Event-Related Potentials in Cognitive Science

Marta Kutas and Steven A. Hillyard

Introduction

How is it that the physicochemical processes in the brain give rise to the phenomena of the mind? This transcendent question can be phrased in the jargon of several disciplines and addressed at many levels, with the choice often dictated by the experimental tools available. Investigators in various arms of the behavioral and neural sciences have for the most part forged ahead independently and found, at best, partial solutions to the problem. It is becoming apparent that the enormity of the endeavor necessitates a liaison—if not a marriage—between disciplines.

Current methods for probing the neural basis of mind all suffer serious limitations, whether in applicability or in logical inference. For example, it is seldom possible to record the activity of single neurons in humans or to use the other intrusive approaches that are available in animal investigations. Studies of natural lesions in patients are rife with interpretive difficulties; simply knowing that damage to a particular cortical area is associated with the debilitation of a particular linguistic function does not necessarily allow one to infer that the damaged area is the “seat” of that function. Thus, in most cases, we must seek out consistencies among the data bases of different experimental approaches, being swayed by the weight of converging evidence to draw conclusions about which parts of the brain do what, how, and when.

Interdisciplinary studies of mind–brain relationships are particularly needed at the interface between the neurophysiology of the cerebral cortex and the psychology of cognitive processes. At present, however, there are few techniques for studying this critical nexus in the human being. The recording of event-related potentials (ERPs) from the intact scalp provides one of the few available windows on the dynamic patterns of information transactions in the brain that cooccur with specific cognitive acts. ERPs are transient voltage fluctuations generated in the brain in conjunction with sensory, motor, or cognitive events. These phasic brain potentials, which reflect the activation pattern of large numbers of neurons in the brain, are embedded in the ongoing electroencephalogram and can be extracted by signal-averaging techniques. A basic assumption underlying the ERP approach is that some part of the neural activity

---

Marta Kutas and Steven A. Hillyard • Department of Neurosciences, School of Medicine, University of California, San Diego, La Jolla, California 92093. Some of the work cited herein was supported by grants from NSF BNS80-05525 and Sloan Foundation B1980-35. M. Kutas is supported by Research Scientist Development Award USPS 1K02MH00322/02.
RESULTS

ingendering the potentials on the scalp has a functional role in some aspect of cognitive activity. From this assumption, it follows that variations in the spatiotemporal properties of ERPs permit inferences to be made about the brain regions subserving particular cognitive functions and the temporal sequencing of the subcomponents of those functions.

While ERP measures provide the luxury of information in both space and time, the technique has serious limitations in each domain. ERP components are generally considered the temporal sequence of peaks and troughs in the ERP waveform (Figure 1). However, because of temporal overlap of the components, a visible peak or trough may represent the summation of activity from different brain sources. Furthermore, the latency of a component often changes with manipulations of stimulus parameters and response requirements. In the spatial domain, the problem of localizing the source generators of an ERP is formidable. Little is known about their physiological bases, and unequivocal inferences about source localization cannot be drawn from an examination of the distribution of the scalp potentials. Finally, we must acknowledge that the ERP reflects primarily synchronous neuronal activity (and not even an exhaustive sample of that), and hence, many information transactions may occur in the brain without an ERP signature at the scalp.

The ERP technique by no means provides any easy solutions to the mind–body problem. Nonetheless, if we are properly sensitive to its limitations and reserved in our inferences, we can gain valuable information using this technique. At present, the use of the ERP approach is primarily as a bootstrap operation, whereby relationships between ERP parameters and cognitive processes are validated so that predictions about ERP variations converge with data from other methods. The ERP method, like

AVERAGED EVENT-RELATED POTENTIALS

Figure 1. Idealized waveform of the auditory ERP—including the brain-stem responses (Waves I–VI); the mid-latency components (Waves N0, P1, N0, P1, N1); the long-latency “vertex potential” (Waves P1, N1, P2, N2); and task-related endogenous components (N1, P2). The anticipated auditory signal is preceded by a slow negative shift (e.g., CNV). Note the logarithmic time base.
most experimental techniques in psychology, is complementary and correlational, but it is capable of providing a physiological litmus test for specific aspects of information-processing theories. While correlations are always subject to different shades of interpretation, we should be open-minded enough to accept their contributions to the weight of converging evidence. In fact, the ERP is becoming a major source of information about the neural bases of cognition by virtue of its strong correlational links with a wide variety of processes, including psychophysical judgment, perception, recognition, selective attention, decision making, orienting reactions, and certain language functions.¹

An exhaustive review of the cognitive ERP literature would be redundant at this point, as many sources exist (Begleiter, 1979; Callaway, Tueting, & Koslow, 1978; Desmedt, 1977a, 1979, 1981; Lehmann & Callaway, 1979; Otto, 1978). We prefer, rather, to give several examples that underscore the utility of the ERP technique in cognitive neuroscience. The general approach implements Mountcastle’s (1976) “combined experiment,” wherein “it is now possible to combine in one experiment the methods and concepts of [psychology and neurophysiology] to yield a deeper insight into the brain mechanisms that govern behavior than is possible with either . . . alone.” The specific questions dealt with in the following sections include (1) how to characterize the specialized processing resources proposed by multiple-capacity models of attention; (2) how to measure the timing of the intermediate stages of information processing; (3) how various linguistic and semantic dimensions are organized in the brain; and (4) how to define the specialized functions of the two cerebral hemispheres in humans.

