A., Ritter, W., & Grose-Fifer, J. (2004). Repetition and semantic priming of nonwords: Implications for theories of N400 and word recognition. *Psychophysiology, 41*, 60-74.
DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience, 8,* 1117-1122.
Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134,* 9-21.
Donchin, E., & Coles, M. G. H. (1988). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences, 11*, 357-374.
Donchin, E., Ritter, W., & McCallum, W. C. (1978). Cognitive psychophysiology: The endogenous components of the ERP. In E. Callaway, P. Tueting, S. H. Koslow (Eds.), *Brain event-related potentials in man* (pp. 349-441). New York: Academic Press.
Doyle, M. C., Rugg, M. D., & Wells, T. (1996). A comparison of the electrophysiological effects of formal and repetition priming. *Psychophysiology, 33*, 132-147.
Elger, C. E., Grunwald, T., Lehnhertz, K., Kutas, M., Helmstaedter, C., Brockhaus, A., Van Roost, D., & Heinze, H. J. (1997). Human temporal lobe potentials in verbal learning and memory processes. *Neuropsychologia, 35,* 657-667.
Elman, J. (2004). An alternative view of the mental lexicon. *Trends in the Cognitive Sciences, 8,* 301-306.
Embick, D., Hackl, M., Schaeffer, J., Kelepir, M., & Marantz, A. (2001). A magnetoencephalographic component whose latency reflects lexical frequency. *Cognitive Brain Research, 10*, 345-348.
Federmeier, K. D., & Kutas, M. (1999a). A rose by any other name: Long term memory structure and sentence processing. *Journal of Memory and Language, 41,* 469- 495.
Federmeier, K. D., & Kutas, M. (1999b). Right words and left words: Electrophysiological evidence for hemispheric differences in meaning processing. *Cognitive Brain Research, 8,* 373-392.
Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs, and class-ambiguous words in context. *Brain, 123*, 2552-2566.
Felser, C., Clahsen, H., & Münte, T. F. (2003). Storage and integration in the processing of filler-gap dependencies: An ERP study of topicalization and *wh*-movement in German. *Brain and Language, 87*, 345-354.

Fernández, G., Heitkemper, P., Grunwald, T., Van Roost, D., Urbach, H., Pezer, N., Lehnertz, K., & Elger, C. E. (2001). Inferior temporal stream for word processing with integrated mnemonic function. *Human Brain Mapping, 14*, 251-260.

Ferretti, T. R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44,* 516-547.

Fiebach, C. J., Schlesewsky, M., & Friederici, A. D. (2001). Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. *Journal of Psycholinguistic Research, 30*, 321-338.

Fiebach CJ, Schlesewsky M and Friederici AD (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German WH-questions. *Journal of Memory and Language*, 47, 250-272.

Fischler, I. S., Bloom, P. A., Childers, D. G., Roucos, S. E., & Perry, N. W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology, 20,* 400- 409.

Fischler, I., Childers, D. G., Achariyapaopan, T., & Perry, N. W., Jr. (1985). Brain potentials during sentence verification: Automatic aspects of comprehension. *Biological Psychology, 21,* 83-105.

Fodor, J. A. (1983). Modularity of mind: An essay on faculty psychology. Cambridge, MA: MIT Press.

Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The turple effect. *Memory and Cognition, 30,* 1106-1117.

Friedrich, M., & Friederici, A. D. (2004). N400-like semantic incongruity effect in 19-month-olds: Processing known words in picture contexts. *Journal of Cognitive Neuroscience, 16,* 1465-1477.

Friederici AD (2002). Towards a neural basis for auditory sentence processing, *Trends in Cognitive Sciences*, 6, 78-84.

Friederici, A. D., von Cramon, D. Y., & Kotz, S. A. (1999). Language related brain potentials in patients with cortical and subcortical left hemisphere lesions. *Brain, 122,* 1033-1047.

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth (2004). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport 15*, 165-169.

Friederici, A. D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: Early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22* (5), 1219-1248.

Friederici, A. D., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31*, 45-63

Friederici, A. D., Hahne, A., & von Cramon, D. Y. (1998). First-pass versus second-pass parsing processes in a Wernicke's and a Broca's aphasic: Electrophysiological evidence for a double dissociation. *Brain and Language, 62,* 311-241.

Friederici, A. D., & Kotz, S. A. (2003). The brain basis of syntactic processes: Functional imaging and lesion studies. *NeuroImage, 20*, S8-S17.
Friederici, A. D., & Mecklinger, A. (1996). Syntactic parsing as revealed by brain responses: First-pass and second-pass parsing processes. *Journal of Psycholinguistic Research, 25*, 157-176.
Friederici, A. D., Mecklinger, A., Spencer, K. M., Steinhauer, K., & Donchin, E. (2001). Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. *Cognitive Brain Research, 11*, 305- 323.
Friederici, A. D., Pfeifer, E., and Hahne, A. (1993). Event-related brain potentials during natural speech processing: Effects of semantic, morphological and syntactic violations. *Cognitive Brain Research, 2,* 183-192.
Frisch, S., Kotz, S.A., von Cramon, D., & Friederici, D. (2003). Why the P600 is not just a P300: The role of the basal ganglia. *Clinical Neurophysiology, 14,* 336-340.
Frisch, S., Hahne, A., & Friederici, A. D. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition, 91*, 191-219.
Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *NeuroReport, 12*, 3391-3394.
Frisch, S., & Schlesewsky, M. (2005). The resolution of case conflicts from a neurophysiological perspective. *Cognitive Brain Research, 25*, 484-498.
Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition, 85*, B83-B92.
Gagnon, L., Goulet, P., Giroux, F., & Joanette, Y. (2003). Processing of metaphoric and non-metaphoric alternative meanings of words after right- and left-hemisphere lesions. *Brain and Language, 87,* 217-226.
Ganis, G., Kutas, M., & Sereno, M. I. (1996). The search for common sense: An electrophysiological study of the comprehension of words and pictures in reading. *Journal of Cognitive Neuroscience, 8*, 89-106.
Garnsey, S. (1985). Function and content words: Reaction time and evoked potential measures of word recognition. (Tech Report No. URCS 29). Rochester N.Y., Univerisity of Rochester.
Gibson, E. (2000). The dependency locality theory: A distance based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (pp. 94-126). Cambridge, MA: The MIT Press.
Gratton, G., & Fabiani, M. (2001). The event-related optical signal: A new tool for studying brain function. *International Journal of Psychophysiology, 42,* 109-121.
Gratton, G., Fabiani, M., Elbert, T., & Rockstroh, B. (2003). Seeing right through you: Applications of optical imaging to the study of the human brain. *Psychophysiology, 40,* 487-491.
Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28,* 267-283.
Guerrera, C. (2004). Flexibility and constraint in lexical access: Explorations in

