Order Number 9218237

# Event-related potentials and semantic processing in Alzheimer's patients and normal controls

Hamberger, Marla Jill, Ph.D.

City University of New York, 1992



U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

# EVENT-RELATED POTENTIALS AND SEMANTIC PROCESSING IN ALZHEIMER'S PATIENTS AND NORMAL CONTROLS

BY

MARLA J. HAMBERGER

A dissertation submitted to the Graduate Faculty in Psychology in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York.

1992

This manuscript has been read and accepted for the Graduate Faculty in Psychology in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

1/10/92
Date

Chair of Examining Committee

1/10/92
Date

Herbert D. Saltzstein
Executive Officer

Jeffrey Rosen, Ph.D
David Friedman, Ph.D.
Walter Ritter, Ph.D.
Yaakov Stern Ph.D.
David Masur, Ph.D.

Supervisory Committee

The City University of New York

Abstract


EVENT-RELATED POTENTIALS AND SEMANTIC PROCESSING IN ALZHEIMER'S PATIENTS AND NORMAL CONTROLS

by

Marla J. Hamberger

Advisor: Jeffrey Rosen

Based on behavioral measures generated by young adults during semantic processing tasks, semantic memory has been conceptualized as a network organized according to semantic relationships. Additionally, in young adults, the N400 component of the event-related brain potential (ERP) is also responsive to semantic relationships. Under certain conditions, N400 amplitude varies inversely with the extent to which a word has been primed by its preceding semantic context, reflecting a gradient of semantic relatedness within the semantic network. Based on their behavioral performance on a wide range of tasks, patients with Probable Alzheimer's disease (PAD) show a collapsing of the semantic gradient, such that items within a category lose their distinction, whereas superordinate information remains relatively intact. Given that N400 provides an index of the gradient of semantic relatedness in normal young adults, and that according to behavioral measures, PAD patients show a collapsing of this gradient, the present study examined whether this collapse would also be reflected in the N400 pattern of PAD patients.

Ten normal young adults, 10 normal elderly, and 6 "mild" PAD patients made speeded (but accurate) sense/nonsense decisions to the final word of a series of highly constrained sentence contexts in which final words

belonged to one of four types: 1) best completion/sensical, 2) related to best completion/sensical, 3) related to best completion/nonsensical and 4) unrelated to best completion/nonsensical. Behavioral and ERP responses were recorded only to terminal words.

In contrast to the behavioral data, PAD patients demonstrated an orderly gradient as a function of semantic relatedness (i.e., N400 was smallest to best completions, and largest to unrelated/nonsensical stimuli). However, as predicted, they committed a disproportionate number of errors (i.e., "sense" responses) to nonsensical sentences terminating in words related to the sentences' best completions. It was speculated that this disruption in semantic processing occurs at some point between the elicitation of N400 and the emission of a behavioral response. Although normal elderly showed an unexpected N400 pattern, this appeared to be associated with strategic processing superimposed upon an intact semantic network.

# ACKNOWLEDGEMENTS

Although I refer to this work as *my* dissertation, there are many people to whom I am indebted for their contributions toward the completion of this project.

First, my deepest thanks to my advisor, Jeffrey Rosen, for his astute guidance from the inception through the completion of this dissertation, for the time and energy he has devoted toward my education, and for his faith in my abilities. He has been a source of endlessly surprising insights, challenge and inspiration, not only during the dissertation process, but throughout my graduate years. His creative, philosophical, yet practical approach to science will continue to influence my thinking and my work.

I thank David Friedman for sharing his expertise in all components of cognitive ERP research and for allowing me to conduct this project in his laboratory. I am also grateful for his daily guidance throughout the study, his careful editing of the dissertation, his patient encouragement, and his empathy.

In his introductory course, it was impossible not to become infected with Walter Ritter's fascination with cognitive ERPs. Had it not been for his exciting introduction to this field, this investigation might never have been conceived. I thank him for his generosity with his time, skill and knowledge, for his insightful comments and challenging questions.

I wish also to thank my outside readers, Yaakov Stern, for permitting me to recruit patients from the Memory Disorders Clinic, and for his useful questions and comments, and David Masur for sharing his knowledge and time during the initial development of this investigation, and again at the defense.