Allocation of Processing Resources

The human brain is a limited-capacity processor. This currently fashionable aphorism refers to the fact that we are limited (by more than our anatomy) in the number of things that we can do simultaneously, and when increasing demands are placed on us, some aspects of our performance deteriorate. It is nonetheless remarkable that in some cases, we can carry out several concurrent tasks without noticeable hesitation. Kahneman’s (1973) original model provided us with a finite, undifferentiated pool of the processing “capacity” or “resources” that could be tapped to perform all types of tasks, with the more “difficult” tasks drawing more resources. Intuitive as it may seem, however, the concept of task difficulty has not been easy to define or quantify.

A commonly used procedure for measuring the difficulty of a task has been to assign a concurrent secondary task and to quantify its impact on the primary task as it becomes progressively more demanding (Kahneman, 1973; Kerr, 1973; Rolfe, 1971). The results of such time-shared (dual-task) experiments have yielded little support for a general capacity model. When task difficulty is varied, some pairs of tasks show

¹ These correlations have been observed for a multiplicity of ERP components including N₁, N₂, N₂, N₄, P₂, P₃₄, P₉₀, Pₓ, CNV, PINV, RP, MP, etc. (For further details see Donchin, Ritter, & McCallum, 1978; Picton & Stuss, 1980).
performance interactions while others do not (see Navon & Gopher, 1979; Sanders, 1979; Wickens, 1980). From such findings have emerged the multiple-capacity models, which include a number of qualitatively distinct pools of processing "resources." The theoretical emphasis has shifted toward the functional characterization of these specialized processing resources and the scaling of their allocations in terms of task difficulty.

This need for converging measures of task difficulty or "work load" in the presence of different putative categories of resources led to dual-task experiments that included the recording of late endogenous ERPs to a secondary (probe) task while the difficulty of a primary sensory and/or motor task was varied. This approach is based on the premise that increasing the utilization of resources by the primary task will take its toll on those ERP components that draw on the same resource pool, thereby revealing the locus of interaction with minimal response interference. The viability of this procedure depends on the identification of a component of the ERP as a relatively pure measure of the availability of a single putative resource. Such a claim has in fact been proffered for the P300 component of the ERP.

The label P3 or P300 has been used to refer to a positive component of the ERP that has a latency anywhere from 300 to 900 msec poststimulus (see Figure 1). There is, however, considerable controversy over the possible presence of multiple positive components within this latency range, each of which may reflect different underlying cognitive processes (Friedman, Vaughan, & Erlenmeyer-Kimling, 1978; Picton & Stuss, 1980; Renault & Leseure, 1978; Roth, 1978). Nonetheless, there is ample evidence (reviewed in Donchin, Ritter, & McCallum, 1978) indicating that there is a modality-independent, late positive component over the centroparietal scalp, elicited by task-relevant, "surprising" stimuli, which can be used as a measure of stimulus evaluation time (e.g., encoding, identification, and categorization) free from response-related factors (e.g., planning, selection, control, and execution). This particular ERP, designated here as the P300, should thus be sensitive to task demands that engage stimulus evaluation processes rather than increased response complexity.

Wickens, Isreal and Donchin (1977) employed a secondary task (counting occasional frequency shifts in a repetitive sequence of tones) to identify the resource demands imposed by manipulating the difficulty of a manual control (joystick tracking) task. They found that the amplitude of the P300 component elicited by the "probe" tones during the tracking task was reduced relative to that obtained in no-tracking control conditions; however, increases in tracking difficulty brought about by raising either the order of the control dynamics or the bandwidth of the signal to be tracked were not reflected in a concomitant attenuation of the P300 amplitude. Isreal, Chesney, Wickens, and Donchin (1980) replicated these findings and noted further that this apparent insensitivity of the P300 to increases in the bandwidth of the tracking task occurred in the face of prolonged reaction times (RTs) to the probe stimuli. These results were consistent with the view that these types of difficulty manipulations affected primarily the response rather than the perceptual aspects of the task.

Other studies of this kind have shown that manipulations of the difficulty of perceptual tasks (e.g., by varying the number of items monitored on a visual display)
were indeed reflected in attenuations of the concurrently recorded P300s to auditory probes (Isreal, Wickens, Chesney, & Donchin, 1980). This interaction of the P300 with perceptual load is consistent with the hypothesized contingency between P300 and processes of stimulus identification and classification (Kutas, McCarthy, & Donchin, 1977; Duncan-Johnson, 1981; McCarthy & Donchin, 1981) and provides converging evidence supporting multiple-resources models of information processing. The most recent ERP studies of resource allocation have varied several dimensions of the tracking task whose loci of interaction were not known a priori but could be inferred on the basis of variations in P300 amplitude (Wickens, Derrick, Gill, & Donchin, 1983).