transposed-letter priming. Doctoral dissertation, University of Arizona.
Guillem, F., N'Kaoua, B., Rougier, A., & Claverie, B. (1996). Differential involvement of the human temporal lobe structures in short- and long-term memory processes assessed by intracranial ERPs. *Psychophysiology, 33,* 720-730.
Gunter, T.C., Friederici, A.D., & Hahne, A. (1999). Brain responses during sentence reading: Visual input affects central processes. *Neuroreport, 10,* 3175-3178.
Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology, 34,* 660-676.
Hagoort, P. (2003a). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience 15*, 883-899.
Hagoort P (2003b). How the brain solves the binding problem for language: a neurocognitive model of syntactic processing, *Neuroimage*, 20, S18-S29.
Hagoort P., & Brown, C. (2000). ERP effects of listening to speech compared to reading: The P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia,* 38, 1531-1549.
Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439-483.
Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp.273-316). Oxford, UK: Oxford University Press.
Hagoort, P., Brown, C. M., & Swaab, T. Y. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia and patients with right hemisphere lesions without aphasia. *Brain, 119,* 627-649.
Hagoort, P., Hald, L., Bastiaansen, M., & Peterson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science, 304,* 438- 441.
Hagoort, P., Wassenaar, M., & Brown, C. (2003a). Real-time semantic compensation in patients with agrammatic comprehension: Electrophysiological evidence for multiple-route plasticity. *Proceedings of the National Academy of Sciences of the U.S.A., 100,* 4340-4345.
Hagoort, P., Wassenaar, M., & Brown, C. (2003b). Syntax-related ERP-effects in Dutch. *Cognitive Brain Research, 16*, 38-50.
Hahne, A., Eckstein, K., & Friederici, A. D. (2004). Brain signatures of syntactic and semantic processes during children's language development. *Journal of Cognitive Neuroscience, 16,* 1302-1318.
Hahne, A., & Friederici, A.D., (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience, 11*, 194-205**.**
Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research 13,* 339-356.
Halgren, E. (1990). Insights from evoked potentials into the neuropsychological

mechanisms of reading. In A. B. Scheibel & A. F. Wechsler (Eds.), *Neurobiology of higher cognitive function* (pp. 103-149). New York: Guilford Press.
Halgren, E., Dhond, R., Christensen, N., Van Petten, C., Marinkovic, K., Lewine, J. D., & Dale, A. M. (2002). N400-like MEG responses modulated by semantic context, word frequency, and lexical class in sentences. *Neuroimage, 17,* 1101-1116.
Halliday, A. M. (1982). *Evoked potentials in clinical testing.* New York: Churchill-Livingstone.
Hämäläinen, M., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography – theory, instrumentation, and application to noninvasive studies of the working human brain. *Reviews of Modern Physics, 65,* 413-497.
Handy, T.C. (Ed.) (2006). *Event-Related Potentials: A Methods Handbook.* Cambridge, Mass: MIT Press.
Helenius, P., Salmelin, R., Service, E., & Connolly, J. F. (1998). Distinct time courses of word and context comprehension in the left temporal cortex. *Brain, 121,* 1133-1142.
Helenius, P., Salmelin, R., Service, E., Connolly, J. F., Leinonen, S., & Lyytinen, H. (2002). Cortical activation during spoken word segmentation in nonreading impaired and dyslexic adults. *The Journal of Neuroscience, 22,* 2936-2944.
Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19*, 59-73.
Holcomb, Grainger Holcomb, P. J., Grainger, J., & O'Rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience, 15*, 938-950.
Holcomb, P. J., Kounios, J., Anderson, J. E., & West, W. C. (1999). Dual-coding, context-availability, and concreteness effects in sentence comprehension: An electrophysiological investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 721-742.
Holcomb, P. J., & McPherson, W. B. (1994). Event-related brain potentials reflect semantic priming in an object decision task, *Brain and Cognition, 24*, 259-276.
Holcomb, P. J., & Neville, H. (1990). Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes, 5*, 281-312.
Holcomb, P. J., & Neville, H. J. (1991). Natural speech processing: An analysis using event-related brain potentials. *Psychobiology, 19,* 286-300.
Hopf, J.-M., Bayer, J., Bader, M., & Meng, M. (1998). Event-related brain potentials and case information in syntactic ambiguities. *Journal of Cognitive Neuroscience,* 10, 264-280.
Janata P (1995). ERP measures assay the degree of expectancy violation of harmonic contexts in music. *Journal of Cognitive Neuroscience*, 7, 153-164.
Just, M.A., & Carpenter, P.A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Kaan, E., Harris, T., Gibson, E., & Holcomb, P. J. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes, 15*, 159-201.

Kaan, E., & Swaab, T.Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15*, 98-110.

Kaan, E., Wijnen, F., & Swaab, T.Y. (2004). Gapping: Electrophysiological evidence for immediate processing of "missing" verbs in sentence.