Throughout my graduate career, and especially during the course of this dissertation, William Dince has been a source of comfort, encouragement, ideas, direction and answers. As I have told him on occasion, if it were not for him, I might still be studying for my second doctoral examination.

A special thanks to Charles L. Brown III, not only for his brilliant computer programming and technical assistance, but also for his many extra hours, his loyal friendship, and for sharing the everyday excitement and frustrations of this project.

I am grateful to Margaret Cramer and her daughters, Tricia and Ina, and to my sister, Karen, for creating many of the stimuli. I thank my friends and colleagues Christopher Christodoulou and Hilary Gomes, who shared in every phase of this dissertation, and who provided ongoing support in more ways than I can possibly mention.

I would also like thank Celia Klin, for her empathy, encouragement, wisdom, bitnets and care packages, Charlotte Trott, for her listening and her optimism. And a special thanks to John, who has lived with this dissertation for as long as I have, for the laundry, dinners, patience, encouragement, warmth, and distractions.

# TABLE OF CONTENTS

# LIST OF TABLES

Table Page

# LIST OF FIGURES

# I. INTRODUCTION

## A. Statement of the Problem

Semantic memory encompasses the meanings of words, concepts and the relationships among them. Although the organizational and functional properties of this system are presently controversial, there is substantial evidence that semantic memory is associatively structured such that concepts are functionally organized along gradients of semantic relatedness. When this system undergoes degeneration, these gradients collapse in a predictable manner.

This notion of a gradient comes out of findings in several different research domains. Neuropsychological study of brain injured patients with selective language impairments (i.e., aphasias) provide a window into the component organization and processes of semantic memory (Benson, 1979). Deterioration due to brain disease demonstrates the unraveling of semantic functioning which was initially intact. Alzheimer's Disease patients show a degenerative process of this sort (Martin and Fedio, 1983). Cross-cultural investigations reveal both the commonalities across as well as the possibilities for the organization of meaning systems (Lakeoff, 1987). Experiments with normal young adults permit observation of an intact semantic system under a variety of carefully manipulated conditions (e.g., task variable, stimulus variables; --Collins and Loftus, 1975; Glass and Holyoak, 1974; McCloskey and Glucksberg, 1979; Rosch,1981). Typical procedures consist of word priming tasks in which a semantic context is provided by either a word or a sentence, and response patterns vary as a function of the relationship between prior context and the response eliciting stimuli. Results from these paradigms have been applied toward the construction of models of semantic organization.

In each of these research strategies, the dependent variables are typically behavioral measures such as written or verbal responses, error patterns, and/or response latencies. Within the past decade however, it has been demonstrated that scalp-recorded cognitive event-related potentials (ERPs) are responsive to semantic processes. Specifically, the amplitude of the N400 component (a negative deflection peaking at approximately 400 milliseconds post stimulus) varies inversely with the extent to which a stimulus (i.e., word) has been primed by its prior semantic context. An advantage of obtaining electrophysiological measures is that ERPs are not contingent upon overt responding, and can be recorded prior to, during, and subsequent to the emission of an overt response. The concurrent recording of behavioral and electrophysiological measures allow for a rich analysis of the nature, sequence, and timing of cognitive events and processes.

There has been little attempt to integrate the information provided by each of these domains of study mentioned above. The present thesis combined a subset of these strategies. The approach of the present study involved the examination of a cognitive system as it breaks down. The population selected was a group of "mild" Alzheimer's patients, because they embody the early stages of degeneration of a semantic system that was at one time, fully developed and intact. On tasks involving semantic processing, patients in the early stages of Alzheimer's type dementia show systematic and reliable patterns of deficits (Nebes, 1989). Thus far, the measures employed have been strictly behavioral.

In normals, the relationships between semantic process and behavioral measures have been extensively studied (Chang, 1986). Similarly, the correspondence between ERP components and semantic processes have recently received much attention (Kutas, 1988). The combination of electrophysiological and behavioral measures during semantic tasks has also been investigated (Hamberger and Friedman, submitted). Given this data base in which a set of relationships between semantic parameters, behavioral responses and electrophysiological measures have been established in a system which is intact, the present study was designed to investigate these relationships in a system undergoing early stages of degeneration. The questions addressed involve where in processing the impairment occurs, what it looks like, and what are its implications for the functional organization of semantic memory.