Further support for the proposition that the amplitude of the P300 mirrors the allocation of perceptual/cognitive resources has come from a study by Kramer, Wickens, Vanasse, Heffley, and Donchin (1981). They found a reciprocal relationship between the amplitudes of the P300s elicited by the primary and the secondary task stimuli as a function of increasing task “difficulty”; that is, while the amplitude of the P300 to secondary probes generally decreased with the order of the tracking dynamics, the P300 to the primary task stimuli showed a concomitant increase (see Figure 2). In this way, studies of the P300 and resource allocation have gone beyond the bootstrap-

![Figure 2](image)

**Figure 2.** Average parietal (Pz) ERPs elicited by the auditory probes (secondary task) and the step changes in the visual target of the primary task, each of which was counted on different experimental runs. The solid-line ERPs were collected during runs in which the subject counted either the auditory signals or the step changes in the visual target without simultaneous manual tracking. The remaining ERPs were obtained during experimental runs in which the subjects performed two tasks concurrently; these runs differed in the associated difficulty of the tracking task, being easiest for first-order regular and hardest for second-order random. (Data presented in Kramer et al., 1981.)
ping phase, providing considerable validation of this physiological measure and convergence with behavioral data.

**Chronometric Analysis of Decision Processes**

The P300 component has been recorded in a wide variety of experimental situations, and numerous attempts have been made to characterize the cognitive processes with which it is affiliated. Candidate hypotheses for the psychological correlates of the P300 include orienting, resolution of uncertainty, delivery of task-relevant information, decision making, context updating, and postdecisional closure of cognitive activity, among others (Desmedt & Debecker, 1979; Donchin, 1979; Friedman, 1978; Tueting, 1978). While this heterogeneity of psychological concepts may seem disconcerting, it is well established that the P300 is an exquisitely sensitive index of the brain’s response to “surprising” stimuli. Thus, for example, Sutton (Sutton, Braren, John, & Zubin, 1965) discovered that when subjects were asked to predict which of a set of stimuli would occur next, the P300 was elicited as an inverse function of stimulus probability: the less likely the event, the larger the associated P300. Since then, Duncan-Johnson and Donchin (1977) have demonstrated that the amplitude of the P300 varied monotonically with the probability of task-relevant stimuli, over a wide operating range. An elegant analysis by Squires, Wickens, Squires, and Donchin (1976) further revealed that P300 amplitude was influenced not only by the overall “global” probability of a particular signal but also by the fine structure of the preceding stimulus sequence. In fact, the P300 has proved sensitive to a variety of probability manipulations, including not only simple event probability, but also sequential, temporal, local, outcome (i.e., combined probability of a particular response and a subsequent stimulus), and contingent (i.e., probability of a stimulus given a particular response) probability (for reviews see Donchin, 1979; Hillyard, Squires, & Squires, 1983; Picton & Stuss, 1980; Pritchard, 1981).

The consensus from such studies is that variations in P300 amplitude are determined primarily by the subject’s “expectancy” for a task-relevant event, which may or may not correspond to its objective probability. In many cases, the task-relevant events are members of a category about which the subject has been asked to make a decision or a covert response (Courchesne, Hillyard, & Courchesne, 1977; Johnson & Donchin, 1980). The defining characteristics of category membership may be based on simple physical cues, such as the frequency of a tone or the hue of a light flash, but similar results ensue for more complex classifications based on semantic rules. Thus, for example, large P300s are elicited by the infrequent occurrences of the words brake, steak, lake, or fake when the subject is required to count all the words that rhyme with cake, or by urge, shove, goad, or nudge when the synonyms of prod are being counted (Kutas & Donchin, 1978).

The dependence of P300 on the “surprise” value of a stimulus implies that the stimulus must be evaluated (encoded, identified, and categorized) before the P300 can be elicited. Donchin and his colleagues have adopted this line of reasoning to propose that while the P300 is probably not a direct sign of the stimulus classification process
per se, its appearance is contingent on the completion of such processes. As a corollary, they further proposed that P300 latency be used as a measure of the time that it takes a subject to categorize events, that is, as a measure of the duration of stimulus evaluation processes.

This interpretation implies that the designation P300 is a misnomer since the latency of this component can vary considerably as a function of the complexity of the categorization rule for the surprising stimulus. A case in point is the experiment by Kutas and Donchin (1978), in which the subjects were presented with random sequences of words under instructions to press a key for one of the word classes. In all cases, the stimulus class occurring with a low frequency (20%) elicited significantly larger P300s than did the remaining stimuli (80%). Moreover, as shown in Figure 3, during the count-only and the RT condition emphasizing accuracy, the latency of the P300 increased as a function of the complexity of the classification rule; it was shortest when the rare stimulus was a single male name, intermediate when the infrequent

---

![Figure 3](image-url)  
*Figure 3. Superimposition of central (Cz) ERP's averaged across five subjects for three semantic categorization tasks of increasing complexity. Three different response regimes were used: count-only (no motor response), RT task emphasizing accuracy, and RT task emphasizing speed. (Data from Kutas & Donchin, 1978.)*
class consisted of a number of different male names, and longest when categorization required determining whether a word was a synonym of the word *prod*.

