The Search for "Common Sense": An Electrophysiological Study of the Comprehension of Words and Pictures in Reading

Giorgio Ganis, Marta Kutas, and Martin I. Sereno
University of California, San Diego

Abstract

■ Event-related brain potentials (ERPs) from 26 scalp sites were used to investigate whether or not and, if so, the extent to which the brain processes subserving the understanding of imageable written words and line drawings are identical. Sentences were presented one word at a time to 28 undergraduates for comprehension. Each sentence ended with either a written word (regular sentences) or with a line drawing (rebus sentences) that rendered it semantically congruous or semantically incongruous. For half of the subjects regular and rebus sentences were randomly intermixed whereas for the remaining half the regular and rebus sentences were presented in separate blocks (affording within-subject comparisons in both cases). In both presentation formats, words and line drawings generated greater negativity between 325 and 475 msec poststimulus in ERPs to incongruous relative to congruous sentence endings (i.e., an N400-like effect). While the time course of this negativity was remarkably similar for words and pictures, there were notable differences in their scalp distributions; specifically, the classic N400 effect for words was larger posteriorly than it was for pictures. The congruity effect for pictures but not for words was also associated with a longer duration (lower frequency) negativity over frontal sites. In addition, under the mixed presentation mode, the N400 effect peaked about 30 msec earlier for pictures than for words. All in all, the data suggest that written words and pictures when they terminate sentences are processed similarly, but by at least partially nonoverlapping brain areas. ■

INTRODUCTION

A symbol can be defined roughly as something that represents an object, an event, or a relationship. The symbolic system par excellence is written language; however, we are able to understand symbols that are not strictly linguistic, for instance, pictures. And, in fact, there may be reason to believe that the mechanisms for the processing of words and pictures are similar if not identical because familiar iconic symbols[1] resemble written words in several respects: both (1) are composed of visual features, (2) have identical referents for different feature combinations (fonts for words; viewpoints for objects), (3) are fixated during recognition, and (4) are rapidly and perhaps automatically recognized once they have been learned. Indeed, there appears to be a historical continuity between iconic symbols and written words stemming from the pictographs of the earliest writing systems to the modern writing systems (Humphreys & Bruce, 1989). Note, however, that there are also some obvious differences between iconic symbols and written words. Perhaps the most important difference for present purposes is the more arbitrary nature of the relationship between a word's meaning and its visual form than between an object and its iconic representation.

These similarities and differences raise the question of the extent to which similar mental operations and/or brain areas are involved in comprehending iconic symbols and written words. For example, how similar are the processes that lead to the meaning of the word "lion" and those that identify the meaning from a picture of the same word; and are these computations performed by the same neural system(s)? Two broad theoretical positions have emerged from attempts to answer these questions. According to the *common semantic system hypothesis* "meaning" is represented in a common system (store) that is equally accessible by word and picture forms (e.g., Pylyshyn, 1980; Anderson & Bower, 1973). In the different variants of such a common system view, it is more or less explicitly acknowledged that pictures and words have different "representational power"; unlike words, pictures are presumed to be inherently better at representing concrete than abstract concepts (see Kolers & Brisson, 1984, for criticism). As a consequence, the range of concepts that is presumed to be equally accessible to both words and pictures is often confined to concrete, imageable objects.[2] More highly articulated versions of the common semantic system hypothesis propose that in the course of seeing or reading, visual information is mapped onto various presemantic perceptual recognition systems that specify the orthographic description of written words and the structural descriptions of objects (Morton & Patterson, 1980; Riddoch, Humphreys, Coltheart, & Funnel, 1988; Nelson, Reed, & McEvoy, 1977; Potter & Faulconer, 1975; Seymour, 1973; Durso & Johnson, 1979; Snodgrass, 1984; Theios & Amrhein, 1989).[3] Finally, some versions of the common semantic system hypothesis maintain that pictures have privileged access to the semantic system (Friedman & Bourne, 1976; Snodgrass and McCullogh, 1986; Caramazza, Hillis, Rapp, & Romani, 1990), leading to the prediction that pictures will be processed more quickly than written words, albeit in the same way.

By contrast, the *multiple semantic systems hypothesis* predicts that words and pictures access different specialized, semantic systems (e.g., Kosslyn and Pomerantz, 1977; Paivio, 1971, 1983, 1986). A prime example is Paivio's (1986) dual coding theory, which posits two structurally and functionally distinct subsystems: a system of logogens for processing words and a system of imagens for processing nonverbal materials; although independent, these two systems can communicate with each other.

Much research in experimental psychology has been aimed at adjudicating between the common and multiple semantic systems positions. Results with isolated words have generally been more variable than those with sentences. For example, isolated words and pictures have been found to prime each other in semantic priming tasks wherein the instructions and the experimental setting encourage semantic analysis (Sperber, McCauley, Ragain, & Weil, 1979; Vanderwart, 1984; Bajo, 1988; Theios & Amrhein, 1989), but not in tasks (such as naming) that did not emphasize semantic analysis (e.g., Durso & Johnson, 1979; Kroll & Potter, 1984; Roediger & Blaxton, 1987; Scarborough, Gerard, & Cortese, 1979; Weldon & Roediger, 1987). Typically, in word pair tasks, cross-modal priming (between words and pictures) has been taken as evidence for common semantic system hypotheses whereas the absence of cross-modal priming has been construed as supporting independent semantic systems.

When semantic analysis is guaranteed by embedding words and pictures into sentence contexts, the results have been most consistent with common semantic system hypotheses. For example, Kroll (1990) found equal facilitation on a reality decision task for both words and pictures when they were preceded by a related compared to an unrelated sentence. Thus, she concluded that the conceptual representations of sentences are built within a semantic system that can be accessed equally by words and pictures. Likewise, Potter, Kroll, Yachzel, Carpenter, and Sherman (1986) argued that since the time it took subjects to decide whether a sentence was plausible or implausible was independent of whether the sentences ended with a word or with a picture, both types of input must be processed in a common, mutually accessible system.

Another potential source of evidence that may differentiate between the common versus multiple semantic systems hypotheses comes from studies of patient populations. However, to date the available evidence can be interpreted as supporting both a common (e.g., Riddoch et al., 1988; Caramazza et al., 1990) and a multiple semantic systems approach (e.g., Shallice, 1993). As is typical of neuropsychological studies, dissociations form the basis of argument in favor of multiple semantic systems. Here we refer to a variety of dissociations in the performance of tasks requiring semantic analysis of visually presented objects and words in patients characterized by modality-specific aphasias (Shallice, 1987; Lhermitte & Beauvois, 1973; Denes & Semenza, 1975), agnosia without alexia (Albert, Reches, & Silverberg, 1975; Damasio, Damasio, & Van Hoesen, 1982; Gomori & Hawryluk, 1984; Karpov, Meerson, & Tonkonough, 1979; Levine & Calvanio, 1989; Mack & Boller, 1977; Ratcliff & Newcombe, 1982; Striano, Grossi, Chiacchio, & Fels, 1981), as well as modality-specific priming (Warrington & Shallice, 1979), modality-specific impairments of semantic memory (Warrington, 1975), and modality-specific category deficits (Warrington & Shallice, 1984; McCarthy & Warrington, 1988, 1990; Hart & Gordon, 1992; Hart, Lesser, & Gordon, 1992). By contrast, findings that lesions in some brain areas produce remarkably similar deficits in both verbal and nonverbal domains have been offered as evidence for the class of common semantic system hypotheses (Hillis, Rapp, Romani, & Caramazza, 1990; Marin, 1987; Sirigu, Duhamel, & Poncet, 1991).

An especially compelling case for a common semantic system can be made from the data on patients presenting with transcortical sensory aphasia (TSA) (Rubens & Kertesz, 1983; Alexander, Hiltbrunner, & Fischer, 1989; Vignolo, 1989; Hart & Gordon, 1990). TSA patients are defined by their poor auditory comprehension in the face of good repetition. However, for the purposes of the present argument it is important to note that a significant majority of TSA patients also exhibit visual object agnosia and visual field defects. Further note that TSA patients often have lesions in the left posterior inferotemporal cortex known to be critically involved in high level visual processes in monkeys. Bilateral lesions of inferotemporal cortex in monkeys result in a severe and lasting deficit in visual object recognition (Gross, 1973; Mishkin, 1982). It is such findings that led Sereno (1991a,b) to propose that the left posterior inferotemporal cortex plays a crucial role in the semantic analysis of both verbal materials and visual objects.

Basal temporal regions such as the fusiform gyrus likewise seem to have the attributes necessary for a common semantic area, namely a region where nonlinguistic visual and written language processing cohabit. On the one hand, electrical stimulation of the left basal temporal region in patients with intractable epilepsy has been found to produce specific language problems such as severe alexia (Luders et al., 1991). In addition, various regions along the fusiform gyrus have been recently found to be responsive to written language (McCarthy et al., 1995; Nobre, Allison, & McCarthy, 1994). On the other hand, these brain regions support visual functions in monkeys as well as in humans (e.g., Allison et al., 1994), suggesting that the basal temporal regions might be involved in written language as well as iconic symbol comprehension.

Complementary to the behavioral and neuropsychological approaches to delineating the nature of semantic representation and processing is the event-related brain potential (ERP) technique. To date, the majority of relevant ERP studies have focused on the semantic analysis of either words or pictures but not both. For example, Kutas and Hillyard (1980, 1984) showed that ERPs to words that were relatively unexpected, improbable, or semantically anomalous in their context were characterized by a large posteriorly distributed, slightly right hemisphere negativity between 300 and 600 msec (N400) postword onset; this negativity was small or absent to highly probable words. Subsequent studies demonstrated that the amplitude of the N400 was an inverse function of the extent to which a word fit its context. Moreover, the results of several studies support the hypotheses that (1) the N400 is sensitive to semantic analysis and (2) N400 elicitation is relatively independent of the specific surface form of the eliciting stimulus. Thus, for example, similar, albeit not identical, ERP effects are obtained whether congruous and incongruous sentences are presented as written text, speech, or the handshapes of American Sign Language (e.g., McCallum, Farmer, & Pocock, 1984; Kutas, Neville, & Holcomb, 1987; Holcomb & Neville, 1991).

Although words that are semantically anomalous in the context of a sentence generate the biggest N400, an equivalent component is also thought to be present in response to words in isolation and in word pairs. For example, in a lexical decision task the amplitude of a negativity between 300 and 600 msec to a word was found to be smaller when it was preceded by a semantically related word (e.g., cat–dog) than by a semantically unrelated word (e.g., table–dog) (e.g., Bentin, 1987; Holcomb, 1988; Holcomb & Neville, 1990). Other studies have shown that N400s are not elicited by physical deviations (Kutas & Hillyard, 1980, 1984) or certain morphological deviations within language (Kutas & Hillyard, 1983) or unexpected notes within familiar melodies (Besson & Macar, 1987; Paller, McCarthy, & Wood, 1992; Verleger, 1990). It is on the basis of these findings, among others, that we propose that variations in N400 amplitude provide a good index of semantic analyses in a way that makes the N400 a good candidate for investigating the question at hand, namely, whether or not words and pictures access a common semantic system.

Most of the ERP studies employing pictorial stimuli have focused on human face processing, on the assumption that the cognitive processes underlying face recognition are quite similar to those mediating word recognition (Barrett, Rugg, & Perrett, 1988). For example, Barrett et al. (1988) had subjects match black and white photographs of two serially presented familiar and unfamiliar faces; the second photograph of each pair was either a different view of the same person (match) or a different person (nonmatch). For familiar faces only, there was a larger negativity peaking around 400 msec for nonmatching (i.e., unprimed) than matching faces. The authors tentatively equated this negativity with the N400 effect seen in various language tasks. The results of a subsequent series of experiments showed that the amplitude of this N400-like effect to familiar faces was also modulated by judgments based on semantic or associative relatedness and not just identity (Barrett & Rugg, 1989). Specifically, the amplitude of an N450 component in the ERP to the faces of individuals with different occupations was larger than that to faces of individuals with the same occupations, when the task required matching on occupation. On the basis of these findings, Barrett and Rugg suggested that the neural processes subserving semantic priming of faces and words overlap.