Katz, A. N., Cacciari, C., Gibbs, R. W., Jr., & Turner, M. (1998). *Figurative language and thought.* New York: Oxford University Press.

Kazmerski, V. A., Blasko, D. G., & Dessalegn-Banchiamlack, G. (2003). ERP and behavioral evidence of individual differences in metaphor comprehension. *Memory and Cognition, 31*, 673-689.

Keurs ter, M., Brown, C. M., Hagoort, P., Stegeman, D. F. (1999). Electrophysiological manifestations of open- and closed-class words in patients with Broca's aphasia with agrammatic comprehension. An event-related brain potential study. *Brain, 122,* 839-854.

Khader, P. Scherag, A., Streb, J., & Rösler, F. (2003). Differences between noun and verb processing in a minimal phrase context: A semantic priming study using event-related brain potentials. *Cognitive Brain Research, 17*, 293-313.

Kim, A., & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52*, 205-222.

King, J. W., & Kutas, M. (1995). Who did what and when? Using word- and clause related ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience, 7*, 378-397.

King, J. W., & Kutas, M. (1998). Neural plasticity in the dynamics of human visual word recognition. *Neuroscience Letters*, 244, 61-64.

Kluender, R., & Kutas, M. (1993a). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5*, 196-214.

Kluender, R., & Kutas, M. (1993b). Subjacency as a processing phenomenon. *Language and Cognitive Processes, 8*, 573-633.

Kluender, R., & Münte, T. F. (1998, March). Subject/object asymmetries: ERPs to grammatical and ungrammatical *wh*-questions. *Poster presented at the 11th Annual CUNY Conference on Human Sentence Processing, New Brunswick, NJ.*

Knight, R.T., Scabini, D., Woods, D. L., & Clayworth, C. C, (1989). Contributions of temporal-parietal junction to the human auditory P3. *Brain Research, 502,* 109-116.

Kojima, T., & Kaga, K. (2003). Auditory lexical-semantic processing impairments in aphasic patients reflected in event-related potentials (N400). *Auris Nasus Larynx, 30,* 369-378.

Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85*, 1-36.

Kotz, S. A., & Friederici, A. D. (2003). Electrophysiology of normal and pathological language processing. *Journal of Neurolinguistics, 16,* 43-58.
Kotz, S. A., Frisch, S., von Cramon, D. Y., & Friederici, A. D. (2003). Syntactic language processing: ERP lesion data on the role of the basal ganglia. *Journal of the International Neuropsychological Society, 9*, 1053-1060.
Kounios, J., & Holcomb, P. J. (1992). Structure and process in semantic memory: Evidence from event-related potentials and reaction times. *Journal of Experimental Psychology, 121,* 460-480.
Kounios, J., & Holcomb, P. J. (1994). Concreteness effects in semantic processing: ERP evidence supporting dual-coding theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 804-823.
Kuperberg, G.R., Kreher, T., Sitnikova, T., Caplan, D., & Holcomb, P.J. (in press), The role of animacy and thematic relations in processing activate English sentences: Evidence from event-related potentials. Brain and Language.
Kuperberg, G. R., Sitnikova, T., Caplan, D., & Holcomb, P. J. (2003). Electrophysiological distinctions in processing conceptual relationships within simple sentences. *Cognitive Brain Research, 17*, 117-129.
Kuperberg, G. R., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P.J. (in press). Neural correlates of processing syntactic, semantic and thematic relationships in sentences. *Language and Cognitive Processes.*
Kutas, M. (1993). In the company of other words: Electrophysiological evidence for single-word and sentence context effects. *Language and Cognitive Process, 8*, 533-572.
Kutas, M., & Hillyard, S. A. (1980a). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203-205.
Kutas, M., & Hillyard, S. A. (1980b). Reading between the lines: Event-related brain potentials during natural speech processing. *Brain and Language, 11,* 354-273.
Kutas, M. & Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. *Memory and Cognition, 11*, 539-550.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161-163.
Kutas, M., Hillyard, S. A., & Gazzaniga, M. (1988). Processing of semantic anomaly by right and left hemispheres of commissurotomy patients. *Brain, 111,* 553-576.
Kutas, M., Lindamood, T. E., & Hillyard, S. A. (1984). Word expectancy and eventrelated brain potentials during sentence processing. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes* (pp. 217-237). Hillsdale, NJ: Erlbaum.
Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science, 197,* 792-795.
Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening, and signing. *Electroencephalography and Clinical Neurophysiology, Supplement 39,* 325-330.