A single trial analysis of these data not only corroborated the reliability of the relationship between the duration of stimulus categorization and P300 latency, but also suggested a possible explanation for the reported variability in P300-RT correlations in such experiments (Kutas et al., 1977). Because RT is the most frequently used measure of information-processing time, the occasional reports of weak P300 latency-RT correlations have been troublesome for P300 advocates. However, Kutas et al. (1977) noted that the P300 latency-RT correlation was altered dramatically by the subject's response strategy. Under instructions to make "speeded" responses, P3-RT correlations were relatively low, while under "accuracy" instructions such correlations were significantly higher. Hence, Kutas et al. concluded that P300 latency-RT correlations were high only when response execution was contingent on the completion of the stimulus identification and classification, again supporting the general proposition that P300 latency, unlike reaction time, could be used to index the duration of stimulus evaluation processes independent of the time invested in response-related processes.

Systematic delays in both RT and P300 latency have been found as a function of the discriminability of the stimuli in an "oddball" detection task (Ford, Roth, & Kopell, 1976; Ritter, Simson, & Vaughan, 1972; Squires, Donchin, Squires, & Grossberg, 1977); the changes in these two measures were not equivalent, however, and the increase in RT was typically larger than the increase in P300 latency. Findings from several variants of the memory-matching paradigm of Sternberg also indicated nonequivalent changes in P300 latency and RT as a function of decision latency (Adam & Collins, 1978; Ford, Roth, Mohs, Hopkins, & Kopell, 1979; Gomer, Spicuzza, & O'Donnell, 1976; Roth, Kopell, Tinklenberg, Darley, Sikora, & Vesecky, 1975). While both P300 latency and RT increased linearly with greater short-term memory load, the slope of this function was steeper for RT than for P300. These results further underscored the dissociation of P300 from response-related processes.

McCarthy and Donchin (1981) directly tested the hypothesis that the slowing of response selection and execution would have no effect on P300 latency, while delays of stimulus categorization would affect both P300 latency and RT. The stimulus evaluation time was varied by embedding target stimuli in a more-or-less discriminable background. Response selection, on the other hand, was manipulated by changing the compatibility between the target stimulus and the response required. The results showed that the P300 to the target words *right* or *left* occurred earlier when the word was embedded in a uniform matrix of *#* signs than when the background consisted of random alphabetical characters. Whether the word *left* required a compatible left-hand response or an incompatible right-hand response, however, did not alter P300 latency. Since both visual "masking" and response incompatibility increased RT to the target words, while the P300 was sensitive only to the perceptual variable, the authors concluded that "P300 can serve as a dependent variable for studies that require, for their effective solution, a measure of mental timing uncontaminated by response selection and execution processes."

Duncan-Johnson and Donchin (1980) capitalized on the established relationship between P300 latency and stimulus evaluation time to tease apart the relative contri-
butions of perceptual facilitation and response priming to the observed reduction in RT to expected events (Hinrichs & Kraiz, 1970; Miller & Anbar, 1981). The data revealed a decrease in both RT and P300 latency as probability increased, with the rate being twice as fast for RT. This result suggests that probable events are indeed identified faster and, in addition, that responses to them are emitted more rapidly. The authors interpreted this pattern of P300–RT covariation as indicating that the major effect of probability on choice RT occurs about equally in stages of stimulus evaluation and response generation.

A striking example of convergence between ERP evidence and psychological theory has been the use of P300 to investigate the locus of interference in the Stroop color-word test (Duncan-Johnson & Kopell, 1981; Warren & Marsh, 1979). Explanations of the Stroop effect (naming the ink color of a printed word is delayed if the word spells a conflicting color name; Stroop, 1935) have been phrased in terms of either perceptual interference (Hock & Egeth, 1970; Seymour, 1977) or response conflict, (Dalrymple-Alford, 1972; Hintzman, Carre, Eskridge, Owens, Shaff, & Sparks, 1972; Keele, 1972; Klein, 1964) and numerous behavioral investigations on this point have not settled the question. Converging measures of RT and P300 latency (Duncan-Johnson & Kopell, 1981), however, have indicated that response competition is the primary source of the Stroop interference effect. This conclusion was based on the observation that the latency of the P300 remained invariant while RTs showed the usual interference effect. In association with the stable P300 latencies in the condition where ink color was the relevant cue, Duncan-Johnson and Kopell found not only costs in the RT to the incongruent stimuli but also benefits to the congruent stimuli in relation to those for neutral words. Moreover, they obtained an increase in P300 latency when the Stroop test used hues more difficult to discriminate (e.g., reddish purple versus bluish purple for the words red and blue). In this way, the measurement of ERP latencies aided in the selection among alternative theories of information processing, and the relative contributions of stimulus and response processes to the timing of the specific act in question could be assessed.

**Language**

**ERPs and Linguistic Categories**

While many ERP studies have adopted language paradigms as a convenient experimental vehicle for evaluating ERP signs of general cognitive functions, several direct methods have been made to relate ERPs to specific linguistic dimensions—from the elementary phonemic and syllabic building blocks to meanings and syntax. Molfee (1978a, b, 1980) has concluded that the ERP can be a sensitive index of changes in the acoustic parameters that are an integral part of speech perception. Based on multivariate techniques of waveform analysis, he has reported that one component of the systematic variability in the ERP reflects the presence or absence of formant transitions, while another represents variations in formant bandwidth. Also working at an elementary level, Wood and colleagues found that the ERPs to CV syllables differed depending on whether they were processed for phonetic or acoustic (e.g., fundamental
frequency) cues; this difference was evident only in the left scalp recordings (Wood, Goff, & Day, 1971; Wood, 1975, 1977).