Barrett and Rugg (1990) likewise obtained a larger N450 to the second of a sequentially presented pair of unrelated (e.g., fork–ring) pictures of common objects relative to the ERPs to related (e.g., fork–spoon) pictures; in this case judgments were based on relatedness. This negativity peaked approximately 50 msec later than is typical of word stimuli and had an amplitude that was roughly equipotential across the scalp. In this same task, Barrett and Rugg (1990) reported an earlier, more frontally distributed negativity (N300) that also differentiated mismatching from matching judgments. They speculated that this component might be specific to picture processing.

Overall, the finding of a consistently greater negativity in the ERP to mismatching than matching stimuli regardless of whether they are written words, spoken words, photographs of faces, or pictures of common objects has been taken as evidence in favor of an amodal semantic system.[4] Note that this conclusion implies that the reported differences in the amplitude, latency, and/or distribution of the negativities with different surface forms and different tasks were considered inconsequential relative to morphological similarity of the different "N400" effects.[5]

A notable exception to this view is found in the work of Holcomb and McPherson (1994), who also observed both an N300 and N400-like in an object decision task, but nonetheless concluded that the N400 effect in an object decision task was *not* the same as that seen in language tasks such as lexical decision. Specifically, Holcomb and McPherson found that (1) the ERPs to semantically unrelated objects were characterized by a larger negativity between 325 and 550 msec than those to related objects; (2) this negativity was larger for pseudoobjects (the equivalent of pseudowords in a lexical decision task) than for unrelated objects; and (3) there was a frontally distributed N300 effect (greater negativity for unrelated objects). Thus, during this object decision task semantic relatedness modulated the amplitude of a negativity with a time course similar to that seen in a lexical decision task. However, unlike N400s to written words, the N400s to pictures were larger over frontal than posterior sites (being almost absent over occipital regions) and were somewhat larger over the left than the right hemisphere. While the apparently more frontal distribution of the N400 could be accounted for in terms of residual overlap of the more frontally distributed N300 as suggested by the authors, this account cannot explain the absence of an N400 effect over occipital sites. Thus, the data seem more consistent with the view that the effects for visual objects and for words differ from each other. On the other hand, the functional equivalence between the object decision and lexical decision tasks on which Holcomb's study was predicated has been questioned (e.g., Kolers & Brisson, 1984). We think a much cleaner and more direct comparison of the ERP results would be possible if the words and pictures were presented to the same subjects within the same experimental paradigm.

Only two studies have reported a direct comparison between words and pictures within the same task. The first was an unpublished study by Kutas[6] including sentences ending with either a semantically congruous word or picture, or incongruous word or picture. There was a similar-looking N400 effect for words and pictures, although the peak latency of the N400s to pictures was shorter than that to words, and pictures were associated with greater frontal negativity overall. There are a number of reasons, however, to be cautious about the interpretation of these results. First, the words and pictures were not equated in size; the pictures were physically much larger than the written words. This difference might explain the shorter N400 peak latency to the pictures. Second, the probability of the sentences ending with a picture (congruous or incongruous) was much lower than the probability that it would end with a word; this imbalance in stimulus probability could have led to differential elicitation of probability-sensitive components such as the P3 (Duncan-Johnson & Donchin, 1977). Third, there was no attempt to equate the imageability of the word and pictures. Fourth, there were only eight recording sites on which to base distributional comparisons.

Recently a direct comparison between the processing of words and pictures in a sentence reading task was reported by Nigam, Hoffman, and Simons (1992). They equated their words and pictures in size. However, there were some major differences with the Kutas study: specifically, in Nigam et al. the modality of the sentence final stimulus was a between-subject factor and therefore whether a sentence ended with a word or a picture was completely predictable. One group of subjects read sentences that ended with congruous or incongruous words and a different group of subjects read sentences that ended with congruous or incongruous pictures. Semantically anomalous words and pictures alike were associated with larger negativities (N400s) than were their congruent counterparts. Moreover, as Nigam et al. observed no significant differences in the amplitude, latency, or scalp distribution of the N400 effects to words versus pictures, they concluded that it reflected activity in an amodal semantic system.

We think that this conclusion was premature for a number of reasons including (1) use of between-subject as opposed to within-subject comparison of the word versus picture data, (2) limited number of recording sites, and (3) potential lack of generalizability given use of blocked presentations for word and picture conditions. In combination with the relatively slow presentation rates and high predictability of the sentence final items, the blocked design may have engendered specific encoding strategies. For example, in the picture condition subjects might have predicted the last word and used this to generate an image. If so, the N400 effect would be influenced by the match/mismatch between the image and the actual picture presented rather than reflecting only the relation between the context and sentence final item.

The present study was designed to eliminate the confounds in the Kutas and Nigam et al. studies and thereby provide evidence pertaining to the question of whether or not words and pictures access a common semantic system. As in the previous studies we chose to use sentential contexts both because they yield large and reliable N400 effects in response to semantic anomalies (e.g., Kutas et al., 1987) and because they encourage semantic processing.

Two variables were manipulated within subjects: *modality* of the target (pictures, words) and *congruity* of the sentence ending (congruous, incongruous). An example of a congruous sentence ending with a picture is shown in Figure 1. An additional variable, "blocking" was manipulated between subjects (mixed modalities and blocked modalities) to assess the effect of the predictability of the modality of the ending. This design yielded a total of eight conditions, and afforded within-subject comparisons between ERPs to words and pictures in two different blocking conditions as well as between subjects comparisons across the blocking conditions.

We expected to find the classic N400 effect in response to written words in both the blocked and mixed conditions, namely, a posteriorly distributed monophasic negativity, slightly larger over right than left hemisphere

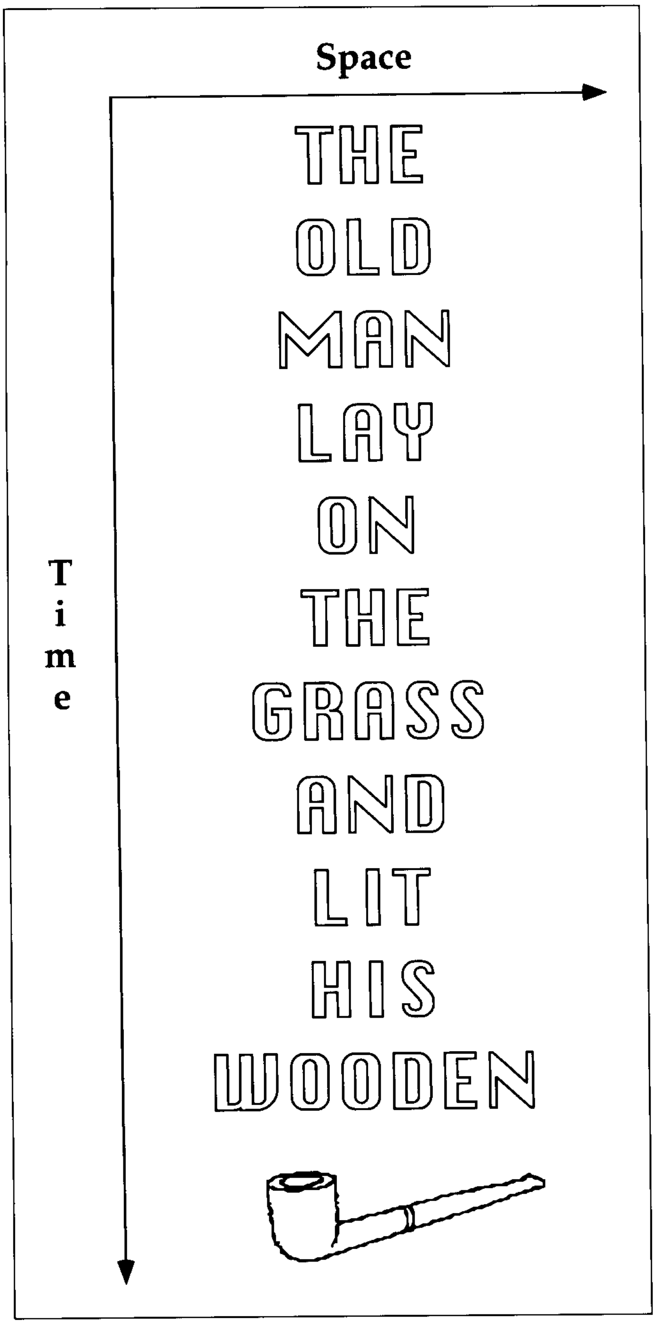

Figure 1. One of the sentences used in the experiment. This is the actual font employed (Chicago). The relative size of the words and the picture is very close to reality.

sites, onsetting around 200-250 msec, and peaking around 400 msec. Overall, we did not expect the N400 effect to differ as a function of blocking condition, although we thought that there would be some indication in the raw ERPs to pictures of the occasional "surprise" if the sentence ended in a picture when the subject expected a word (mixed condition). In general, we will take "similar" N400 effects for words and pictures as evidence for a common semantic system theory and "different" N400 effects as evidence against it. More specifically, (1) identical N400 effects (onset and/or peak latency, amplitude, and distribution) for words and pictures will be interpreted as favoring a common semantic system theory (e.g., Theios & Amrhein, 1989); (2) an N400 effect that differs only in onset and peak latency, most likely a shorter latency for pictures than for words, would be viewed as compatible with common semantic system theories that afford pictures privileged access to semantic representations (e.g., Potter & Faulconer, 1975); (3) N400 effects for words and pictures that differ only in amplitude but not in spatial distribution, will also be considered as consistent with common semantic system theories; specifically those that posit different access procedures for words and pictures (e.g., Caramazza et al., 1990); (4) absence of an N400 effect for pictures will be taken as evidence for the presence of multiple semantic systems; (5) insofar as the N400 effects to words and pictures are similar in morphology but not in spatial distribution we will assume that both are processed for meaning in similar ways but in different brain areas; and (6) N400 effects for words and pictures that differ both in morphology and distribution will be taken as strong evidence for multiple semantic system theories that assume differential processing of words and pictures by different brain systems (e.g., Paivio, 1990).

RESULTS

A similar sequence of early components characterized the ERPs to words in all the conditions.[7] At frontal sites (Figs. 2 and 3, e.g., frontal sites 2-10) these included a negativity peaking around 150 msec (N1) followed by a positivity peaking around 250 msec (P2). At occipital sites (Figs. 2 and 3, e.g., occipital sites 5-7) the earliest evoked potentials (EPs) were the P1-N1-P2 sequence: a positive deflection peaking around 130 msec (P1), followed by a negative deflection peaking around 210 msec (N1) and by a positive one peaking around 280 msec (P2). Notice that there were some other, smaller deflections triggered by the offset of the previous word (which occurs 50 msec prior to the onset of the final item) that preceded and partially overlapped the P1. The early potentials were followed by a broadly distributed negative deflection peaking around 400 msec (N400) that was larger for the incongruous than congruous endings, and was maximal at parietal sites (Figs. 2 and 3, e.g., site 26). At posterior sites the N400 was followed by a late positive deflection starting around 500 msec whose peak latency seemed to be later for incongruous than for congruous endings (Figs. 2 and 3, e.g., sites 5-7).