B. Semantic Priming

Laboratory experiments support the intuitive notion that the recognition and/or interpretation of a word is influenced by its prior semantic context. It has been reliably demonstrated under a variety of conditions, that words which are preceded by semantically related words are recognized faster and more accurately compared to words preceded by unrelated words. For instance, in a lexical decision task which requires subjects to quickly determine whether a letter string is a word or a nonword, reaction time to decide that "doctor" is a word will be faster when it is presented subsequent to a "related" word such as "nurse" than when it follows an unrelated word such as "butter" (Meyer, Schvaneveldt and Ruddy, 1974). Similarly, tachistoscope recognition time (Morton, 1969) and the time to pronounce a word aloud is faster when preceded by a semantic associate (Meyer, Schvaneveldt and Ruddy, 1974; DeGroot, 1983). These effects can be robust (e.g., 50 milliseconds) even when the "prime" (i.e., the first word presented) is pattern masked to prevent conscious perception (DeGroot, 1983; Marcel, 1985; Fowler, Wolford, Slade and Tassinary, 1981). These facilitory effects on naming and lexical decision latencies also occur when the preceding context consists of a sentence frame (Fischler and Bloom, 1979, Kleiman, 1980) (e.g., The sky turned dark and it started to - RAIN), regardless of whether the context is presented in the visual (Schwantes, 1985) or auditory (Kintch and Mross, 1985) modality. A century ago, William James wrote, "Each word is doubly awakened; once from without by the lips of the talker, but already before that from within by the premonitory processes irradiating from the previous words..."(1890, vol I, p.450) Although the terminology has changed, this phenomenon still receives much attention, and is currently known as "semantic priming".

When a normal adult produces the first word that comes to mind in response to a stimulus word (e.g., "dog"), s/he will most likely name a paradigmatic associate (i.e., another member of the same semantic category; e.g., "cat") or some other semantic associate (e.g., "animal", "mammal", etc.), (Nelson, 1986). Several explanations could account for the conscious use of semantic context (i.e., the stimulus word) and the ways in which it might influence an individual's response. However, as described above, semantic context also influences performance 1) when semantic processing is not necessary for task completion, 2) when semantic context is irrelevant to the task, and 3) when semantic context is not consciously perceived. These results have been interpreted as reflecting the organization of the mental representation of word meanings. These data provide strong evidence that semantic memory is to some extent, associatively structured (i.e., semantically related concepts , e.g., "cat" and "dog", are functionally proximal, compared to unrelated items, e.g., "cat" and "hammer", which are functionally distant)

Consequently, a variety of models were created to represent the organization of semantic memory, and the processes by which it operates. The utility of a model is that it attempts to make explicit the structural properties and processing assumptions of a theory, and provides a framework in which to integrate a set of relevant observations. The model can then be evaluated for its integrative capacity and its ability to make predictions under a variety of conditions.

The present study investigated semantic priming in a population in which the semantic system has begun to break down. The background and rationale for this investigation will be presented as follows: 1) The domain of semantic memory for the current investigation will be defined, 2) The methodologies and findings in the literature will be presented, 3) Several models of semantic memory will be described to provide a context for the ways in which a normally functioning semantic system has been conceptualized, 4) Speculations will be made regarding the nature of semantic impairment as inferred from these models, 5) Behavioral data from Alzheimer's patients in which the deterioration of the semantic system is manifest will be presented, 6) The use of the N400 component of the event-related potential (ERP) in studies of normal semantic functioning will be described, and 7) These pieces will be integrated in the rationale for the current investigation in which electrophysiological and behavioral measures were recorded from Alzheimer's patients and normal controls during a semantic priming task.

C. Semantic Memory: Experimental Literature in Normal Adult Subjects

A major issue in the semantic memory literature concerns the scope of its domain. In its broadest sense, semantic memory encompasses the store for all meaningful material, including knowledge of word meanings, as well as transient memories of sentences encountered in everyday experience (Kintch, 1974). In a slightly narrower sense, Tulving (1985) proposed a hierarchical classification scheme in which memory is decomposed into three separable subsystems. Briefly, "procedural memory" affords the retention of complex stimulus response patterns, "semantic memory" is characterized by the internal representation of the world that is not perceptually present, and "episodic memory" refers to the more classical notion of memory in that it encompasses information about personally experienced events and their temporal relations in subjective time. The most restrictive definition limits semantic memory to a system containing knowledge about the language, specifically, the meanings of words, relations, and rules for operating on them (Smith, 1982). The present investigation operates within the latter characterization of semantic memory, with the understanding that the information contained within this system is necessarily highly integrated with these other types of information.