At the word level, Chapman has carried out an extensive inquiry into the ERP signs of connotative semantic meaning. His approach has been to average ERPs to words belonging to six different semantic categories, consisting of the high and low values along Osgood's three orthogonal dimensions: evaluative, potency, and activity (Chapman, Bragdon, Chapman, & McCrary, 1977; Chapman, McCrary, Chapman, & Bragdon, 1978; Chapman, McCrary, Chapman, & Martin, 1980). While the ERPs to each of the six semantic word categories were found to be quite similar, differences were revealed if the ERP to each semantic category was compared against that of the grand waveform (averaged across all six categories). Chapman concluded that semantic content is manifested in slight amplitude and latency deviations from the average ERP. Significantly, these ERP signs were invariant across subjects.

While Chapman has searched for semantic universals across many physically different words, Brown and his colleagues have hunted for the telltale ERP signs of meaning by using physically identical pairs of words (homophones), which can take on different meanings depending on context. Across several languages and auditory and visual modalities, and using a host of analytic techniques, Brown has obtained consistent ERP differences, particularly over the left anterior scalp regions, in association with the noun and verb forms of a number of different homophones (Brown & Lehmann, 1979; Brown, Lehmann, & Marsh, 1980; Brown, Marsh, & Smith, 1973, 1976, 1979; Marsh & Brown, 1977). At present, it is unknown whether these ERP differences are specific to the differential semantic meanings induced or to the syntactic shift between the noun and verb categories of grammar.

Another important aspect of language amenable to the ERP approach is the role of context in establishing expectancies. Such expectancies influence our recognition and memory of letters, words, sentences, and prose passages (see Blank & Foss, 1978). We have investigated the role of contextual factors in language by comparing the ERPs elicited by confirmations and violations of linguistic expectancies during visual sentence processing (Kutas & Hillyard, 1980a,b,c, 1981). In these studies, the subjects were asked to read sentences, presented one word at a time, in order to answer subsequent questions about their contents. While the majority were simple, meaningful sentences, some percentage (25–50%) of them were completed by semantically inappropriate words that rendered the sentence nonsensical. A comparison of the ERPs recorded in association with these two types of endings indicated that the response to semantically anomalous words was characterized by a negative component between 300 and 600 msec poststimulus (N400), which was not evident in response to the semantically appropriate endings.

Further studies have demonstrated that the N400 did not occur in response to all types of “surprising” words in this linguistic context. Thus, when semantically appropriate words were unpredictably presented in these sentences in oversize, boldface print, these “surprises” were manifest in the enhancement of a late positive complex of waves (Kutas & Hillyard, 1980b). Moreover, these two distinct ERP effects (N400 and late positivity) seemed to be elicited concurrently by the same word when it was both semantically inappropriate and surprisingly large (Kutas & Hillyard, 1980a).
Figure 4. Grand average ERPs (over six subjects) elicited by semantically and grammatically deviant words presented in prose passages. Deviant words are superimposed on ERPs elicited by control words that occurred in comparable positions in other sentences of the passages (see footnote 4).

The majority of these experiments employed a series of unrelated seven-word sentences with the deviations restricted to occasional terminal positions. However, we have also examined the ERP signs of various linguistic deviations in variable-length sentences in prose passages. Scattered throughout these texts, read by the subjects as a source of answers to subsequent multiple-choice questions, were a number of semantic and grammatical anomalies. As can be seen in Figure 4, very similar brain responses were elicited by semantically inappropriate words, whether they occurred at the ends or in the interiors of sentences; in both cases, the N400 has a broad central-parieto-occipital distribution. It is also clear that grammatical aberrations (e.g., the unexpected occurrence of an incorrect noun–verb number correspondence or an incorrect verb tense) were not associated with the same ERP configuration as were the semantic anomalies. Thus, these data are consistent with the view that N400 elicitation within a linguistic context appears to be contingent on some aspect of semantic processing.

Function and Content Words

Linguists and psycholinguists have long realized that words not only contribute their individual meanings but also provide structure to the language. To reflect the fact that some words assume more of a structural role than others, a fundamental

---

2 These data are a subset of the ERPs collected in a prose reading task (Kutas & Hillyard, in press).
division of the lexicon into two vocabularies has been proposed: the content words, or contentives, generally bear reference and carry the major semantic information load, while the function words, or functors, generally provide syntactic structure in language and signal the relations between content words. Content words—consisting of the major lexical categories of noun, main verb, adjective, and most adverbs—are members of the “open-class” category, so called by virtue of its openness to new members as warranted by language usage. On the other hand, function words—consisting of the minor lexical categories of determiners, auxiliary verbs, prepositions, conjunctions, articles, and pronouns, as well as bound morphemes that bear grammatical reference—form the small, relatively fixed set of the “closed-class” word-category. The open class consists of a large number of elements that vary greatly in length and frequency of occurrence, and that generally lack a one-to-one mapping upon a particular grammatical category. By contrast, the members of the closed class are typically short and limited in number, and they occur with relatively high frequency in a constant grammatical class. The two classes also differ in their predictability in a Cloze procedure (Aborn, Rubenstein, & Sterling, 1959) and in the type of production errors typically associated with each (Garrett, 1975, 1976).