Kutas, M. and Van Petten, C., Psycholinguistics electrified: event-related brain potential investigations. In M. A. Gernsbacher (Ed.), Handbook of Psycholinguistics, Academic Press, San Diego, 1994, pp. 83-143.
Kwon, H., Kuriki, S., Kim, J. M., Lee, Y. H., Kim, K., & Nam, K. (2005). MEG study on neural activities associated with syntactic and semantic violations in spoken Korean sentences. *Neuroscience Research, 51,* 349-357.
Lelekov T, Dominey PF and Garcia-Larrea L (2000). Dissociable ERP profiles for processing rules vs instances in a cognitive sequencing task. *Neuroreport:* 11, 1129-1132.
Makeig, S., Bell, A. J., Jung, T. P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. In D. Touretzky, M. Mozer & M. Hasselmo (Eds.), *Advances in Neural Information Processing Systems: Vol. 8* (pp. 145-151). Cambridge: MIT Press.
Makeig, S., Debener, S., Onton, J., & Delorme, A. (2004). Mining event-related brain dynamics. *Trends in Cognitive Science, 8,* 204-210.
Marchand, Y., D'Arcy, R. C. N., & Connolly, J. F. (2002). Linking neurophysiological and neuropsychological measures for aphasia assessment. *Clinical Neurophysiology, 113,* 1715-1722.
McCallum, W. C., Curry, S., & Hutch, S. (Eds.) (1993), Slow potential changes in the human brain. New York, N.Y., U.S., Plenum Press, 303 pp.
McCallum, W. C., Farmer, S. F., & Pocock, P. V. (1984). The effects of physical and semantic incongruities on auditory event-related potentials. *Electroencephalography and Clinical Neurophysiology, 59,* 477-488.
McCarthy, G., Wood, C. C., Williamson, P. D., & Spencer, D. D. (1989). Task dependent field potentials in human hippocampal formation. *Journal of Neuroscience, 9,* 4253-4268.
McCarthy, G., Nobre, A.C., Bentin, S., & Spencer, D.D., (1995), Language-related field potentials in the anterior-medial temporal lobe: I. Intracranial distribution and neural generators. *Journal of Neuroscience, 15*, 1080-1089.
McKinnon, R., & Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes, 11*, 495-523.
McRae, K., de Sa, V., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General, 126,* 99-130.
Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A.D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory and Cognition, 23* (4), 477-494.
Mills, D. L., Coffey-Corina, S. A., & Neville, H. J. (1997). Language comprehension and cerebral specialization from 13-20 months. *Developmental Neuropsychology, 13,* 397-446.
Mills, D. L., Plunkett, K., Prat, C., & Schafer, G. (2005). Watching the infant brain learn words: Effects of vocabulary size and experience. *Cognitive Development, 20,* 19-31.
Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F.

(2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month olds. *Journal of Cognitive Neuroscience, 16,* 1452-1464.
Morton, J. (1969). The interaction of information in word recognition. *Psychological Review, 76,* 165-178.
Müller, H. M., King, J. W., & Kutas, M. (1997). Event-related potentials to relative clause processing in spoken sentences. *Cognitive Brain Research, 5*, 193-203.
Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology, 47,* 36-45.
Münte, T. F, Matzke, T.F., & Johannes, S. (1997) Brain activity associated with syntactic incongruencies in words and pseudo-words. *Journal of Cognitive Neuroscience, 9*, 318-329.
Münte, T. F., Heinze, H. -J., Matzke, M., Wieringa, B. M., & Johannes, S. (1998). Brain potentials and syntactic violations revisited: No evidence for specificity of the syntactic positive shift. *Neuropsychologia, 36*, 217-226.
Münte, T. F., Schiltz, K., & Kutas, M. (1998). When temporal terms belie conceptual order. *Nature, 395*, 71-73.
Münte, T. F., Wieringa, B. M., Weyerts, H., Szentkuti, A., Matzke, M., & Johannes, S. (2001). Differences in brain potentials to open and closed class words: Class and frequency effects. *Neuropsychologia, 39,* 91-102.
Neely, J. H., & Keefe, D. E. (1989). Semantic context effects on visual word processing: A hybrid prospective/retrospective processing theory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 24, pp. 207-248). San Diego: Academic Press.
Neville, H. J., Mills, D. L., & Lawson, D. (1992). Fractionating language: Different neural systems with different sensitive periods. *Cerebral Cortex, 2,* 244-258.
Neville, H. J, Nicol, J. L., Barss, A., Forster, K. I., & Garrett, M. F. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience, 3*, 151-165.
Niedeggen M and Rosler F (1999). N400-effects reflect activation spread during arithmetic fact retrieval. *Psychological Science*, 10, 271-276.
Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature, 372,* 260-263.
Norris, D. (1986). Word recognition: Context effects without priming. *Cognition, 22,* 93-136.
Nunez, P. L. (1981). *Electric fields of the brain.* New York: Oxford University Press.
Nunez-Pena, M. I., & Honrubia-Serrano, M. L. (2004). P600 related to rule violation in an arithmetic task. *Cognitive Brain Research, 18*, 130-141.
Olichney, J., Van Petten, C., Paller, K., Salmon, D., Iragui, V., & Kutas, M. (2000). Word repetition in amnesia: Electrophysiological evidence of spared and impaired memory. *Brain, 123*, 1948-1963.
Osterhout, L., Bersick, M., & McKinnon, R. (1997). Brain potentials elicited by words: Word length and frequency predict the latency of an early negativity.

Biological Psychology, 46, 143-68.
Osterhout, L., & Hagoort, P. (1999) A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS-P300 debate, *Language and Cognitive Processes, 14*, 1-14.
Osterhout, L., & Holcomb, P. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31*, 785-806.
Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes, 14*, 283-317.
Paller, K.A., Kutas, M., & Mayes, A.R. (1987) Neural correlates of encoding in an incidental learning paradigm, *Electroencephalography and Clinical Neurophysiology, 67,* 360-371.
Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45,* 255-287.
Palohati, M., Leino, S., Jokela, M., Kopra, K., & Paavilainen, P. (2005). Event-related potentials suggest early interaction between syntax and semantics during on-line sentence comprehension. *Neuroscience Letters, 384,* 222-227.
Patel, A. D., Gibson, E., Ratner, J., Besson, M., & Holcomb, P. J. (1998). Processing syntactic relations in language and music: An event-related potential study. *Journal of Cognitive Neuroscience, 10*, 717-733.
Patterson, K., & Hodges, J. R. (2000). Semantic dementia: One window on the structure and organisation of semantic memory. In L. S. Cermak (Ed.), *Handbook of Neuropsychology, 2nd Edition, Volume 2, Memory and its disorders* (pp. 313-333). Amsterdam: Elsevier.
Perea, M., & Lupker, S.J. (2003). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory and Cognition, 31*, 829-841.
Phillips C, Kazanina N, Abada SH (2005). ERP effects of the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22, 407-428.
Pineda, J. A., Swick, D., & Foote, S. L. (1991). Noradrenergic and cholinergic influences on the genesis of P3-like potentials. In C. H. M. Brunia (Ed.), *Event-related brain research* (pp. 165-172). Amsterdam: Elsevier.
Plante, E., Van Petten, C., & Senkfor, A. J. (2000). Electrophysiological dissociation between verbal and nonverbal processing in learning disabled adults. *Neuropsychologia, 38,* 1669-1684.
Pylkkänen, L, Stringfellow, A., & Marantz, A. (2002). Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language, 81,* 666-678.
Pynte, J., Besson, M., Robichon, F-H., & Poli, J. (1996). The time-course of metaphor comprehension: An event-related potential study. *Brain and Language, 55*, 293-316.
Ratcliff, R., & McKoon, G. (1988). A retrieval theory of priming in memory. *Psychological Review, 95,* 385-408.