Pictures in both blocking conditions elicited frontal N1-P2s similar to those for words (Figs. 2 and 3, e.g., sites 2-10). The P1/N1 deflections were also visible at occipital sites; however, the P2 appeared to be masked by a large posterior positivity peaking around 380 msec that was not present in the response to words (Figs. 2

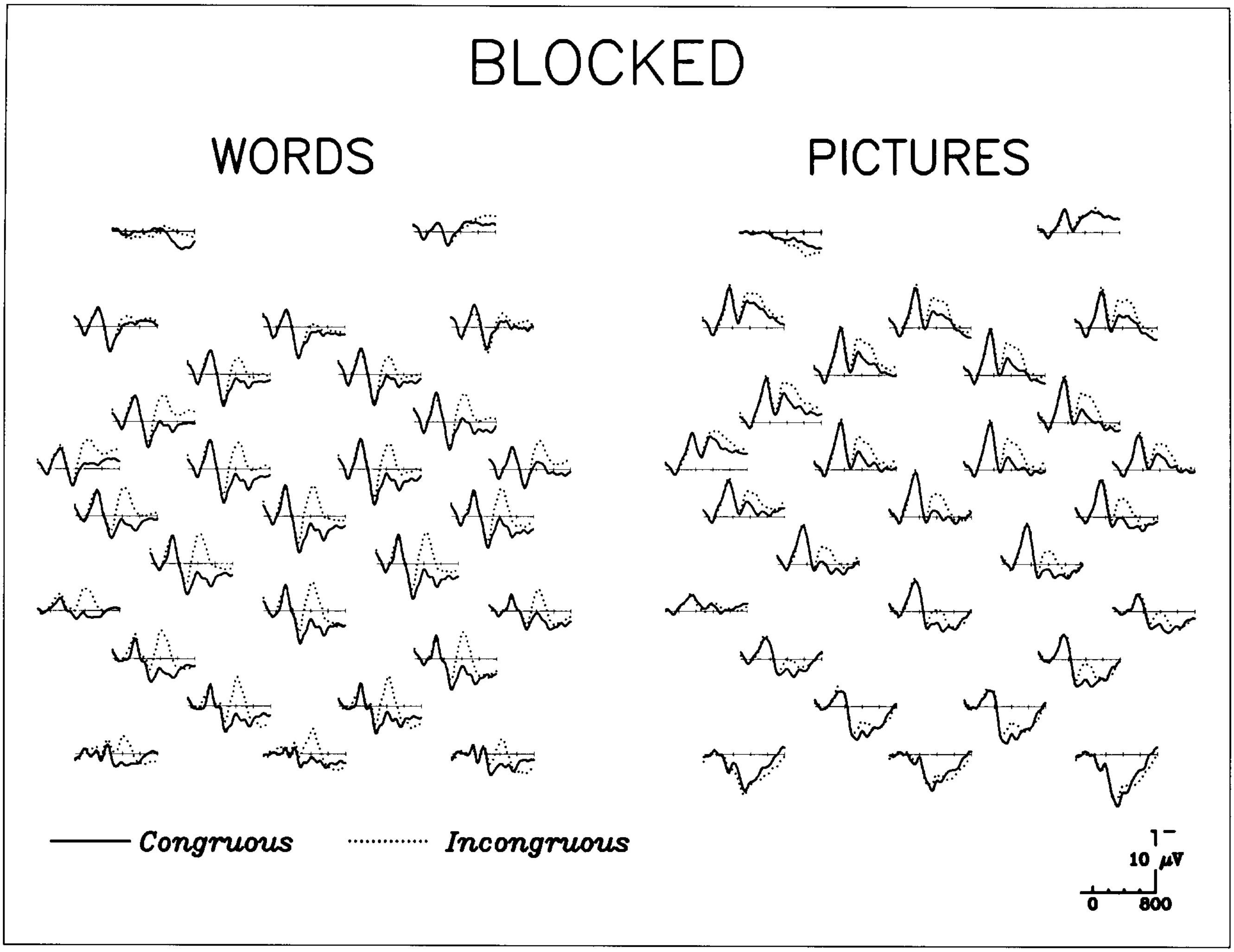

Figure 2. Grand average (N = 12) ERP waveforms obtained in the blocked condition for words (left) and pictures (right). Congruous endings are indicated by a solid line, while incongruous ones are indicated by a dotted line. The locations and labels of the electrodes are shown in the Method section, Figure 9. The topmost left and right plots in each panel show horizontal and below the eye data, respectively (see Method section for details). Note that in this and following plots, negative is up.

and 3, e.g., sites 5–7). A broadly distributed negative deflection peaking around 400 msec (N400) that was larger for incongruous than congruous endings and was maximal at central and anterior sites followed the early potentials (Figs. 2 and 3, e.g., sites 21–25). This N400 was followed by a late positive deflection starting around 550 msec that appeared to be largest at posterior sites in the mixed presentation condition (Figs. 2 and 3, e.g., sites 5–7).

Mean Amplitude Measures

The mean amplitude data in three time windows were subjected to repeated measures ANOVA, with degrees of freedom adjusted when necessary by the Geisser-Greenhouse procedure (Geisser & Greenhouse, 1959). For the main analyses on raw ERPs four independent variables were used, unless otherwise specified: blocking (mixed vs. blocked), congruity (congruous vs. incongruous), modality (pictures vs. words), and site (26 electrode sites). An alpha of 0.05 was used in all the analyses.

Early Congruity Effects

To examine congruity effects preceding the N400 an ANOVA was performed on the mean amplitude of the raw waves between 175 and 325 msec. Pictures were generally more negative than words [main effect of modality, $F(1,22) = 20.83$] and incongruous items were more negative than congruous ones [main effect of congruity, $F(1,22) = 29.69$]. Neither the early potentials [main effect of site, $F(25,550) = 11.58$] nor the congruity effect were uniform across scalp sites [congruity by site interaction, $F(25,550) = 3.95$]. Thus, both the early EPs and the congruity effect to words tended to be more negative at posterior than at anterior sites, while the opposite was true for pictures [interaction modality by site, $F(25,550) = 70.86$; interaction modality by congru-

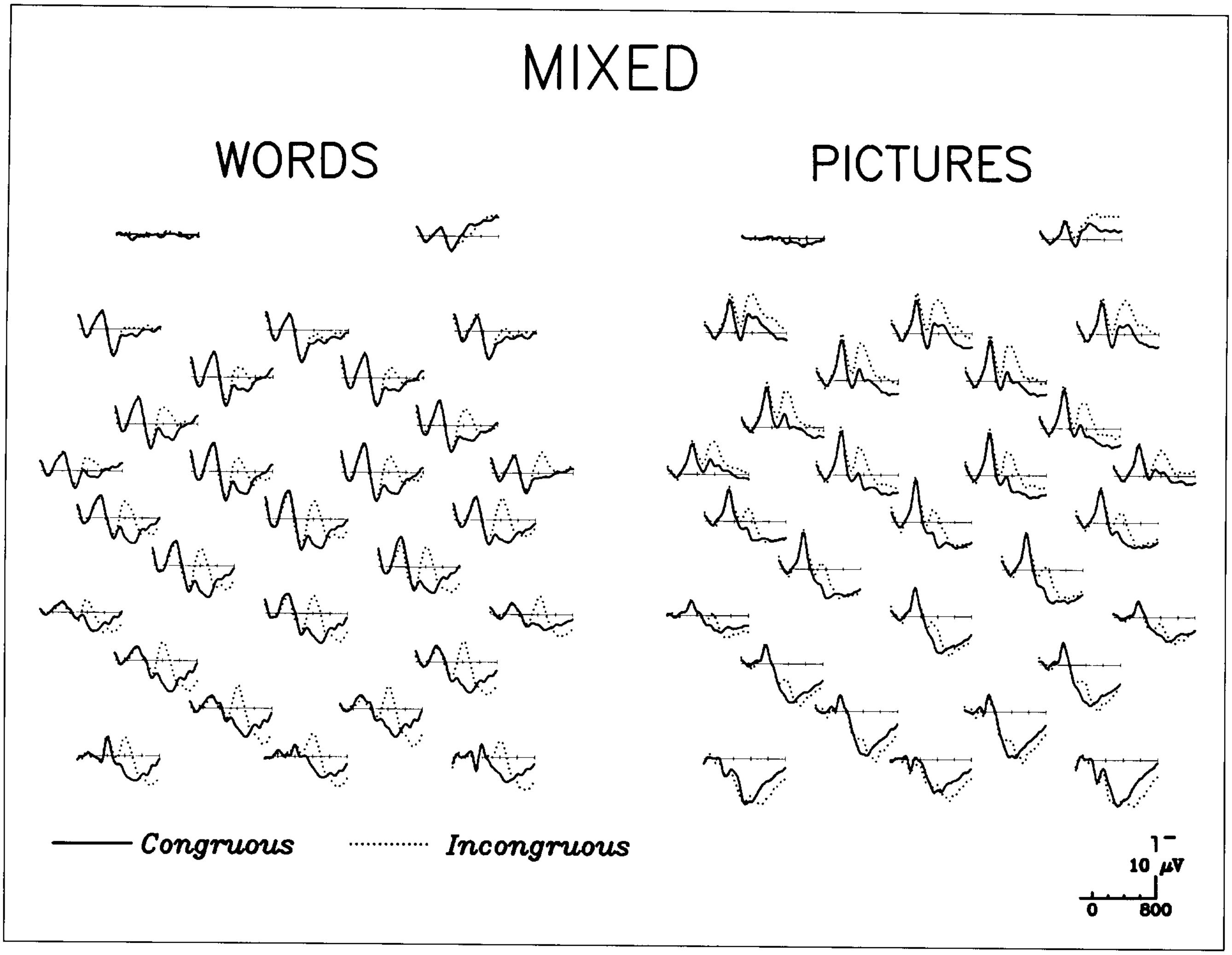

Figure 3. Grand average (N = 12) ERP waveforms obtained in the mixed condition for words (left) and pictures (right). Congruous endings are indicated by a solid line, while incongruous ones are indicated by a dotted line.

ity by site, $F(25,550)$ = 10.34]. Blocking took part in a marginally significant three way interaction with congruity and site [$F(25,550)$ = 2.20, p = 0.047]. Follow-up analyses were conducted on words and pictures in both blocking conditions, separately. There was a significant congruity effect for pictures, largest at frontal sites [interaction congruity by site: blocked $F(25,275)$ = 3.61]; mixed $F(25,275)$ = 14.14]. There was also a marginally significant effect of congruity for words in the blocked condition, with a central/posterior maximum [interaction congruity by site, $F(25,275)$ = 2.89, p = 0.049], but not in the mixed condition.

To better examine whether there was an earlier congruity effect for pictures than for words, we conducted the same analysis on an earlier time window, specifically 150-275 msec. The outcome of the main analysis was very similar to that described above, with the only difference being that no effects of blocking were found. Follow-up analyses performed on words and pictures separately indicated that words did not show a congruity effect in this time window [main effect of congruity, $F(1,22)$ = 0.45, N.S.; interaction congruity by site, $F(25,550)$ = 0.75, N.S.], whereas pictures did but mainly at frontal sites [interaction congruity by site, $F(25,550)$ = 11.30]. This early congruity effect for pictures tended to be larger in the mixed than in the blocked condition [interaction blocking by congruity by site, $F(25,550)$ = 2.80, p=0.03]. An additional follow-up analysis was performed on the picture data for blocked and mixed conditions, separately. Both analyses showed an interaction of congruity by site, reflecting the presence of a congruity effect for pictures in this time window with a frontal maximum. In summary, between 150 and 275 msec there was an early, frontally distributed congruity effect for pictures but not for words.