Recently, it has been suggested that this division in the vocabulary derived from a linguistic analysis has a cognitive counterpart vis-à-vis information-processing theories of language (Bradley, 1978; Bradley, Garrett, & Zurif, 1980; Zurif, 1980). In particular, Bradley found that RTs for deciding whether a given array of letters is a real word varied systematically as a function of its frequency of occurrence in the language for the closed-class items but not for the open-class items. In addition, she observed that subjects took more time to classify a nonword that began with the letters of a real word than to classify one beginning with a nonsense syllable, but only if that initial real word was a member of the open class. Bradley and associates have suggested that these results imply different neural organizations for contacting and retrieving open- and closed-class words.

Neuropsychological support for this view has come from the administration of these paradigms to aphasic patients (Bradley, 1978; Bradley et al., 1980; Friederici & Schoenle, 1980). Both Bradley and Zurif have reported that, unlike neurologically intact subjects, Broca’s aphasics processed these two vocabulary classes equivalently. Zurif (1980) emphasized that failure to distinguish between open- and closed-class words seemed to be specific to aphasics with damage to the anterior regions of the left hemisphere. Patients with posterior lesions in the left hemisphere (Wernicke’s aphasics) showed a normal differentiation between these two vocabulary classes. Data of this sort reinforce the proposition that open- and closed-class words are processed differently in the brain, and they suggest the existence of a “natural neurological separation between the function of processing sentence form and that processing semantic representations” (Zurif, 1980, p. 310).

Despite their interdisciplinary appeal, these data are not without controversy (Gordon & Caramazza, 1982), and their interpretations leave a number of questions unanswered. For example, the suggestion that the grammatical impairment exhibited by Broca’s aphasics relates to the manner in which the open- and closed-class voca-
bularies are used in sentence processing has been questioned on the grounds that major sentential stress falls more often on open-class than on closed-class words, thereby creating a potential confounding of stress and word class (Kean 1977; Swinney, Zurif, & Cutler, 1980). Thus, Kean (1977, 1980) has argued that the agrammatism of Broca’s aphasics is more aptly viewed as a phonological than as a grammatical deficit. Moreover, there exists the problem of classifying the brain’s response to words that could be members of either the open or the closed class in isolation (e.g., can, will). Are there multiple representations for such words? Thus, some controversy remains concerning whether language processing can indeed be characterized in terms of distinct systems of semantics and grammar that are subserved by functionally and/or neuroanatomically discrete regions of the brain.

Such uncertainties underscore the need for physiological data to converge with the evidence for the differential cerebral representation of mechanisms subserving referential (open-class) and syntactic (closed-class) systems. A straightforward approach to this question would be to compare the scalp-recorded ERPs elicited by words belonging to these two vocabulary classes. Such a comparison is provided in Figure 5. Inspections of these waveforms reveals that members of the two vocabulary classes elicited ERPs with distinctly different scalp distributions. Open-class (content) words were associated with a greater positive shift in the recording epoch from 200 to 700 msec poststimulus than were the closed-class (function) words; this difference was most pronounced over the frontal regions of the scalp.

The ERPs to these two lexical categories also differed in their lateral distribution. While ERPs elicited by closed-class words appeared to be symmetrical over the two cerebral hemispheres, the ERPs to open-class words included a positivity in the 400–700 msec poststimulus epoch (shaded area in Figure 5) that was greater over the left than the right hemisphere, particularly over the temporoparietal regions.

In previous studies, the degree of hemispheric asymmetry of this late positivity had been found to be influenced by the subject’s family history of left-handedness, being greater for subjects without left-handers in their immediate family (Kutas & Hillyard, 1980a,b). A detailed examination of these waveforms indicated that this difference in the degree of asymmetry appeared to be due primarily to additional late positivity over the left hemisphere in response to open-class words. The ERPs to open- and closed-class words over the right temporoparietal area were not noticeably different. This pattern of ERP lateralization parallels Bradley’s finding that words from the two vocabulary classes were equally difficult to recognize when presented to the left visual field, while open-class words were identified more accurately than closed-class words in the right visual field. These different ERP patterns to open- and closed-class words provide neurophysiological evidence of the separate neural systems subserving the processing of these two vocabulary classes that were originally described by the linguists and are currently utilized by psycholinguists, cognitive psychologists, and aphasologists.