Rösler F, Pechmann T, Streb J, Roder B and Hennighausen E (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language*, 38, 150-176.

Rösler, F., Streb, J., Haan, H. (2001). Event-related brain potentials evoked by verbs and nouns in a primed lexical decision task. Psychophysiology, 38, 694-703.

Rossi, S., Gugler, M.F., Hahne, A., & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters, 384*, 228-233.

Rugg, M. D., & Nagy, M. E. (1987). Lexical contribution to nonword-repetition effects: Evidence from event-related potentials. *Memory and Cognition, 15*, 473-481.

Salmelin, R., Service, E., Kiesila, P., Uutela, K., & Salonen, O. (1996). Impaired visual word processing in dyslexia revealed with magnetoencephalography. *Annals of Neurology, 40*, 157-162.

Schendan, H., Ganis, G., & Kutas, M. (1998). Neurophysiological evidence for visual perceptual organization of words and faces within 150 ms. *Psychophysiology, 35*, 240-251.

Schmitt, B. M., Münte, T. F., & Kutas, M. (2000). Electrophysiological estimates of the time course of semantic and phonological encoding during implicit picture naming. *Psychophysiology, 37*, 473-484.

Shao, J., & Neville, H. J. (1996). ERPs elicited by semantic anomalies: Beyond the N400. In *Proceedings of the Third Meeting of the Cognitive Neuroscience Society* (p. 73). San Francisco, CA.

Sharkey, N. E. (1989). The lexical distance model and word priming. In *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 860-867). Hillsdale, NJ: Erlbaum.

Simon,G., Bernard,C., Largy, P., Lalonde, R., & Rebai, M. (2004). Chronometry of visual word recognition during passive and lexical decision tasks: An ERP investigation. International *Journal of Neuroscience, 114*, 1401-1432.

Simos, P. G., Basile, L.F. H., & Papanicolaou, A. C. (1997). Source localization of the N400 response in a sentence-reading paradigm using evoked magnetic fields and magnetic resonance imaging. *Brain Research, 762,* 29-39.

Smith, M. E.,& Halgren, E. (1987). Event-related potentials during lexical decision: Effects of repetition, word frequency, pronounceability. and concreteness. In R. Johnson Jr., J. W. Rohrbaugh, & R. Parasuraman (Eds.), *Current trends in event-related potential research: Electroencephalography and Clinical Neurophysiology* (Suppl. 40, pp. 417-421). Amsterdam: Elsevier.

Smith, M. E.,Stapleton, J. M., & Halgren, E. (1986). Human medial temporal lobe potentials evoked in memory and language tasks. *Electroencephalography and Clinical Neurophysiology, 63,* 145-159.

Spencer, K. D., Dien, J., & Donchin, E. (2001). Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology, 38*, 343-358.

St. George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive*

Neuroscience, 6, 71-83.
Stockall, L., Stringfellow, A., & Marantz, A. (2004). The precise time course of lexical activation: MEG measurements of the effects of frequency, probability, and density in lexical decision. *Brain and Language, 90*, 88-94.
Streb, J.,Henninghausen, E., & Rosler, F. (2004). Different Anaphoric Expressions are Investigated by Event-Related Brain Potentials. *Journal of Psycholinguistic Research, 33*, 175-201.
Swaab, T.Y.,Baynes, K., & Knight, R. T. (2002). Separable effects of priming and imageability on word processing: An ERP study. *Cognitive Brain Research, 15*, 99-103.
Swaab, T. Y.,Brown, C., & Hagoort, P. (1997). Spoken sentence comprehension in aphasia: Event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience, 9,* 39-66.
Swick, D. (2004). ERPs in neuropsychological populations. In T. C. Handy (Ed.), *Eventrelated potentials: A methods handbook* (pp. 299-322). Cambridge, MA: MIT Press.
Swick, D., Kutas, M., & Knight, R.T. (1998). Prefrontal lesions eliminate the LPC but do not affect the N400 during sentence reading. *Journal of Cognitive Neuroscience Supplement (March 18),* 29.
Tarkiainen, A., Cornelissen, P. L., & Salmelin, R. (2002). Dynamics of visual feature analysis and object-level processing in face versus letter-string perception. *Brain, 125*, 1225-1136.
Tarkiainen, A., Helenius, P., Hansen, P. C., Cornelissen, P. L., & Salmelin, R. (1999). Dynamics of letter string perception in the human occipitotemporal cortex. *Brain, 122*, 2119-2132.
Tartter, V. C., Gomes, H., Dubrovsky, B., Molholm, S., & Stewart, R. V. (2002). Novel metaphors appear anomalous at least momentarily: Evidence from N400. *Brain and Language, 80*, 488-509.
Ueno, M., & Kluender, R. (2003a). Event-related brain indices of Japanese scrambling. *Brain and Language, 86*, 243-271.
Ueno, M., & Kluender, R. (2003b). On the processing of Japanese *wh*-questions: Relating grammar and brain. *Proceedings of the 22nd West Coast Conference on Formal Linguistics, 22,* 491-504.
van Berkum, J. J. A., Brown, C. M., & Hagoort, P. M. (1999a). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language, 41*, 147-182.
Van Berkum, J. J. A., Brown, C. M., Hagoort, P., & Zwitserlood, P. (2003). Event-related brain potentials reflect discourse-referential ambiguity in spoken-language comprehension. *Psychophysiology*, 40, 235-248.
van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory and Cognition, 31,* 443-467.
van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999b). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11*, 657-671.