Historically, models of semantic memory arose from studies of semantic categorization, investigated principally via the speeded verification and falsification of sentences. Like the semantic priming paradigms (e.g., lexical decision tasks), the variable of interest is reaction time rather than error rate, as error rates are exceptionally low (e.g., less than 5%), and more likely reflect misreadings rather than memory failure (Chang, 1986). Sentences are typically of the form, "An x is a y", with the associative relationship between subject and predicate varied along dimensions of relatedness. The sentences involve ordinary, commonplace knowledge, so there is little difference among sentences in either truth value or difficulty level. For instance, "A robin is a bird" and "A robin is an animal" are equally true. Nonetheless, the former is consistently verified approximately 50 milliseconds (msec) faster than the latter (Chang'1986). Similarly, reaction time for false sentences vary with respect to the relationship of subject and object, e.g., "All chairs are bats" is falsified faster than "all birds are bats". (These effects will be addressed below)

C1. Findings

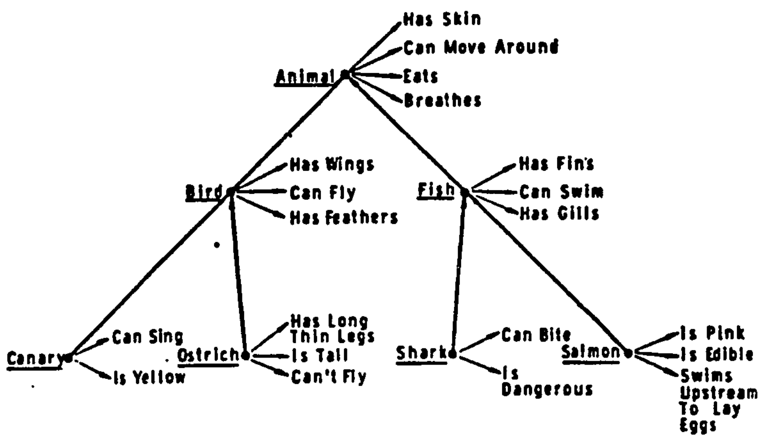
Two critical findings from sentence verification procedures pertain to typicality for "true" sentences, and relatedness for "false" sentences, (although it is likely that these two effects are themselves related). Typicality refers to the extent to which an exemplar is considered "typical" or representative of a particular category (Rosch, 1973). Categorical structure has been conceptualized along a gradient of centrality in which typical members are more central than atypical members (Rosch,1975, 1981, Lakeoff, 1987). Typicality has been assessed by rating procedures (McCloskey and Glucksberg, 1979; Rosch and Mervis, 1975), in which subjects rate exemplar-category pairs along dimensions of "similarity", "relatedness" or "goodness of example", and by production frequency measures (Holyoak and Glass, 1975), in which subjects produce exemplars in response to category names. For true sentences, "typical" category members are verified more quickly than atypical members. Since "robin" is considered to be a more typical "bird" than is "chicken", "A robin is a bird" is verified faster than "A chicken is a bird". The term, "relatedness" has been used somewhat inconsistently in the literature. For instance, relatedness has been operationalized as, "associatedness " (Holyoak and Glass, 1975), "similarity of meaning" (McCloskey and Glucksberg, 1979), and as "the ease with which concepts can be mentally transformed into one another" (Smith, 1978). These different conceptualizations of relatedness have produced dissimilar results, e.g., black and white are highly associated, although black and dark are more similar in meaning. Nevertheless, when applied specifically to the falsification of sentences, the relatedness effect cuts across most of these discrepancies, and refers to an effect in which responses are slower when subject and predicate are in some sense related. For example, "A bat is a bird" is disconfirmed slower than "A bat is a chair".