N400 Congruity Effect

The time window 325-475 msec was chosen to capture the N400 effect without contamination from preceding and following components (such as P2, N300, and late positive components, e.g., Holcomb & McPherson, 1994;

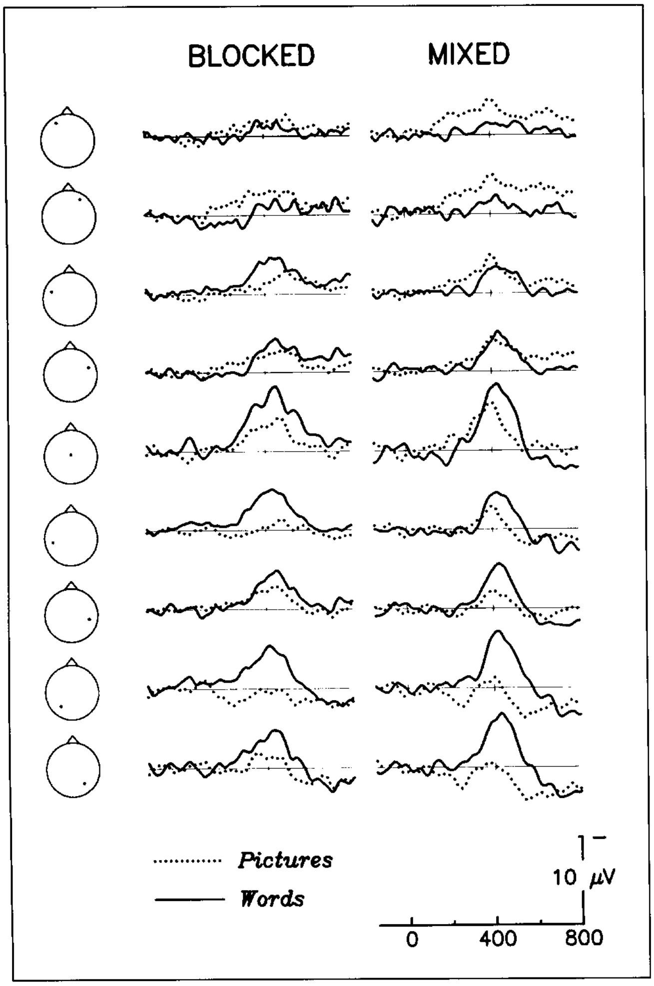

Figure 4. Difference ERP waveforms (incongruous minus congruous) obtained in the blocked (left) and mixed (right) conditions for words (solid line) and pictures (dotted lines). Only nine representative electrodes are shown for clarity.

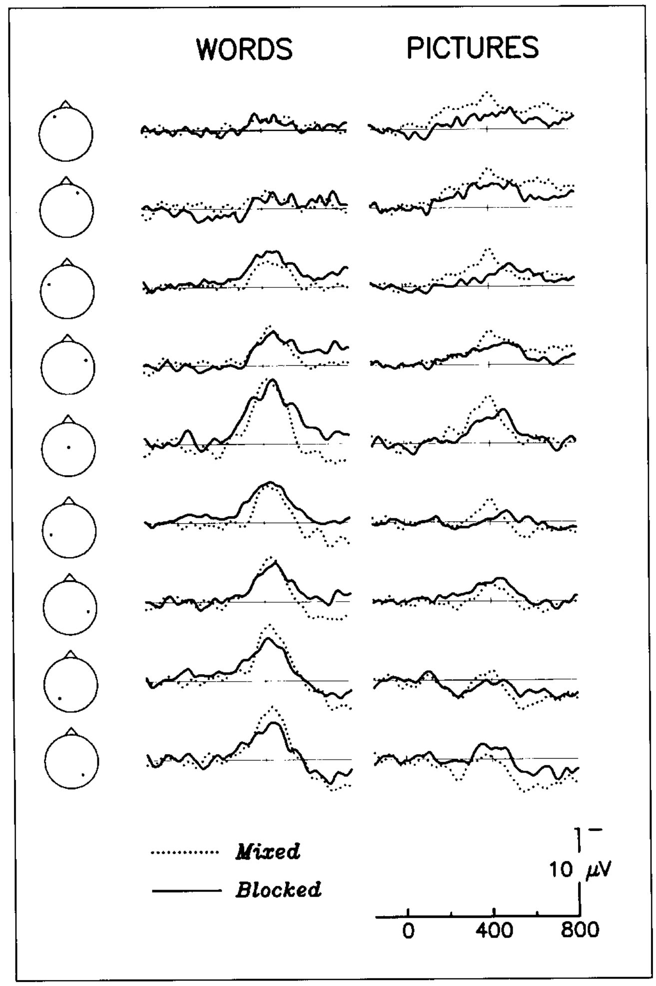

Figure 5. Difference ERP waveforms (incongruous minus congruous) obtained for words (left) and pictures (right) in the blocked (solid line) and mixed (dotted line) conditions.

Heit, Smith, & Halgren, 1990). First, an ANOVA was performed on all the difference waveform data (Figs. 4 and 5 show difference waveforms from selected sites). The three independent variables were blocking (blocked vs. mixed), modality (pictures vs. words), and electrode site (26 locations). As can be seen in Figure 4, the N400 effect was larger for words (−4.08 μV) than for pictures [−2.57 μV; main effect of modality $F(1,22) = 9.52$] and was largest at central recording sites [main effect of site, $F(1,22) = 13.82$]. Figure 4 also hints at the difference in the distribution of the N400 effect for words and pictures [modality by site interaction, [$F(25,550) = 16.35$].

This difference can be better seen in Figure 6, where the distributions of the N400 effect for words and pictures in both blocking conditions are compared directly. To determine whether this interaction reflected a real difference in the scalp distributions, the approach recommended by McCarthy and Wood (1985) was used. Thus, the data for each subject were normalized across electrode sites, within each modality separately, thereby eliminating any systematic amplitude differences between modalities, and allowing a purer assessment of the differences in scalp distributions. These normalized data were then subjected to an ANOVA, with blocking (blocked vs. mixed), modality (words vs. pictures), and site (26 electrode sites) as independent variables. There was a significant interaction of modality by site [$F(25,550) = 12.49$], indicating that the N400 effect for words and pictures was differentially distributed across the scalp.

To better analyze the modality by site interaction, planned comparisons were performed on the data from each symmetric pair of lateral sites (11 total) with blocking (blocked vs. mixed), modality (words vs. pictures), and hemisphere (left vs. right) as independent variables. As can be seen in Figure 6, the N400 effect for pictures was larger than that for words over frontal sites [e.g., pair 2-10, main effect of modality, $F(1,22) = 6.62$], while the reverse pattern held over parietal sites [e.g., pair 14-17, main effect of modality, $F(1,22) = 25.99$] and occipital sites [e.g., pair 15-16, main effect of modality, $F(1,22) = 26.17$]. This difference can be better appreciated in the maps of Figure 7, where the normalized amplitudes of

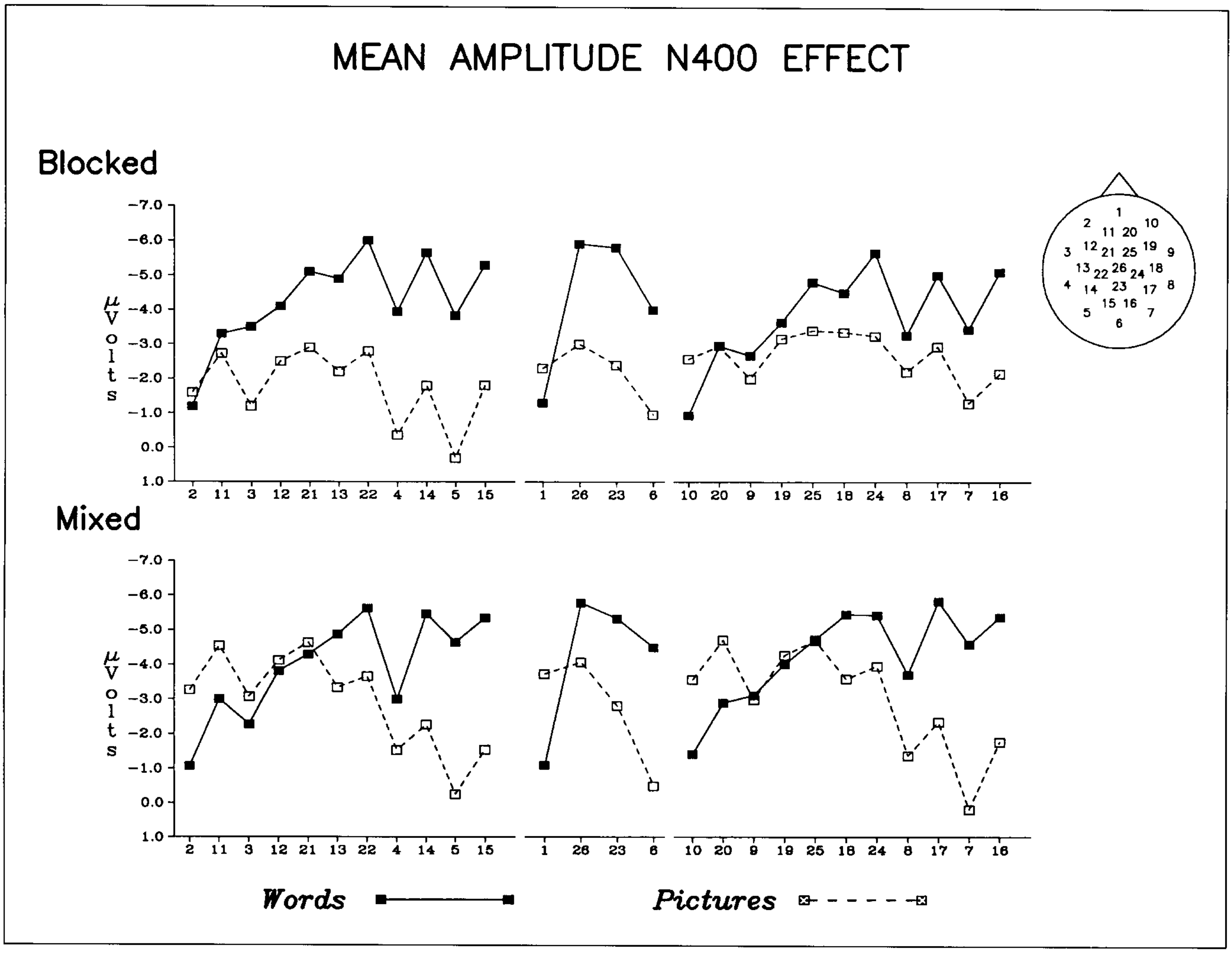

Figure 6. Distribution of the N400 congruity effect (mean amplitude of the difference waves between 325 and 475 msec) in the blocked and mixed conditions, for words (solid line) and pictures (dashed line). Each panel plots electrodes starting from frontal sites. Corresponding positions in the leftmost and rightmost panel within each blocking condition refer to symmetric pairs of electrodes (e.g., electrode pair 2-10). The central panel within each blocking condition corresponds to midline electrodes. A diagram depicting the electrode positions is included for reference.

the congruity effect between 325 and 475 msec are plotted for all four blocking by modality conditions. Marginally significant blocking by modality by hemisphere interactions were found for five pairs of electrodes: temporal pairs 4-8 (p = 0.041) and 3-9 (p = 0.047), temporoparietal pair 13-18 (p = 0.047), occipitotemporoparietal pair 14-17 (p = 0.042), and parietal pair 21-25 (p = 0.025). The triple interactions are due to the fact that at these sites the congruity effect for pictures tended to be symmetric in the mixed condition, but larger on the right side in the blocked condition, whereas the opposite pattern was observed for words (see Figs. 6 and 7).

Finally, to better assess the hemispheric asymmetries, separate analyses were performed on words and pictures with site (11 possible symmetric pairs) and hemisphere (left or right) as variables. As is evident in Figures 6 and 7, the N400 effect was asymmetric only for pictures and only in the blocked condition [main effect of hemisphere, $F(1,11)$ = 5.71], with the right hemisphere being more negative than the left one. In all other conditions, the N400 effect was symmetric.

An ANOVA performed on the mean amplitude of the raw ERPs between 325 and 475 msec confirmed the results of the previous analysis. To allow a better comparison with the results of previous studies, all analyses were repeated for the time window 300-500 msec. Neither the results nor the conclusions to be drawn were affected in any way by this change in the measurement window. Thus, they will not be discussed further.