van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17*, 701-718.
van den Brink, D., Brown, C. M., & Hagoort, P. (2001). Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. *Journal of Cognitive Neuroscience, 13,* 967-985.
van den Brink, D., & Hagoort, P. (2004). The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. *Journal of Cognitive Neuroscience, 16,* 1068-1084.
van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22*, 241-255.
Van Petten, C. (1989). *Context effects in word recognition: Studies using event-related brain potentials*. Doctoral dissertation, University of California, San Diego.
Van Petten, C. (1993). A comparison of lexical and sentence-level context effects and their temporal parameters. *Language and Cognitive Processes, 8,* 485-532.
Van Petten, C. (1995). Words and sentences: Event-related brain potential measures. *Psychophysiology, 32*, 511-525.
Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Timecourse of word identification and semantic integration in spoken language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25,* 394-417.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory and Cognition, 18*, 380-393.
Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open and closed class words. *Memory and Cognition, 19*, 95-112.
Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & McIsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3,* 131-150.
Van Petten, C., & Luka, B. (in press). Neural localization of semantic context effects in electrophysiological and hemodynamic studies. *Brain and Language.*
Van Petten, C., & Rheinfelder, H. (1995). Conceptual relationships between spoken words and environmental sounds: Event-related brain potential measures. *Neuropsychologia, 33,* 485-508.
Van Petten, C., Weckerly, J., McIsaac, H. K., & Kutas, M. (1997). Working memory capacity dissociates lexical and sentential context effects. *Psychological Science, 8*, 238-242.
Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38*, 41-63.

Wassenaar M, Brown CM and Hagoort P (2004). ERP effects of subject-verb agreement violations in patients with Broca's aphasia. *Journal of Cognitive Neuroscience*, 16, 553-576.

Waters, G.S., & Caplan, D. (1996) The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review, 103*, 761-772.

West, W.C., & Holcomb, P.J. (2000). Imaginal, semantic, and surface-level processing of concrete and abstract words: An electriophysiological investigation. *Journal of Cognitive Neuroscience, 12,* 1024-1037.

Wicha, N. Y., Bates, E. A., Moreno, E. M.,& Kutas, M. (2003). Potato not pope: Human brain potentials to gender expectation and agreement in Spanish spoken sentences. *Neuroscience Letters, 346*, 165-168.

Wicha, N. Y., Moreno, E.M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex, 39*, 483-508.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience, 16*, 1272-1288.

Ziegler, J. C., Besson, M., Jacobs, A. M., & Carr, T. H. (1997). Word, pseudoword, and nonword processing: A multitask comparison using Event-Related Brain Potentials. *Journal of Cognitive Neuroscience, 9*, 758-775.

FIGURE CAPTIONS

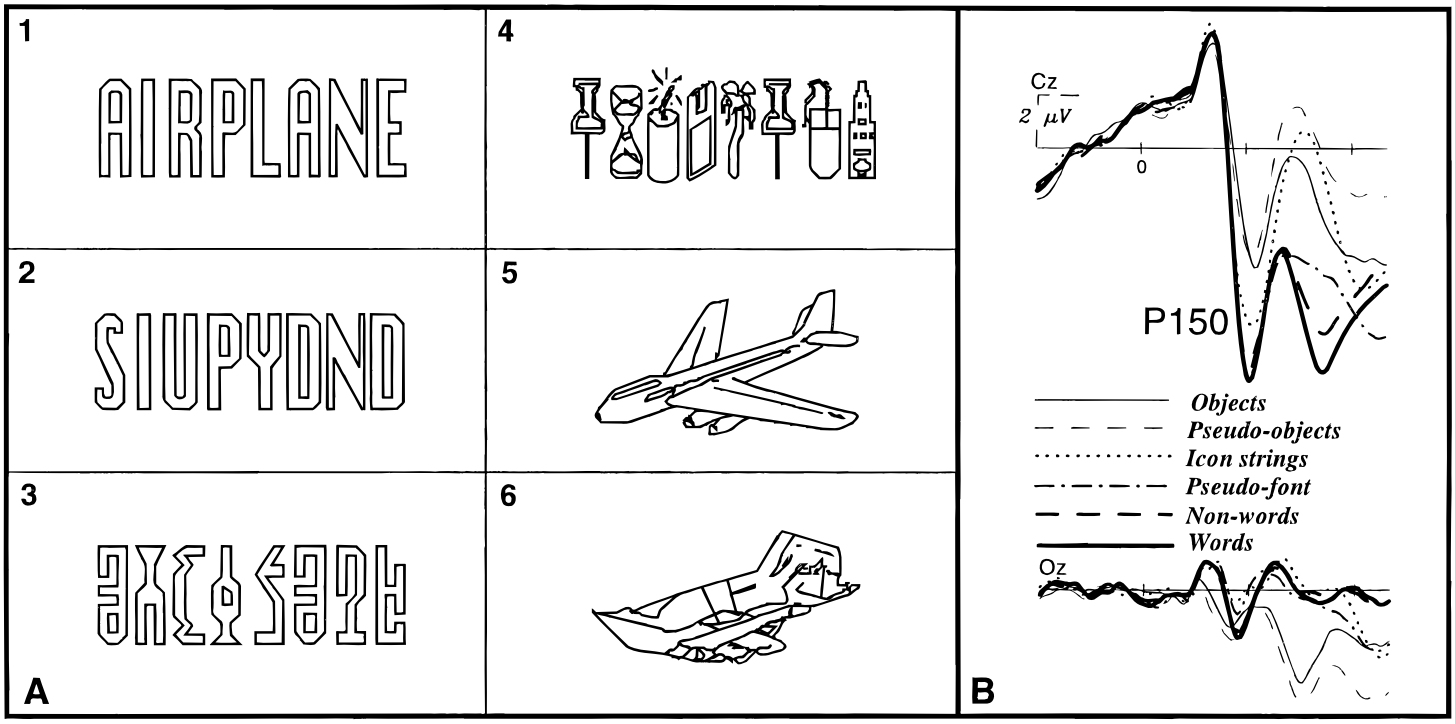
Figure 1. A) Examples of the six different image types presented to subjects in a random sequence: (1) words, (2) nonwords, (3) pseudofont strings, (4) icon strings, (5) real objects, and (6) pseudo-objects. B) First 350 ms of the grand average ERP to these stimuli from midline central (Cz) and occipital (Oz) sites. Note that P150 amplitude differentiates word-like (or letter stringlike) images from nonlinguistic object-like stimuli. Reprinted with permission from Schendan et al. 1998.