C2. Models of Fully Developed Semantic Memory

The individual models differ in their fundamental assumptions concerning the content, structure and processes encompassed in semantic memory, and therefore, in how they account for these findings. Most models of semantic memory deal primarily with nouns, thereby, neglecting most other parts of speech. Additionally, since they are based on performance of young healthy college students, they represent a fully developed, intact semantic system. There is little or no account within the theories themselves for the development of semantic representation, or for flexibility and change over time. Most models can be classified as either "computational" or "network" types. These types are distinct in that network models prestore semantic information (i.e. relations among concepts) as a critical component of the model's structure, whereas computational models derive this information through computational comparison processes.

In a typical computational model, each word is represented by a set of attributes. For instance, in the "feature comparison model" (Smith et al., 1974), each attribute is weighted as a function of the degree to which it defines category membership. Some attributes are "defining" whereas others are only "characteristic". For example, "animate" would be considered a defining attribute of "bird", and therefore weighted more heavily than "feathered", which is only a "characteristic" attribute of "bird". Items possessing both defining and characteristic attributes are likely to be typical category members (e.g., "robin" for bird) whereas items with only defining attributes are likely to be less typical members (e.g., ostrich or chicken for bird). Also, items which possess common attributes will be more similar or *related* to each other (e.g., chair and bench compared to chair and table). Hence, the quantity and quality (i.e., characteristic or defining) of attributes and their relative weights underlie relatedness and centrality gradients within categories. During sentence verification, attributes of the subject and predicate undergo a comparison process, which yields either positive (true) or negative (false) evidence. The process terminates when a positive or negative estimate exceeds some criterial value. Since typical members possess relatively more defining and characteristic attributes, positive evidence for typical members accumulates more rapidly compared to atypical members, and are therefore verified quickly.

Unlike computational models, the relational properties between words in network theories are inherent in their structure (see figure 1). The best known model of this type is Collins and Quillian's (1972) hierarchically organized network model, later expanded upon by Collins and Loftus (1975). In contrast to computational models, the network model was not developed in order to explain experimental data, rather, it was designed as a model for incorporating human semantic structure and processes into a computer. Thus, it was intended to address more general applications of semantic processing.




## Figure 1a

A memory representation according to Collins and Quillian (1969) heirarchical network model. (From "Retrieval times from semantic memory", by A.M. Collins and M.R. Quillian, 1969, Journal of Verbal Learning and Verbal Behavior, 8, p. 241. Copyright 1969 by Academic Press. Reprinted by permission.)