The generality of these findings has been tested with words presented in isolation to the right and left visual fields (Neville, personal communication, 1982). Preliminary data indicate that ERP differences between open- and closed-class words, in some
respects like those described above, were also obtained under these circumstances. This finding leaves the door open for ERP investigations of specific types of words within a vocabulary class, the effects of word position and stress, and other such factors. Moreover, ERP differences along these dimensions might be used to determine the influence of an item’s membership in a specific category versus its structural role within a sentence. Further confirming evidence would come from recordings of ERPs to physically (e.g., phonologically and orthographically) identical words that are grammatically ambiguous in isolation but that can assume either an open- or a closed-class function depending on sentential context.
Hemispheric Specialization

The functional asymmetry of the left and right cerebral hemispheres is a robust instance of the localization of brain function in humans that has formed the basis of many ERP investigations. The extensive literature relating ERPs to functional hemispheric specialization has been critically reviewed in some detail (see Desmedt, 1977b; Donchin, Kutas, & McCarthy, 1977; Friedman, Simson, Ritter, & Rapin, 1975a; Galambos, Benson, Smith, Schulman-Galambos, & Osier, 1975; Hillyard & Woods, 1979; Neville, 1980). To give a brief overview, a number of studies have reported amplitude asymmetries in the ERPs recorded over the left and right hemispheres during tasks requiring some linguistic or semantic analysis (Morrell & Salamy, 1971; Matsumiya, Tagliasco, Lombroso, & Goodglass, 1972; Preston, 1979; Thatcher, 1977). In general, the ERPs over the left hemisphere were more sensitive to the linguistic nature of the eliciting stimuli than were ERPs over the right hemisphere (Brown et al., 1973, 1976; Buchsbaum & Fedio, 1969, 1970; Wood et al., 1971; Wood, 1975). These reported asymmetries have been relatively small, however, and many of the pertinent studies could be criticized on methodological and/or statistical grounds. Moreover, a number of other studies have obtained a surprising absence of hemispheric asymmetry in the ERPs recorded in tasks requiring linguistic analysis (Friedman et al., 1975a,b; Galambos et al., 1975; Shelburne, 1972, 1973; Smith, Nielson, & Thistle, 1975). Given the sound psychological and neurological foundations of the functional distinction between the hemispheres, such negative ERP findings have left many eager investigators shaking their heads in dismay.

It might be the case, however, that these results are only superficially negative. Let us assume, for the sake of argument, that the ERP reflects a highly complex spatiotemporal pattern of brain activity that is an essentially valid reflection of important neuronal computations underlying cognitive acts. If Lenneberg (1973) was correct in proposing that "all the specialized activities in all the different parts of the nervous system can be viewed as a single configuration and the activity patterns of the brain can be seen as a series of moment to moment transitions from configuration to configuration," then the ERP may well be a valid indicator of what is actually going on, that is, a myriad of information-processing acts, most of which are not strictly lateralized in the intact brain even when it is dealing with linguistic material. This view fits in nicely with data indicating that whatever the true basis of the functional differences between the left and right hemispheres, it is not as simple as the verbal/nonverbal distinction (see, for example, Bryden & Allard, 1976; Shankweiler & Studdert-Kennedy, 1967).

Very subtle patterns of lateralization of various processes in the intact brain may be difficult to discern in the ERPs, at least without austere experimental controls. There are some indications, however, that global measures of the differential engagement of the two hemispheres may be evidenced in the ERPs to neutral probe stimuli (Galin & Ellis, 1975; Shucard, Shucard, & Thomas, 1977). For example, Shucard et al. (1977) found that late components of an auditory EP were reduced over the left hemisphere during a verbal task and over the right hemisphere while listening to music.

It is generally necessary to set up either a forced competition between the hemi-
spheres or a difficult task in order to reveal lateralized performance measures (which in most cases are still slight and statistical). When such behavioral tasks are obtained concurrently with ERP recordings (the epitome of the "combined experiment"), parallel ERP and behavioral asymmetries have been observed consistently. Thus, Neville and her associates (Neville, Kutas, & Schmidt, 1983a) recorded ERPs during a task in which the subjects were required to identify (in writing) a word that had been flashed briefly to either the right or the left visual field or pairs of words that had been presented simultaneously, one to each visual field. This task engendered a large behavioral asymmetry: every subject correctly identified significantly more words in the right than in the left visual field after both unilateral and bilateral word presentations.

As is evident in Figure 6, the ERPs elicited by these words displayed large and consistent lateral asymmetries. Over the posterior scalp, the asymmetries reflected the crossed anatomical organization of the visual system, whereas over the frontal and temporal scalp areas, the left-hemisphere ERP was consistently more negative in the 300-500 msec region than was the right hemisphere, irrespective of the field of presentation. The degree of ERP asymmetry was largest when the word identification was the best, namely, with unilateral right-visual-field presentations. Moreover, the ERPs from a control experiment, in which all stimulus parameters and response requirements were similar but the subjects were required to make same/different judgments about strings of nonalphabetic symbols, did not show this pattern of asymmetries. The implication is that the large asymmetries in the ERPs to words are specifically associated with some of the linguistic aspects of reading.