In summary, in the 325-475 msec time window (1) a congruity effect was present for words and pictures in both blocking conditions; (2) the congruity effect was larger for pictures than for words at frontal sites, and larger for words than pictures at occipital sites; and (3) the only hemispheric asymmetry was found for pictures

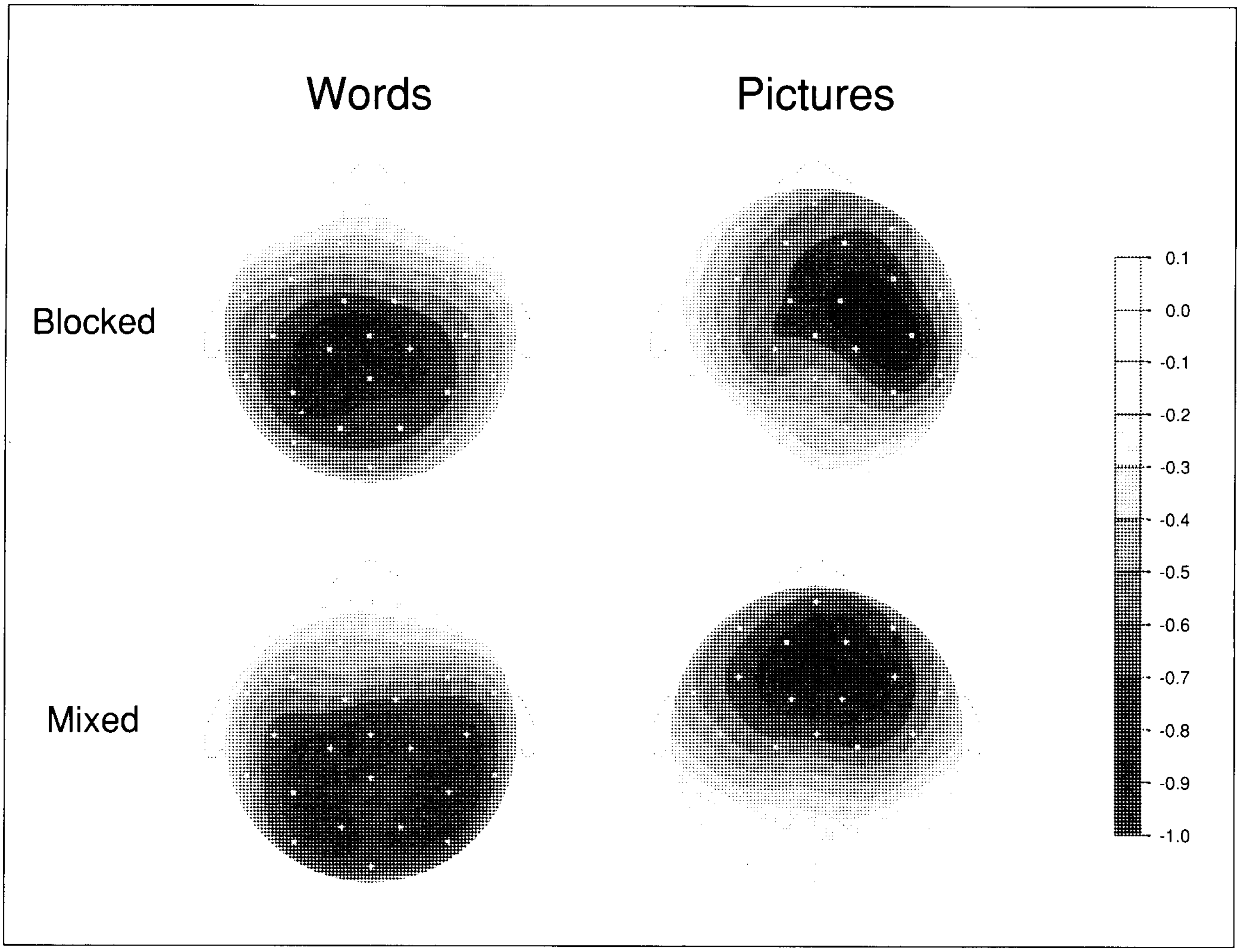

Figure 7. Isopotential gray-scale maps of the normalized distribution of the N400 effect (mean amplitude of the difference waves between 325 and 475 msec) in the blocked (top) and mixed (bottom) conditions for words (left) and pictures (right). The original scattered data (26 scalp sites) were interpolated with a spherical spline algorithm (Hassainia et al., 1994).

in the blocked condition, with greater negativity over the right than left hemisphere.

Late Congruity Effects

An ANOVA was performed on the mean amplitude of the averaged ERPs from 575 to 800 msec, which was meant to capture late congruity effects (Figs. 2 and 3). This late congruity effect was negative at anterior sites and positive at posterior sites [interaction congruity by site, $F(25,550) = 21.91$]. Moreover, these late potentials were more positive to pictures than those to words at posterior sites [interaction modality by site, $F(25,550) = 4.06$]. Finally, these late congruity effects seemed to differ in distribution for words and pictures [interaction modality by congruity by site, $F(25,550) = 3.13$]. Follow-up analyses revealed no significant effects of congruity for words but significant effects for pictures in both blocking conditions [blocked: interaction congruity by site, $F(25,275) = 5.20$; mixed: interaction congruity by site, $F(25,275) = 16.34$].

Latency Measure

On visual inspection there appears to be a shift in the peak latency of the N400 effect between words and pictures in the mixed condition (see Fig. 4): that is, the N400 difference ERP peaks at a shorter latency for pictures than for words. An ANOVA with blocking (blocked vs. mixed), modality (pictures vs. words), and site (26 electrode sites) as independent variables was performed on the N400 difference peak latencies. The analyses revealed a significant interaction between blocking, modality, and site [$F(25,550) = 2.77$]. Planned follow-up analyses of the mixed condition showed that the N400 peak latency was significantly shorter for pictures than for words [397 vs. 414 msec, respectively; $F(1,11) = 10.49$]; the size of this latency difference was about 30 msec at parietal sites and was not uniform across scalp electrodes [modality by site interaction, $F(25,275) = 3.01$]. The blocked condition analysis revealed no significant difference in the peak latency of the N400 effect to pictures and words [410 vs. 416 msec, respectively; main effect of modality, $F(1,11) = 0.31$, N.S.; modality by site interaction, $F(25,275) = 1.15$, N.S.].

The finding that the N400 effect, at least in the mixed presentation condition, is earlier for pictures than for words was present in 8 out of 12 subjects (mean = 28 msec), but absent in the remaining four subjects (mean = −3 msec). This might have been due partly to the difficulty of locating the peak of the N400 at those sites where the effect was small; to circumvent this problem, we restricted our ANOVA to data from the parietal sites (electrodes 22, 26, and 24), where the N400 effect was reliable for both words and pictures. With modality (pictures vs. words) and site (3 electrode sites) as the independent variables, we obtained a significant main effect of modality [$F(1,11) = 9.64$]. For this analysis, the advantage for pictures was present in 10 of the 12 subjects. Perhaps these differences reflect the different strategies that individual subjects used to encode the words and pictures.

Postrecording Recognition Test

For sentences ending with a word, on average subjects recognized 57% of the congruous and 63% of the incongruous endings. For sentences ending with a picture, subjects recognized 72% of the congruous and 79% of the incongruous endings. An ANOVA was performed on the *d* primes calculated by taking into account both hits and false alarms. The independent variables were modality (words vs. pictures), congruity (congruous vs. incongruous), and blocking (blocked vs. mixed). Overall, subjects performed above chance [average $d' = 2.0729$; $F(1,22) = 6300, p < 0.0001$]. Pictures were recognized better than words [$F(1,22) = 30.96, p < 0.0001$]. Incongruous items were recognized better than congruous ones [$F(1,22) = 19.48, p < 0.001$]. There were no effects of blocking [$F(1,22) = 0.16$, N.S.].

DISCUSSION

The classic N400 effect for words was replicated in both blocked and mixed presentation modes. Likewise, an N400-like effect was observed for pictures, also in both presentation modes. The time course of these effects was in large part independent of whether the sentence endings were words or pictures, indicating a similarity in the temporal course of the neural processes underlying the semantic analysis of both in sentential contexts. These similarities, however, were accompanied by notable differences in the scalp distributions of the congruity effects for words and pictures, implying at least partial nonoverlap of responsible neural structures. Two aspects of these topographical differences are worth discussing; specifically, the greater posterior extent of the word N400 effect and the greater frontal extent of the picture congruity effect.

At parietal and occipital sites the N400 effect was larger for words than for pictures. This is consistent with Holcomb and McPherson's (1994) finding that pictures in an object-decision task elicited virtually no N400 activity over posterior sites. Given the difference between their task and ours, however, this comparison must be made with caution. There was also a visible trend for the N400 congruity effect to be smaller for pictures than for words at the occipital site (Oz) in the Nigam et al. (1992) study, although their analysis revealed no statistical differences (see their Fig. 3, p. 18).

Our finding that the N400 at posterior sites was smaller for pictures than for words when they terminated written sentences nonsensically is compatible with at least three different hypotheses. For one, it may be the case that the neural generators of the N400 to words and pictures actually do differ and this difference in the activated brain areas is manifest in the different scalp distributions.

Another possibility, however, is that exactly the same neural structures (i.e., parts of the anterior fusiform gyrus according to McCarthy et al., 1995, Nobre et al., 1994) were engaged by the processing of semantically anomalous words and pictures alike but overlap of a late posterior positivity (of the P3 family for instance) resulted in the apparent reduction of the N400 effect to pictures at posterior scalp sites. A late positive component, with a distribution consistent with that reported for the visual P3, was in fact present in the raw ERPs over posterior sites (Onofrj et al., 1990). Certainly, by some views of the P3, a relatively unexpected switch from the printed words of the sentence context to the terminal picture would be sufficient to elicit some P3 activity. At least two factors could modulate the amplitude of such a P3: the relative probability and the predictability of the stimulus modality of the sentence final item. Overall, pictures were much less frequent than words, and more so in the mixed (1/16) than in the blocked presentation conditions (1/8). All else being equal, this factor would produce greater P3 activity for pictures than for words, and more so in the mixed than in the blocked presentation mode (e.g., Donchin & Coles, 1991). Moreover, the sentence final item's modality was predictable in the blocked but not in the mixed conditions; note, however, that even in the blocked conditions the exact moment at which the modality switch occurred was not easily predictable as the sentences were not constant in length. This factor in isolation would probably produce little if any P3 activity to word endings but moderate P3 activity to picture endings in the blocked presentation mode; large P3s would be produced in the mixed presentation mode for both word and picture endings, as they were equally unpredictable.

An explanation combining the contributions of these two factors therefore would predict the smallest P3 for words in the blocked condition and the largest P3 for pictures in the mixed condition. Pictures in the blocked and words in the mixed conditions would generate P3s intermediate in amplitudes to these extremes, depending on the relative contributions of the two factors. In our data there were only nonsignificant trends consistent with these predictions.

To account for the smaller posterior N400 effect to pictures than words, the overlapping P3 would have to be modulated not only by stimulus modality, but also semantic congruity. If the P3 were the same size for congruous and incongruous endings, then its contribution would not be evident in the N400 difference wave (the P3s would cancel out in the subtraction). That is to say, the N400 and the P3 would be additive with respect to congruity. On the other hand, if the P3 were larger for incongruous than congruous endings (i.e., if the N400 and P3 were interactive with respect to congruity), then the N400 difference ERP would appear to be smaller than it really is. Thus, a stronger modulation of the P3 by congruity for pictures than for words could result in an apparent reduction of the N400 effect to pictures relative to words. Although some modulation of the P3 by congruity is likely, the evidence to date points to an additive rather than interactive relationship. In an experiment using short Kanji sentences that were either true (red is red) or false (red is blue), Katayama and Yagi (1992) observed both an N400 and a late positivity. The colors for verification were presented either as words or as colored patches. The results showed larger negativities (presumed N400s) to mismatching than matching colors regardless of modality as well as a large posterior positivity to the color patches that was equivalent for matches and mismatches. Similarly, Kutas and Hillyard (1980) found that the N400 elicited by semantic incongruity and P560 elicited by a physical deviation (change in font size) were additive when semantically incongruous words were presented in unexpectedly large type. It is on the basis of these findings that we think that the smaller posterior N400 effect to pictures than words cannot be explained fully by an overlapping P300.