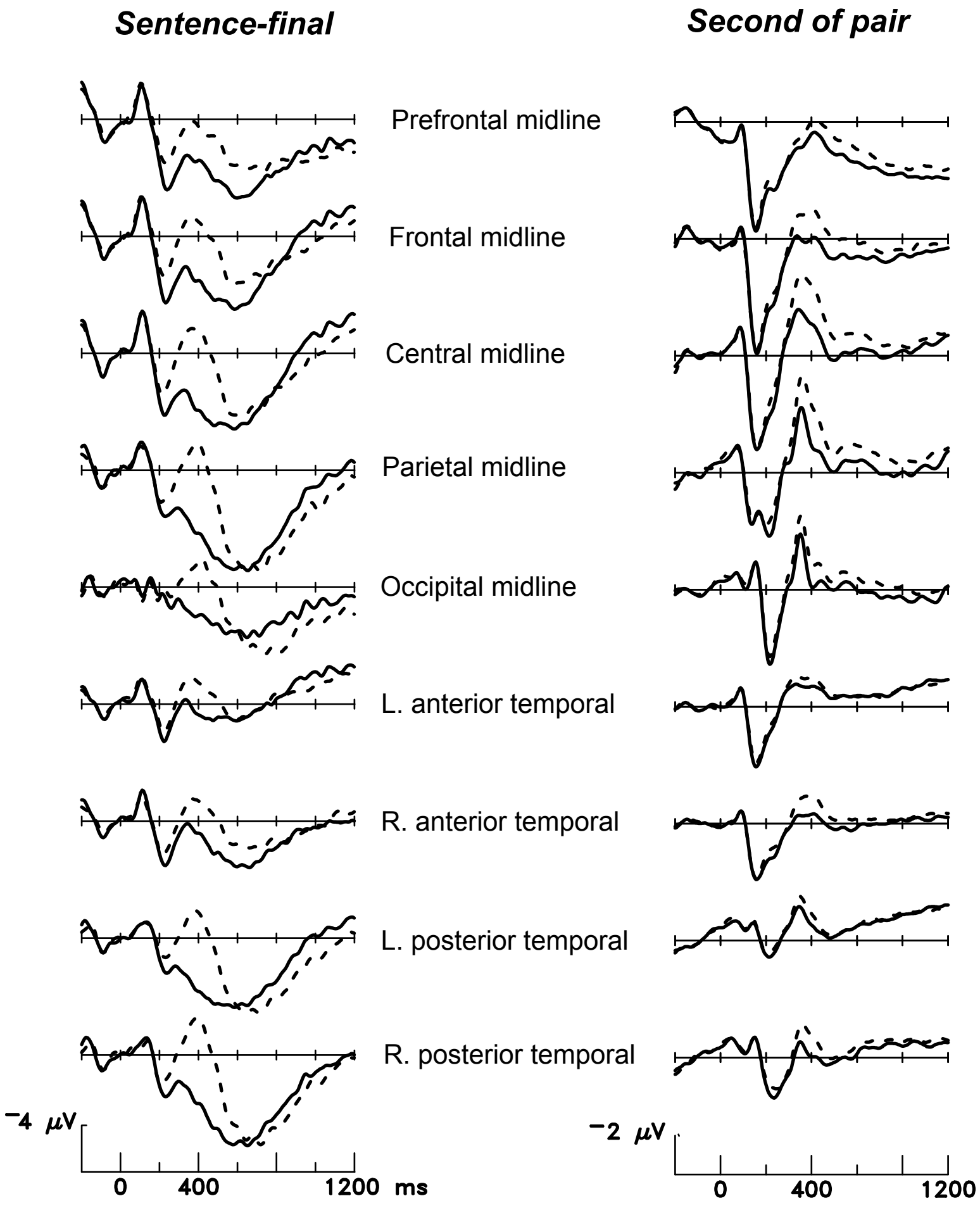
Figure 2. Left: grand average visual ERPs (N=16) elicited by high-cloze congruent and incongruent word completions of sentences read one word at a time, followed by a delayed "sensible and grammatical" judgment. judged Unpublished data from Groppe and Kutas. Right column: grand average ERPs (N=30) to the second words of semantically related and unrelated visual word pairs. Participants performed a delayed letter search task (decide if a letter presented after the second word was present in either word of the pair). Unpublished data from Luka and Van Petten. Note that for both data sets decision-related potentials were postponed beyond the epochs shown here. In contrast to the predictable sentence completions, even related words elicit substantial N400 activity. Also note the right-greater-than-left asymmetry of both N400 context effects.