These data are of particular interest when compared with ERPs obtained in the same paradigm in congenitally deaf adults whose primary language is American Sign

![Figure 6](image-url)  
*Figure 6. Comparison of the left (solid line) and right (dotted line) hemisphere ERPs elicited by four-letter words presented in random order to either the left or the right visual fields or both fields simultaneously. The ERPs are grand averages over 10 hearing subjects and 8 deaf subjects. (Data presented in Neville, Kutas, & Schmidt, 1983a, b.)*
Language (Neville, Kutas, & Schmidt, 1983b). Briefly, the deaf adults differed from hearing subjects not only in their lack of behavioral asymmetries but also in the specific nature of the ERP asymmetries elicited by the English words (Figure 6). Deaf adults did not exhibit either the negative-positive shift in the left-hemisphere ERP or the greater differential between the hemispheres in response to words going to the right visual field as opposed to the left visual field or bilaterally, which were both so prominent in the hearing subjects.

Whatever the exact nature of the process underlying these electrophysiological asymmetries, these data indicate that the functional cerebral specialization during the identification of English words is different in normally hearing and congenitally deaf adults. ERP data of this sort, used in conjunction with behavioral studies of the effects of differential auditory and language experience, should help resolve inconsistencies in our knowledge of cerebral specialization, particularly in evaluating specific hypotheses about the functions of the left hemisphere and how they develop.3

While the predominance of the left hemisphere in language reception and production is hardly disputable, there is considerably less agreement about the language capabilities of the right hemisphere (Caramazza, Gordon, Zurif, & DeLuca, 1976; Dennis & Whitaker, 1976; Eisenson, 1962; Searleman, 1977). One approach to this problem has been opened through the study of commissurotomized (split-brain) patients whose hemispheres have been surgically separated (Gazzaniga, 1970; Sperry, 1974; Zaidel, 1978). We have attempted to use the N400 component of the ERP in such patients as a probe of the sophistication of right-hemisphere language functions. This approach was based on the premise that the elicitation of the N400 within our experimental design requires both an appreciation of semantic contexts and a realization that a particular word is inappropriate in a given context. As an initial step, we wanted to investigate the relative differences in the left and right hemispheres' processing of semantic information.

We did so by comparing the ERPs elicited by words that were semantically appropriate to the sentence context and those that were not, with the inappropriate word being routed to either the left or the right hemisphere (or in some cases to both). This procedure required a modification of the original N400-eliciting paradigm, wherein all but the final words of the sentences were presented aurally and, hence, to both hemispheres simultaneously. The last word of each sentence was flashed to each hemisphere separately; it was either the same or different and semantically appropriate or not in each field, following a randomized schedule.

The results from our investigation of five commissurotomized patients in this paradigm4 converge nicely with the behavioral assessments of the right hemisphere's language capacities in these patients. Whereas N400s were generated by all five of the patients when the semantically anomalous word was seen only by the left heme-

---

3 Some of the more popular formulations are (1) processes that require analytic versus holistic analysis; (2) the processing of complex acoustic information containing rapid-frequency transitions; (3) the perception of temporal sequences; and (4) processes involved in the grammatical recoding of language information, etc.

4 These data are part of an ongoing collaborative effort between our laboratory and Dr. Michael Gazzaniga and colleagues at Cornell Medical School.
Figure 7. Vertex ERPs elicited by simultaneous presentations of words to the two visual fields completing an auditory carrier sentence. The comparison is between predictable words that completed the sentences appropriately (normal words) and those that were semantically anomalous (deviants), which occurred at random in either the right (RVF) or the left visual field (LVF). These two patients represented the extreme cases, one having behaviorally well-established language capabilities in both hemispheres (P.S.) and the other having minimal language functioning in the right hemisphere (N.G.).

sphere, N400s to incongruent words presented to the right hemisphere were observed only in those two patients who had been shown to have a high degree of right-hemisphere language skills (see the two examples in Figure 7). While the scalp distributions of these ERPs recorded from the right and left sides of the scalp were puzzling in some cases, these findings illustrate how the recording of ERPs affords an evaluation of the cognitive and linguistic capabilities of the isolated hemispheres and their interaction through subcortical systems.

At this point, let us recall our original goal of understanding how the machinations of the brain give rise to the mind—not an easy task if we consider the definition of each in the Devil's Dictionary (Bierce, 1978): Brain: “an apparatus with which we think that we think”; and Mind: “a mysterious form of matter secreted by the brain. Its chief activity consists in the endeavor to ascertain its own nature, the futility of the attempt being due to the fact that it has nothing but itself to know itself with.” Hence the need for the cognitive neuroscience approach; may its emergent properties outwit if not transcend its methodological and inferential limitations.

References

Event-Related Potentials in Cognitive Science

Brown, W. C., Marsh, J. T., & Smith, J. C. Principal component analyses of ERP differences related to the meaning of an ambiguous word. Electroencephalography and Clinical Neurophysiology, 1979, 46, 709–714.
Desmedt, J. E. Language and hemispheric specialization in man: Cerebral ERPs: Progress in clinical neurophysiology (Vol. 3). Basel: Karger, 1977. (b)


Friedman, D., Simson, R., Ritter, W., & Rapin, I. The late positive component (P300) and information processing in sentences. *Electroencephalography and Clinical neurophysiology*, 1975, 38, 255–262. (b)


Johnson, R., & Donchin, E. P300 and stimulus categorization: Two plus one is not so different from one plus one. *Psychophysiology*, 1980, 17, 167–178.