Another potential, albeit less interesting, explanation is that the distributional differences are a consequence of the fact that word and picture conditions included different sentence contexts. We conducted an additional experiment on seven new subjects to rule out this possibility. For this control experiment, each sentence of the mixed condition that had ended with a picture in the original experiment was now ended by a printed word (i.e., the name of the picture). Thus, subjects experienced all of the sentence fragments from the mixed presentation mode, except all the sentences terminated with words. If some uncontrolled difference between the sets of sentences used in the word and picture conditions were responsible for the topographic ERP difference between words and pictures in the original experiment, then a similar pattern of results should obtain even if the pictures were replaced by their names. This was not the case. As can be seen in Figure 8, semantically incongruous minus congruous words gave rise to an N400 effect that had the same scalp distribution for both sets of sentences in the control experiment as it had for words in the original experiment. An ANOVA of the control data revealed no main effect of set [$F(1,6) = 0.06$, N.S.], or interaction with congruity [$F(1,6) = 0.55$, N.S.], and no interaction between set, congruity and site [$F(25,150) = 0.45$, N.S.]. It seems unlikely, therefore, that any of the topographic differences in N400 scalp distribution for pictures and words were due to some differences in the materials used.

The other main topographic difference between words and pictures was the presence of a larger congruity effect for pictures at frontal sites. This is generally consistent with previous reports of a greater frontal negativity between 300 and 500 msec for nonmatching than matching pictures (e.g., Barrett & Rugg, 1990; Holcomb & McPherson, 1994) as well as for incongruous relative to congruous pictures in sentences (compared to congruous ones, Kutas & Van Petten, 1990). Nigam et al. (1992) did not observe this frontal difference; however, they did not record at frontal sites.

At frontal sites the congruity difference ERP for pictures onset around 150 msec (before that for words) and peaked around 400 msec. Perhaps this indicates that while the meaning of pictures was accessed more quickly than that of words, its integration within a context occurred more gradually. The quicker access for pictures would follow from the less arbitrary relationship between pictures and their referent in the real world than that between written (or spoken) words and their real-world referents. For example, pictures of animals have some visual features in common that distinguish them from utensils, which simply are not evident in the actual names. Given that the brain tends to make use of whatever information is available as quickly as it can to decipher meaning, it is not unreasonable to think that such perceptual aspects of a picture may serve to initiate a congruity effect even before the object has been fully identified. Note that the frontal electrodes (sites 2 and 10) are situated almost in front of the temporal poles and thus could easily pick up activity from these brain regions.[8] There is evidence that the temporal poles are involved in human object recognition (e.g., Sirigu et al., 1991). Neurophysiological studies on monkeys also seem consistent with this account.[9] For example, Mikami, Nakamura, and Kubota (1993) found a high percentage of neurons that responded maximally to just one stimulus (stimulus selective) as well as "categorically selective" neurons in monkey temporal poles. The categorically selective neurons gave virtually maximal responses to all the members of one of four categories (human faces, monkey faces, objects, abstract

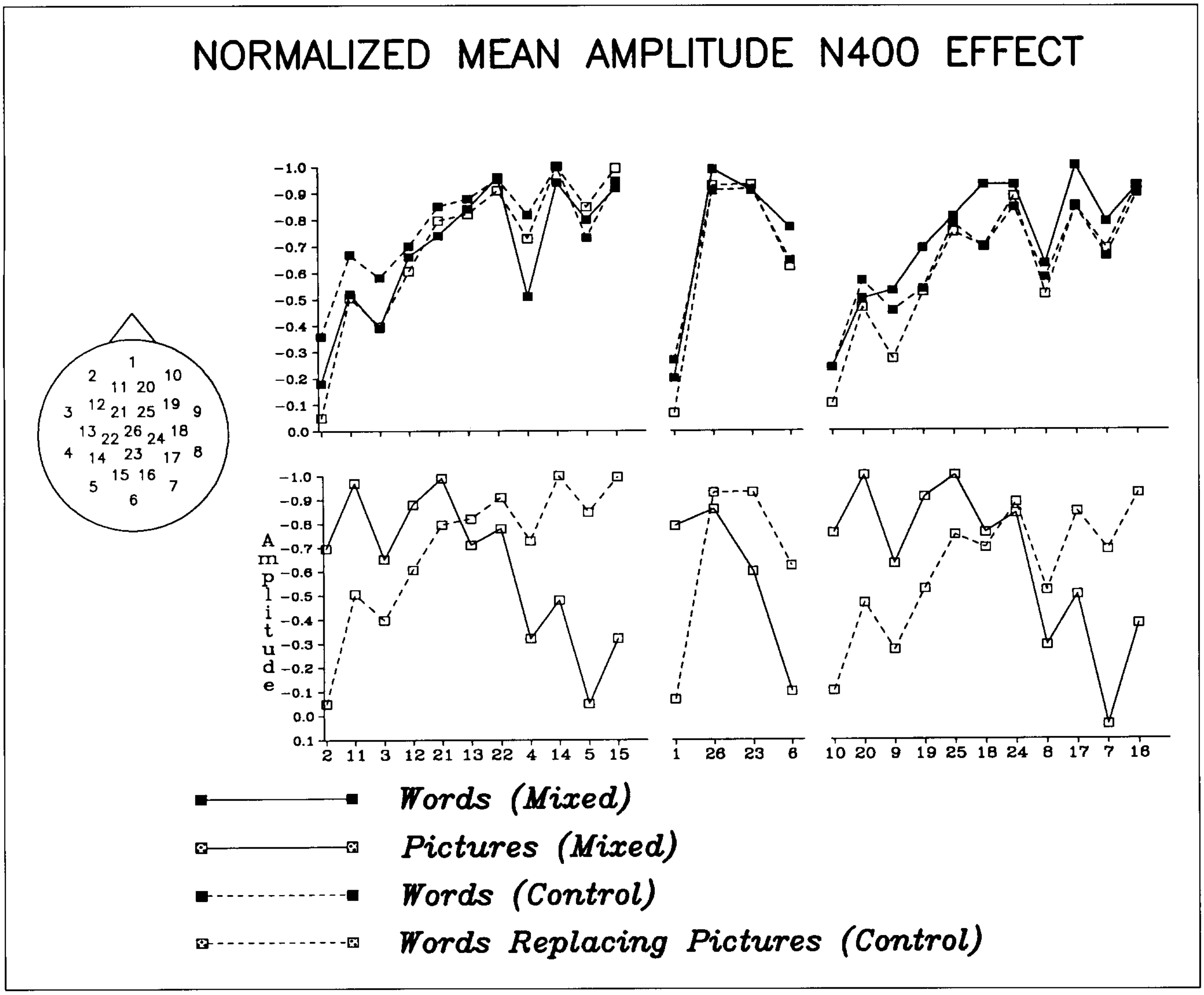

Figure 8. Comparisons between the normalized distribution of the N400 effect (mean amplitude of the difference waves between 325 and 475 msec) obtained in the control experiment (dashed lines) and in the mixed condition (solid lines). In the control experiment the same set of words used in the mixed condition was compared to the set of words obtained by replacing all the pictures in the mixed condition by their names. The top panel shows that the N400 to both sets of words in the control experiment does not differ from that to words in the mixed condition. The bottom panel shows that the difference in N400 distribution for words and pictures is still present when the same set of context sentences is used. See text for details.

patterns), but not to members of the remaining categories, indicating at least some kind of presemantic categorization in such areas.

Blocking had an effect on hemispheric asymmetry; specifically, we found that the N400 congruity effect for pictures in the blocked condition was larger over the right hemisphere. Although Nigam et al. (1992) did not report any significant hemispheric asymmetries, the congruity effect for pictures tended to be bigger over the right hemisphere (Wernicke's sites, see their Fig. 3, p. 18). Their null result could have been due to a lack of statistical power (only two electrodes were compared) and the fact that the electrodes were not sufficiently lateral (30% of interaural distance).

Blocking also affected the timing of the N400 congruity effect. In the mixed presentation mode, the N400 congruity effect for pictures had a shorter peak latency than that for words, suggesting that the semantic analysis of pictures may proceed more quickly. This result seems consistent with several reports that pictures are categorized more quickly than words (e.g., Friedman & Bourne, 1976; Snodgrass & McCullogh, 1986; Caramazza et al., 1990).[10] The latency difference was not significant in the blocked condition, mainly due to a slowing of processing for pictures. The absence of a latency difference in the blocked condition could reflect either real differences in the underlying processes, or greater variability in the peak latency of the N400 under these circumstances. Visual inspection of the ERPs (Fig. 5) reveals that the N400 in fact looked less peaked in the blocked than in the mixed condition, which might thus be a consequence of greater latency variability in the blocked condition.

Overall, the present data contradict the conclusion of Nigam et al. (1992) that there is no reliable difference between the congruity effects for words and pictures in sentences. One possible explanation for this discrepancy is that the between-subjects variability in the study of Nigam et al. was very high. The amplitude of their N400s to words was smaller (especially at occipital sites) than that in the present experiment and in other reports, while the amplitude of their N400s to pictures was comparable to ours. This suggests that, perhaps, the subjects in the two main conditions of the Nigam study were not well-matched. One potential source of unaccounted variance, among others that are known to affect N400 amplitude, may be family history of left-handedness (Kutas & Van Petten, 1990). An indirect example of the problem of between-subject variability can be seen in the present study: the mean amplitude of N400 effect for exactly the same stimuli presented to two different groups of subjects was markedly different: 4.1 μV (mixed condition) versus 2.8 μV (control condition). One way of dealing with this objection in future studies would be to add a control condition that the subjects from all the different groups experience; insofar as the ERPs to the control stimuli do not differentiate the groups, a stronger case can be made for whatever differences are obtained in the data for the between-subject comparisons.

Our study was based on the rationale that similar congruity effects for words and pictures would provide evidence for a common semantic system hypothesis while different congruity effects would provide evidence for a multiple semantic systems hypothesis. As noted in the introduction, this rationale and its variants have often been used in addressing the issue of common versus multiple semantic systems, both in the behavioral and in the ERP literature. Although an in-depth analysis of this rationale goes beyond the purpose of this paper, we want to emphasize that the issue hinges critically on what it is meant by "same semantic system." This is unarguably a difficult issue as reflected by recent animated debates on the topic (e.g., Caramazza et al., 1990; Shallice, 1993). Despite the progress in our understanding of how aspects of the external world are encoded in the brain (see, for example, the very elegant studies by Tanaka, 1993 or Sakai & Miyashita, 1991 on monkey inferotemporal cortices), at present there seems to be a lack of information about the neural basis of knowledge representations, upon which we could start to formulate more refined questions that could be investigated empirically. In this paper we have shown by using ERP evidence that *if* we accept this rationale then the conclusion must be that there are multiple semantic systems.

SUMMARY AND CONCLUSIONS

In summary we found that (1) the time course of the N400 congruity effect for words and pictures was very similar; by contrast, (2) the scalp distributions of the N400 congruity effect for pictures and words differed significantly in that pictures were associated with a larger frontal effect than words, and words were associated with a larger effect than pictures posteriorly; (3) pictures were characterized by an early congruity effect that did not seem to be present for words; and finally (4) the peak latency of the N400 congruity effect was earlier for pictures than for words but only in the mixed presentation mode. According to the rationale that similar congruity effects reflect the existence of a common semantic system, the distributional differences are inconsistent with a strong common semantic system view. However, the similarity in the time course of the word and picture congruity effects also argues against a view that words and pictures are processed in radically different ways within semantic systems with different internal or functional organizations. Our results are most consistent with a view of semantic analysis that holds that the meaning of words and pictures is determined by functionally similar neural systems that are at least partially nonoverlapping.

METHOD

Subjects

Twenty-eight native speakers of English were paid $5.00/hour to participate in the experiment (12 males and 16 females, 18 to 30 years of age, mean 21). All the subjects were right handed with no family history of left handedness. Fourteen subjects were assigned to the blocked modality condition and 14 other subjects to the mixed modality condition. The data from three subjects (two in the mixed condition and one in the blocked condition) had to be discarded because the recordings were contaminated with eye movements. One subject's data (in the blocked condition) were lost because of a disk writing error. All the analyses were performed on the remaining 24 subjects, 12 in each condition.