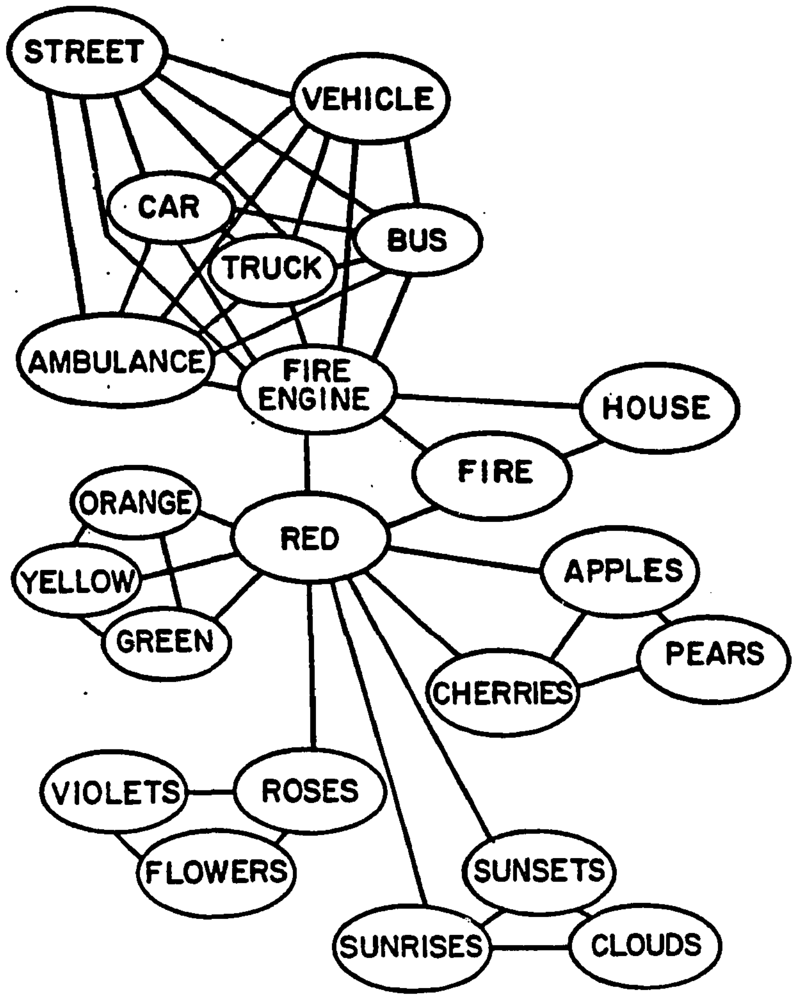



**Figure 1b**

A schematic representation of semantic relatedness. A shorter line represents greater relatedness. (From " A spreading activation theory of semantic processing", by A.M. Collins and E.F. Loftus, 1975, Psychological Review, 82, p. 412. Copyright 1975 by the American Psychological Association. Reprinted by permission.)

The fundamental elements of this particular model (Collins and Loftus, 1975), are "nodes", which represent words/concepts, "links" which are the connections between nodes, and the hierarchically organized network, which maps out its structure. The fundamental processes are search, which occurs along the links from node to node, and activation, which is initiated by the processing of a particular node (via reading, hearing, rehearsing, etc...) and spreads first to nodes directly connected to the original node, and then to all nodes connected to these nodes etc,... in a decreasing gradient. Recognition occurs when activation exceeds some unspecified "firing threshold". The strength of this activation is attenuated both by time and by intervening nodes. However the breadth of this "spreading activation" is not specified within the theory.

Attribute information about concepts is not represented within the concept nodes themselves, i.e., nodes have no internal structure or content. Instead, attributes are encoded by links to other nodes. Two types of relations are distinguished: 1) subset relations which are encoded within the hierarchical structure, labeled "is a" (e.g, a robin is a bird), and 2) attribute relations (e.g., a robin "has wings", "can fly", etc...). These relations are represented directly (i.e., they are not computed). Hence, the full meaning of any concept is represented by the entire complex of labeled relations and their connections leading to and from a node in the network. (The boundaries of this complex are not specified). The hierarchical structure mandates that a word is more proximal to an immediate superordinate (e.g., robin-bird) than to a more distant superordinate (e.g., robin-animal, see figure 1a). In relating a word to a distant superordinate, the intermediate link(s) must be traversed. This occurs in real time, thus, partially accounting for the faster verification time for "a robin is a bird" compared to "a robin is an animal." More important than its hierarchical structure (as stated by Collins and Loftus, 1975), is that the network is organized according to semantic similarity in that the more properties two concepts have in common, the more links there are between their two nodes via these properties, and therefore, the two concepts are more closely related. For example, "fire engine" is related to both "truck" and "red", but shares relatively more interconnections with "truck", and is therefore more closely related to "truck", (see figure 1b). This notion of aggregates and clustering of interlinkages underlies gradients of relatedness. Additionally, the strength or "criteriality" of a link, which is not necessarily symmetrical, encodes how essential a particular link is to the meaning of a concept. For example, it is more criterial that a butterfly is an insect than it is that an instance of an insect is a butterfly. Thus, the number, type, and criteriality of links all contribute to typicality gradients. For instance, "robin" a highly typical bird possesses many criterial properties for the category "bird" whereas "ostrich", an atypical bird possesses fewer criterial attributes.

Semantic priming has been explained in terms of spreading activation within a semantic network (e.g., DeGroot, 1983; Lupker, 1984; Neely, 1976). When a word is presented, activation spreads outward from its corresponding node along its links to other nodes. The activation "tags" these related nodes as well as the links on which it has spread. The theory offers two possible mechanisms to account for semantic priming: Upon presentation of a related concept, the same spreading activation processes is initiated, except this second concept is already in an activated state, and is therefore recognized faster than if it had not been primed by a related concept. Alternatively, it is possible that the second concept has not yet been activated by the spreading activation originating from the first concept. Since the second concept is in some way related to the first, the activation spreading from the second concept will intersect with that of the first. In both cases, the summation of activation that exceeds the firing threshold for recognition occurs more rapidly due to the activation of the initial concept. The magnitude of priming (i.e., how much faster a word will be recognized, a sentence will be verified, etc.) is determined by the number of common links shared by the two concepts, the strength (criteriality) of those links, the number of intervening nodes between the two concepts, and the amount of time that has lapsed between the presentation of the first and second concept.