Figure 3. Grand average (N=21), midline parietal site. Spoken sentence were completed with congruous words, words sharing onset phonemes with the congruous word, or words sharing final phonemes (rhyming) with the congruous words. ERPs to the rhyming completions were indistinguishable from ERPs to incongruous completions that were phonemically dissimilar to the congruous completion. Left panel: conventional average in which time 0 corresponds to the onset of critical words. Note the auditory N1 component peaking ~100 after acoustic onset. N400 onset and peak are delayed when the early portion of the presented word sounds like a congruous completion. Right panel: same data averaged with respect to the word isolation point, established in a gating experiment. N1 is not visible here because it is elicited by acoustic onsets, and the isolation points for the words occur at variable times post-onset; the averaging process leads to a "smearing" of the N1 across several hundred milliseconds, so that no peak is visible. Left panel shows that the N400 sentence congruity effect begins before the presented word can be uniquely identified (at the isolation point). Data from Van Petten, Coulson, Rubin, Plante, & Parks, 1999.

Figure 4. Representative data to grammatical violations (incorrect pronoun case markings, verb number mismatches) completing written sentences. Far left: grand average ERPs (N=16) to final words of grammatical and ungrammatical sentences presented one word at a time (2 wds/sec), intermingled with semantically congruent and anomalous sentences (see Figure 3). Middle:

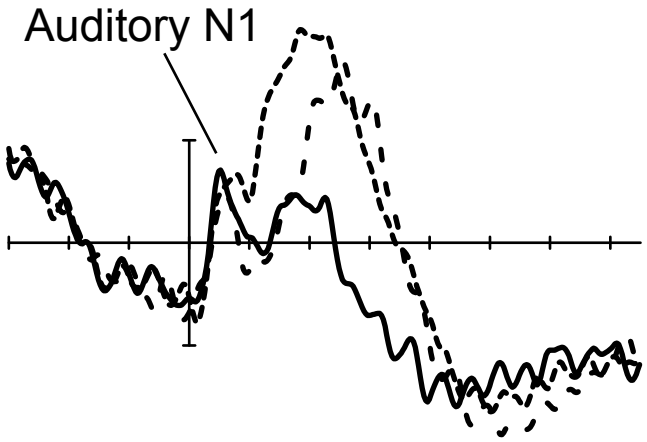
difference ERP created by point by point subtraction of ungrammatical minus grammatical ERPs; LAN effect is evident over the left frontal site and the P600 over a midline parietal site. Far right: topographical maps of the spatial distribution of the mean amplitudes (shaded area) for the LAN and P600 effects. Unpublished data from Groppe and Kutas.


1
AIRPLANE
2
SIUPYDND
3
4
5
6
A
Cz
2 μV
0
P150
Objects
Pseudo-objects
Icon strings
Pseudo-font
Non-words
Words
Oz
B

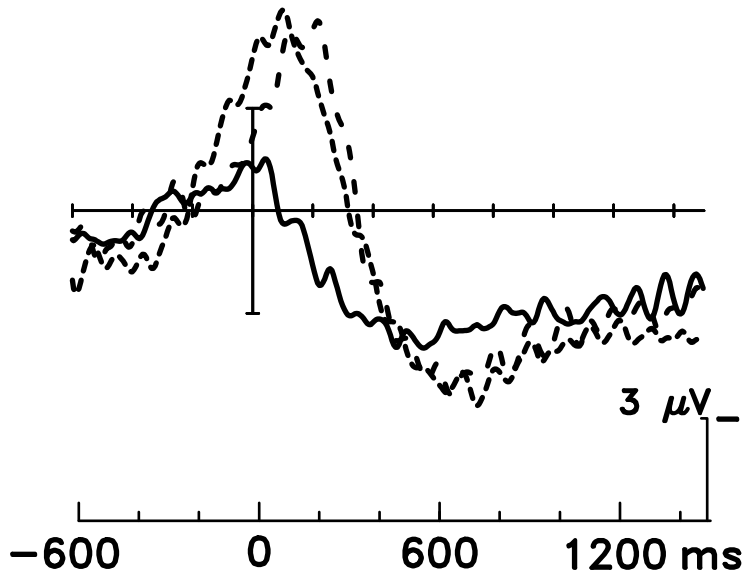

Sentence-final
Second of pair
Prefrontal midline
Frontal midline
Central midline
Parietal midline
Occipital midline
L. anterior temporal
R. anterior temporal
L. posterior temporal
R. posterior temporal
−4 μV
0
400
1200 ms
−2 μV
0
400
1200

The main highway was flooded so they had to take a long ...

Timelocked to word onset



Timelocked to isolation point

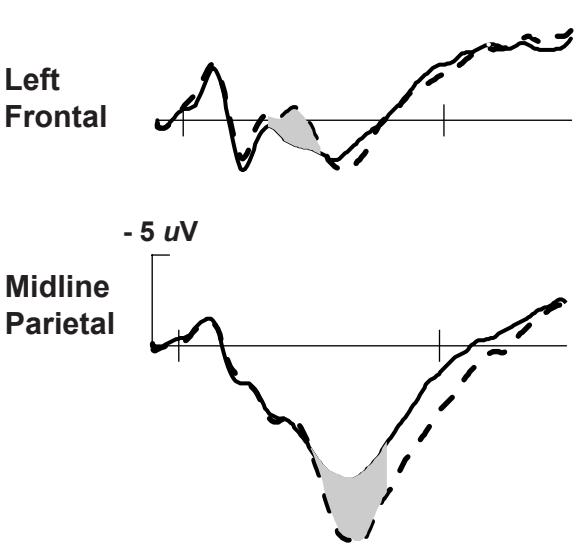


------- ***Incongruous, rhyme*** **(CONTOUR)**

- - - - ***Incongruous, onset similar*** **(DETAIL)**

——— ***Congruous*** **(DETOUR)**

"My niece wove a pot holder for..."



Grammatical Ending (*me*)
Ungrammatical Ending (*I*)

GRAMMATICALITY EFFECTS