Materials

A total of 140 experimental sentences was used. The same set of sentences was used in the mixed and in the blocked conditions. They were divided into four sets of 35 sentences each, according to the type of ending: (1) congruous/word, (2) congruous/picture, (3) incongruous/word, (4) incongruous/picture. The cloze probability of the endings was determined by calculating the percentage of individuals who completed the sentence with that ending. Cloze data were collected from a different group of 60 UCSD undergraduates. Each sentence had

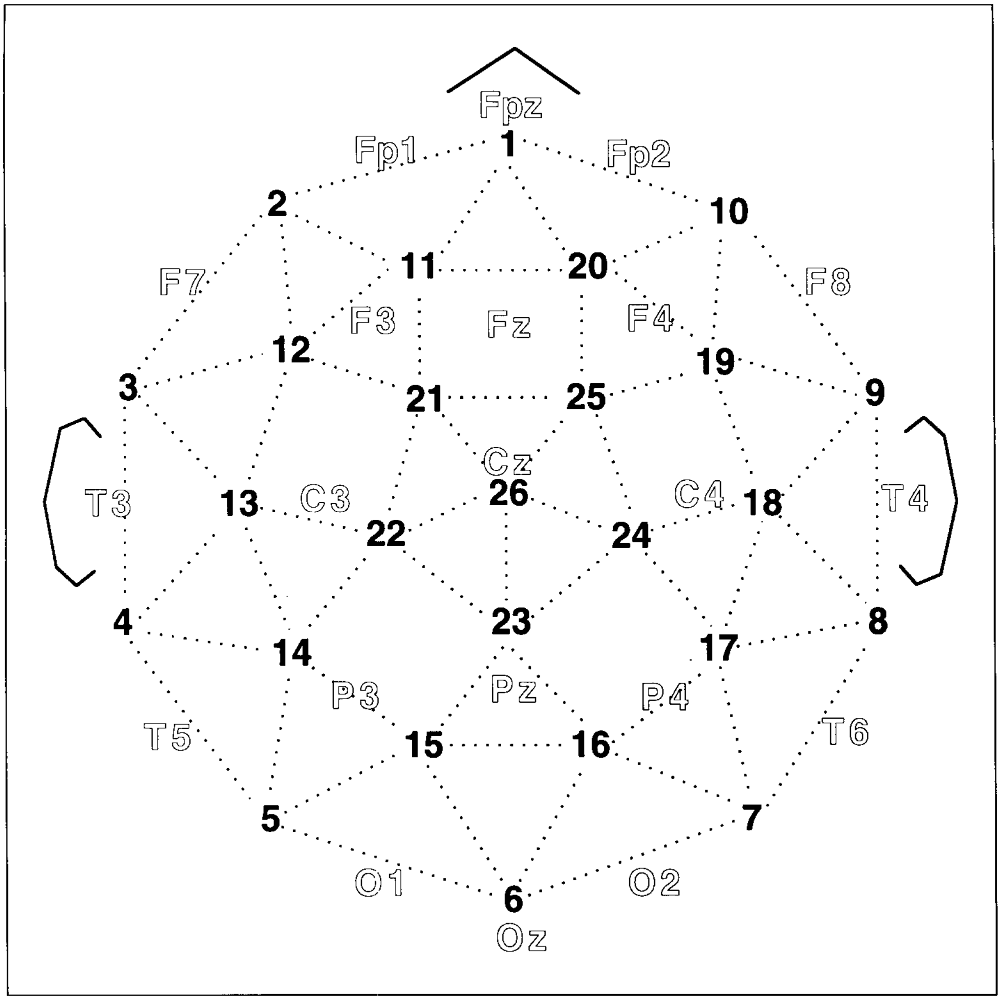

Figure 9. Schematic of the electrode array used in the experiments (thick labels). The electrodes are arranged on a series of four equally spaced concentric rings. The outermost two rings are composed of 10 equally spaced electrodes. The third ring is composed of 5 electrodes, and the last one is composed of a single electrode. The electrode sites corresponding to the International 10–20 system have been superimposed for comparison (thin labels). Note that electrode 1 corresponds exactly to Fpz in the 10–20 system, electrode 6 to Oz, and electrode 26 to Cz.

an ending with a cloze probability greater than 0.7 (mean = 0.85). In half of the sentences, the high cloze ending was replaced with an item that could not be integrated in the context of the sentence (incongruous ending). All the word endings had an imageability score greater than one standard deviation above the average on Paivio's imageability scale. Word endings in the various conditions were matched for frequency of occurrence (median = 37), length (mean length was 5.5), concreteness (mean = 590), and imageability. Picture endings were matched across conditions for familiarity and visual complexity (Snodgrass & Vanderwart, 1980). To minimize early visual processing differences all endings, words, and pictures were matched for average area (the area of the smallest rectangle around each word/picture was always about 9 square deg, while the area subtended by words varied from about 5 to 13 square deg, with an average of 9 square deg, Theios & Amrhein, 1989), and line thickness (a rough matching for spatial frequency). The average length and structural complexity of the sentences also were matched across conditions. Finally, when a picture sentence required a plural ending, two instances of the required concept were depicted to avoid case mismatches in the picture conditions.

Procedure

Each subject was tested in one experimental session lasting about 90 min, divided into two subsessions. The order of the subsessions was counterbalanced across subjects. The subjects were instructed to read the sentences and to try to understand them. They were told there would be a later test on the materials seen during the ERP session. They were also instructed not to blink during the presentation of the sentence. Each session began with a short practice, followed by the experimental sentences. Each trial was preceded by a fixation stimulus (red cross). Sentences were presented one word at a time in the center of the screen (Fig. 1). The duration of each word was proportional to its length, so as to simulate fixation times during normal reading; it varied between 215 and 365 msec/word. The ISI between successive items was 50 msec. The sentence final item was presented for a duration of 1000 msec, to avoid stimulus offset ERPs. The final word was followed by a green sign, after which the subject was allowed to blink. At the end of the ERP session, subjects were administered a recognition test on the endings of the sentences to verify they had paid attention to the stimuli during the ERP session.

Recordings

The EEG was recorded by means of Ag/AgCl electrodes from 26 scalp sites (Fig. 9). This set of electrode sites offers full coverage of the head, allowing more reliable current source density analysis. All these electrodes, as well as the right mastoid electrode, were referred to the left mastoid. All the data were rereferenced off-line to the average of the activity at left and right mastoids. Blinks and vertical eye movements were monitored by an electrode placed on the lower orbital ridge, also referred to the left mastoid. Lateral eye movements were monitored by two electrodes placed lateral to each eye, and the left eye electrode was used as reference. The EEG was amplified by Grass amplifiers with a 0.01 to 100 Hz half-amplitude cutoff bandpass. The sampling rate was 250 Hz. ERPs were averaged off-line for an epoch of 1024 msec, beginning 150 msec before the onset of sentence ending. Trials contaminated by eye movements or amplifier blocking (about 10%) were rejected off-line. Electrode impedances were kept below 5 kΩ.

Acknowledgments

This work was supported by grants NICHD (HD 22614) and NIA (AG 08313) to M. Kutas. Giorgio Ganis was supported by Grant NIH/NIDCD 5 T32 DC00041-01 to CRL (Center for Research in Language) and funds from the San Diego McDonnell-Pew Center for Cognitive Neuroscience. We would like to thank Haline Schendan for helpful discussion as well as for obtaining the line drawings by Snodgrass and Vanderwart and converting them into a usable format. Thanks also to Ron Ohst and Paul Krewski for keeping the machines running and for help with programming. The software used to create the typographic maps was developed by the first author and is freely available. For information on how to obtain a copy, please contact the first author. The software was developed on an SGI workstation (purchased with McDonnell-Pew funds) but runs on many other UNIX platforms.

Requests for reprints should be sent to Giorgio Ganis, Department of Cognitive Science, University of California at San Diego, La Jolla, CA 92093-0515; email: ganis@cogsci.ucsd.edu.

Notes

1. Of course, there are many kinds of iconic symbols [see Kolers and Brisson (1984) for a discussion]. In this paper "iconic symbols" refer to simple line drawings of objects. Since in most of the relevant literature "line drawing" and "picture" are used interchangeably, we will uphold that tradition.
2. On this view, pictures and concrete words would share a semantic system, whereas concrete words would differ from abstract words.
3. These variants build in the necessary differences in visual processing, but maintain that the final product of presemantic visual processes is nonetheless a common semantic representation. Presumably, the mode of entry is lost in this representation.
4. This despite the fact that very few of the studies directly compared the ERP effects to the different stimuli in the same set of subjects.
5. Until recently a similar logic has allowed the various late positivities to be subsumed under the P3 rubric.
6. Work done in collaboration with S. A. Hillyard.
7. Note that all the potentials are relative to the baseline, that is, the average potential during the 150-msec period preceding the onset of the stimulus.
8. We are of course aware that the relationship between scalp distributions and neural generators is not a trivial one and that our suggestion is at the moment speculative.
9. Of course there are some potential problems in drawing parallels between the function of human and nonhuman brain areas based on anatomical location. Although there is some evidence that areas such as V4 and some inferotemporal cortices are located in different places relative to sulcal/gyral landmarks in human and nonhuman primates (e.g., Heywood et al., 1993; Haxby et al., 1993), there is no direct evidence that this is the case for temporal pole areas.
10. Notice that the presence of a strong semantic context, as in our experiment, might produce a ceiling effect, thereby reducing the advantage for pictures. Perhaps the use of a weak semantic context would reveal a larger advantage for pictures.

REFERENCES

Albert, M. L., Reches, A., & Silverberg, R. (1975). Associative visual agnosia without alexia. *Neurology, 25,* 322-326.

Alexander, M. P., Hiltbrunner, B., & Fischer, R. S. (1989). Distributed anatomy of transcortical sensory aphasia. *Archives of Neurology, 46,* 885-892.

Allison, T., Ginter, H., McCarthy, G., Nobre, A. C., Puce, A., Luby, M., & Spencer, D. D. (1994). Face recognition in human extrastriate cortex. *Journal of Neurophysiology, 71*(2), 821-825.

Anderson, J. R., & Bower, G. H. (1973). *Human associative memory.* New York: Wiley.

Bajo, M. (1988). Semantic facilitation with pictures and words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*(4), 579-589.

Barrett, S. E., & Rugg, M. D. (1989). Event-related potentials and the semantic matching of faces. *Neuropsychologia, 27,* 913-922.

Barrett, S. E., & Rugg, M. D. (1990). Event-related potentials and the semantic matching of pictures. *Brain and Cognition, 14,* 201-212.

Barrett, S. E., Rugg, M. D., & Perrett, D. I. (1988). Event-related potentials and the matching of familiar and unfamiliar faces. *Neuropsychologia, 26,* 105-117.

Bentin, S. (1987). Visual word perception and semantic processing: An electrophysiological perspective. *Israel Journal of Medical Sciences, 23,* 138-143.

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

Caramazza, A., Hillis, A. E., Rapp, B. C., & Romani, C. (1990). The multiple semantics hypothesis: Multiple confusions? *Cognitive Neuropsychology, 7*(3), 161-189.

Damasio, A. R., Damasio, H., & Van Hoesen, G. W. (1982).

Prosopoagnosia: Anatomical basis and behavioral mechanisms. *Neurology, 32,* 331-341.
Denes, G., & Semenza, C. (1975). Auditory modality-specific anomia: Evidence from a case of pure word deafness. *Cortex, 11,* 401-411.
Donchin, E., & Coles, M. G. (1991). Behavior, cognition and event-related brain potentials. *Behavioral and Brain Sciences, 14*(4), 735-739.
Duncan-Johnson, C. C., & Donchin, E. (1977). The variation of event-related potentials with subjective probability. *Psychophysiology, 14*(5), 456-467.
Durso, F. T., & Johnson, M. (1979). Facilitation in naming and categorizing repeated pictures and words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 5,* 449-459.
Friedman, A., & Bourne, L. E., Jr. (1976). Encoding the levels of information in pictures and words. *Journal of Experimental Psychology: General, 105,* 169-190.
Geisser, S., & Greenhouse, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95-112.
Gomori, A. J., & Hawryluk, G. A. (1984). Visual agnosia without alexia. *Neurology, 34,* 947-950.
Gross, C. G. (1973). Inferotemporal cortex and vision. *Progress in Psychobiology and Physiological Psychology, 5,* 77-123.
Hart, J., & Gordon, B. (1990). Delineation of single-word semantic comprehension deficits in aphasia, with anatomical correlation. *Annals of Neurology, 27*(3), 227-231.
Hart, J., & Gordon, B. (1992). Neural subsystems for object knowledge. *Nature (London), 359,* 60-64.
Hart, J., Lesser, R. P., & Gordon, B. (1992). Selective interference with the representation of size in the human direct cortical electrical stimulation. *Journal of Cognitive Neuroscience, 4*(4), 337-344.
Hassainia, F., Medina, V., Donadey, A., & Langevin, F. (1994). Scalp potentials and current density mapping with an enhanced spherical spline interpolation. *Medical Progress through Technology,* 20(1-2), 23-30.
Haxby, J. V., Grady, C. L., Horwitz, B., Salerno, J., Ungerleider, L. G., Mishkin, M., & Schapiro, M. B. (1993). Dissociation of object and spatial visual processing pathways in human extrastriate cortex. In B. Gulyas, D. Ottoson, & P. E. Roland (Eds.), *Organisation of the human visual cortex.* Oxford: Pergamon Press.
Heit, G., Smith, M. E., & Halgren, E. (1990). Neuronal activity in the human medial temporal lobe during recognition memory. *Brain, 113,* 1093-1112.
Heywood, C. A., & Cowey, A. (1993). Color and face perception in man and monkey: The missing link. In B. Gulyas, D. Ottoson, & P. E. Roland (Eds.), *Organisation of the human visual cortex.* Oxford: Pergamon Press.
Hillis, A. E, Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology,* 7(3), 191-243.
Holcomb, P. J. (1988). Automatic and attentional processing: An event-related potential analysis of semantic priming. *Brain and Language, 35,* 66-85.
Holcomb, P. J., & McPherson, B. (1994). Event-related brain potentials reflect semantic priming in an object decision task. *Brain and Cognition, 24*(2), 259-276.
Holcomb, P. J., & Neville, H. J. (1990). Semantic priming in visual and auditory lexical decision: A between modality comparison. *Language and Cognitive Processes, 5,* 281-312.
Holcomb, P. J., & Neville, H. J. (1991). The electrophysiology of spoken sentence processing. *Psychobiology, 19,* 286-300.
Humphreys, G. W., & Bruce, V. (1989). *Visual cognition: Computational, experimental and neuropsychological perspectives.* London: Lawrence Erlbaum Associates Ltd.
Karpov, B. A., Meerson, Y. A., & Tonkonough, I. M. (1979). On some peculiarities of the visuomotor system in visual agnosia. *Neuropsychologia, 17,* 281-294.
Katayama, J., & Yagi, A. (1992). Negative brain potentials elicited by an unexpected color patch or word. *Electroencephalography and Clinical Neurophysiology, 83,* 248-253.
Kolers, P. A., & Brisson, S. J. (1984). Commentary: On pictures, words, and their mental representation. *Journal of Verbal Learning and Verbal Behavior, 23,* 105-113.
Kosslyn, S. M., & Pomerantz, J. R. (1977). Imagery, propositions and the form of internal representations. *Cognitive Psychology, 9,* 52-76.
Kroll, J. F. (1990). Recognizing words and pictures in sentence contexts: A test of lexical modularity. *Journal of Experimental Psychology, 16*(5), 747-759.
Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical object and reality decisions. *Journal of Verbal Learning and Verbal Behavior, 23,* 39-66.
Kutas, M., & Hillyard, S. A. (1980). Event-related potentials to semantically inappropriate and surprisingly large words. *Biological Psychology, 11,* 99-116.
Kutas, M., & Hillyard, S. A. (1983). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203-205.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature (London), 307*(5954), 161-163.
Kutas, M., Neville, H. J., & Holcomb, P. J. (1987). A preliminary comparison of the N400 response to semantic anomalies during reading, listening and signing. In R. J. Ellingson, N. M. F. Murray, & A. M. Halliday (Eds.), *The London symposia.* Amsterdam: Elsevier Science Publishers, B.V.
Kutas, M., & Van Petten, C. (1990). Electrophysiological perspectives on comprehending written language. In P. M. Rossini & F. Mauguiere (Eds.), *New trends and advanced techniques in clinical neurophysiology.* Amsterdam: Elsevier Science Publishers, B.V.
Levine, D., & Calvanio, R. (1989). Prosopoagnosia: A defect in visual configural processing. *Brain and Cognition, 10,* 149-170.
Lhermitte, F., & Beauvois, M. F. (1973). A visual-speech disconnection syndrome: Report of a case with optic-aphasia, agnosic alexia, and color agnosia. *Brain, 96,* 695-714.
Luders, H., Lesser, R. P., Hahn, J., Dinner, D. S., Morris, H. H., Wyllie, E., & Godoy, J. (1991). Basal temporal language area. *Brain, 114,* 743-754.
Mack, J. L., & Boller, F. (1977). Associative visual agnosia and its related deficits: The role of the minor hemisphere in assigning meaning to visual perceptions. *Neuropsychologia, 15,* 345-349.
Marin, O. S. M. (1987). Dementia and visual agnosia. In G. W. Humphreys & M. J. Riddoch (Eds.), *Visual object processing: A cognitive neuropsychological approach.* London: Lawrence Erlbaum Associates Ltd.
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, R. A., & Warrington, E. K. (1988). Evidence for modality-specific meaning systems in the brain. *Nature (London), 334,* 428-430.
McCarthy, R. A., & Warrington, E. K. (1990). The dissolution of semantics. *Nature (London), 343,* 599.
McCarthy, G., & Wood, C. (1985). Scalp distributions of event-related potentials: An ambiguity associated with analysis of

variance models. *Electroencephalography and Clinical Neurophysiology, 62,* 203-208.

McCarthy, G., Nobre, A. C., Bentin, S., & Spencer, D. D. (1995). Language-related field potentials in the anterior-medial temporal lobe. Intercranial distribution and neural generators. *Journal of Neuroscience, 15*(2), 1080-1089.

Mikami, A., Nakamura, K., & Kubota, K. (1993). Comparisons of single neuronal activities in the temporal cortex of the amygdala of the Rhesus monkey during visual discrimination and memory task. In T. Ono, L. R. Squire, M. E. Raichle, D. I. Perrett, & M. Fukuda (Eds.), *Brain mechanisms of perception and memory.* Oxford: Oxford University Press.

Mishkin, M. (1982). A memory system in the monkey. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 298*(1089), 83-95.

Morton, J., & Patterson, K. E. (1980). A new attempt at an interpretation, or, an attempt at a new interpretation. In M. Coltheart, K. E. Patterson, & J. C. Marshall (Eds.), *Deep dislexia.* London: Routledge and Kegan Paul.

Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory, 3,* 485-497.

Nigam, A., Hoffman, J. E., & Simons, R. F. (1992). N400 to semantically anomalous pictures and words. *Journal of Cognitive Neuroscience, 4,* 15-22.

Nobre, A. C., Allison, T., & McCarthy, G. (1994). Word recognition in the human inferior temporal lobe. *Nature (London), 372,* 260-262.

Onofrj, M. C., Ghilardi, M. F., Fulgente, T., Nobilio, D., Bazzano, S., Ferracci, F., & Malatesta, G. (1990). Mapping of event-related potentials to auditory and visual odd-ball paradigms. *Electroencephalography and Clinical Neurophysiology. Supplement, 41,* 183-201.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Paivio, A. (1983). The empirical case for a dual coding. In J. C. Yuille (Ed.), *Imagery, memory and cognition: Essays in honor of Allan Paivio* (pp. 307-332). Hillsdale, NJ: Erlbaum.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Paller, K. A., McCarthy, G., & Wood, C. C. (1992). Event-related potentials elicited by deviant endings to melodies. *Psychophysiology, 29*(2), 202-206.

Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature (London), 253,* 437-438.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115,* 281-294.

Pylyshyn, Z. W. (1980). Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences, 3,* 111-132.

Ratcliff, G., & Newcombe, F. (1982). Object recognition: Some deductions from the clinical evidence. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions.* New York: Academic Press.

Riddoch, M. J., Humphreys, G. W., Coltheart, M., & Funnel, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. *Cognitive Neuropsychology, 5*(1) 3-25.

Roediger, H. L., & Blaxton, T. A. (1987). Effects of varying modality, surface features, and retention interval on priming in word-fragment completion. *Memory and Cognition, 15,* 379-388.

Rubens, A. B., & Kertesz, A. (1983). The localization of lesions in transcortical aphasias. In A. Kertesz (Ed.), *Localization in Neuropsychology* (pp. 245-268). New York: Academic Press.

Sakai, K., & Miyashita, Y. (1991). Neural organization for the long- term memory of paired associates. *Nature, 354,* 152-155.

Scarborough, D. L., Gerard, L., & Cortese, C. (1979). Accessing lexical memory: The transfer of word repetition effects across task and modality. *Memory and Cognition, 7,* 3-12.

Sereno, M. I. (1991a). Language and the primate brain. *Proceedings of the 13th Annual Conference of the Cognitive Science Society,* (pp. 79-84), Hillsdale, NJ: Lawrence Erlbaum Associates.

Sereno, M. I. (1991b). Four analogies between biological and cultural/linguistic evolution. *Journal of Theoretical Biology, 151,* 467-507.

Seymour, A. (1973). A model for reading, naming and comparison. *British Journal of Psychology, 64,* 35-49.

Shallice, T. (1987). Impairments in semantic processing: Multiple dissociations. In M. Coltheart, R. Job, & G. Sartori (Eds.), *The cognitive neuropsychology of language.* London: Lawrence Erlbaum Associates Ltd.

Shallice, T. (1993). Multiple semantics: Whose confusions. *Cognitive Neuropsychology, 10*(3), 251-261.

Sirigu, A., Duhamel, J. R., & Poncet, M.(1991). The role of sensorimotor experience in object recognition. A case of multimodal agnosia. *Brain, 115,* 645-665.

Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior, 23,* 3-22.

Snodgrass, J. G., & McCullogh, B. (1986). The role of visual similarity in picture categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition, 12*(1), 147-154.

Snodgrass, J. G., & Vanderwart, M. A. (1980). Standardized set of 260 pictures: Norms for name agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174-215.

Sperber, R. D., McCauley, C., Ragain, R., & Weil, C. M. (1979). Semantic priming effects on picture and word processing. *Memory and Cognition, 7,* 339-345.

Striano, S., Grossi, D., Chiacchio, P., & Fels, A. (1981). Bilateral lesion of the occipital lobes. *Acta Neurologica (Napoli), 36,* 690-694.

Tanaka, K. (1993). Neuronal mechanisms of object recognition. *Science, 262,* 685-688.

Theios, J., & Amrhein, P. C. (1989). Theoretical analysis of the cognitive processes of lexical and pictorial stimuli: Reading, naming, and visual and conceptual comparisons. *Psychological Review, 96,* 5-24.

Vanderwart, M. (1984). Priming by pictures in lexical decision. *Journal of Verbal Learning and Verbal Behavior, 23,* 67-83.

Verleger, R. (1990). P3-evoking wrong notes: unexpected, awaited, or arousing? *International Journal of Neuroscience, 142*(4), 225-244.

Vignolo, L. A. (1989) Non-verbal conceptual impairment in aphasia. In F. Boller & J. Grafman (Eds.), *Handbook of Neuropsychology,* (Vol. 2), Amsterdam: Elsevier.

Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology, 27,* 635-657.

Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain, 102,* 43-63.

Warrington, E. K., & Shallice, T. (1984). Category specific semantic impairments. *Brain, 107,* 829-854.

Weldon, M. S., & Roediger, H. L. (1987). Altering retrievial demands reverses the picture superiority effect. *Memory and Cognition, 15,* 269-280.