Spreading activation theory has been applied to single word priming as in lexical decision tasks, and to priming among concepts with sentences. For sentence verification tasks, an additional criterion mechanism is proposed, which accumulates positive and negative evidence. Concepts sharing many attribute links of high criterialities generate strong positive evidence (i.e., true) whereas the absence of common links would provide negative evidence (i.e., false). Thus, "A robin is a bird" would be verified faster than "A robin is a house" due to the activation of related concepts and, the accumulation of positive evidence. Also, since "bat" and "bird" share many similar attributes and are therefore related, "A bat is a bird" would generate positive evidence, which would delay falsification. This phenomenon is referred to as the "false-relatedness" effect (Kintch, 1980).

Semantic priming effects are best understood in the context of a semantic network and spreading activation theory. However, while the theory accounts for a number of findings, it lacks precision. For instance, 1) the description of a node is vague i.e., a node lacks internal structure, yet, the meaning of a concept is determined by a single node's connections with other nodes. 2) When a concept is primed, "activation tags are spread by tracing an expanding set of links in the network out to some unspecified depth" (Collins and Loftus, 1975), and 3). There are no limits or rules specifying the types of links that are possible or how they are differentially represented. Another problem for most semantic memory models is their unidirectionality, i.e., an external stimulus initiates internal processing (e.g., spreading activation, feature comparison processes), terminating with a decision and typically, an overt response. Human beings do not function in this linear fashion. Rather, cognitive processing involves a dynamic interaction between internal processing systems and the external environment. This interaction holds implications for mechanisms of learning and change. Associations between words/concepts are necessarily learned. Thus, a comprehensive theory of semantic memory should make sense at the level of on line processing, and at a broader level, capable of stability, consistency, and relatively slower change over time as a function of development and experience.

Although principally a model of online processing during word recognition, Grossberg and Stone's (1986) Adaptive Resonance Theory incorporates principles of development and change, and interactions between stimuli, working memory and long term verbal memory. In addition, it provides a more detailed description of its fundamental processing assumptions and basic elements. Thus, the theory is mentioned here because it provides a more detailed account of semantic representation and relevant processing mechanisms, and perhaps, the incorporation of some of its principles into the spreading activation theory would provide a more plausible description of semantic organization.

The model is described at two levels: 1) the microtheory, which pertains to online processing at the level of working memory, and 2) the macrotheory, which describes language related visual, verbal and motor representations. At both levels, the most fundamental meaningful unit consists of a pattern of activation across nodes in a network. Each of these two levels are further decomposed into a series of substages, which are defined by circular reactions and feedback loops between these patterns of activation. For example, an external stimulus (e.g. a word) is represented as a pattern of activation at the earliest substage. The last substage at this level contains the final pattern that will be encoded in short term memory (STM), and which has the capacity for storage/learning in long term memory (LTM). Each substage, however, is gated by "top-down" signals from LTM. That is, a pattern at an earlier substage sending input signals to the next substage is transformed by a pattern stored in LTM. This process recurs in a cyclical fashion until there is a match between bottom up and top down signaling (When this occurs, the signal is amplified and is referred to as a "resonance"). Rather than a single node or "memory trace", a pattern of activation across a network is chosen as the basic computational unit to allow for this cyclical matching process between top-down and bottom up information. In this way, recognition does not require a perfect match between external input and mental representation. Instead, a partial match can elicit recognition, and adjustments can be made through learning in long term memory if a mismatch persists.

The system is designed such that the pattern of activation at the last substage, (i.e., that which will be encoded in long term memory) is largely influenced by "top-down" processes (i.e., expectation, familiarity). The particular pattern reaching this stage will be a function of the content stored in long term memory, attention, and the amount of time in which the stimulus has been processed. Thus, while the system is capable of change, it is biased toward stability, and continuously generates expectation.

The macrocircuit operates according to the same principles of the microcircuit. The boxes in figure 2 labelled, "Ai" represent the macrostages for the elaboration of speech, audition and language, with each ascending stage representing a higher level of organization. Also shown are the substages governing language-related motor representations, and connections with the visual system. The cyclical feedback loops promote development, and both the visual and auditory systems have access to the semantic network.

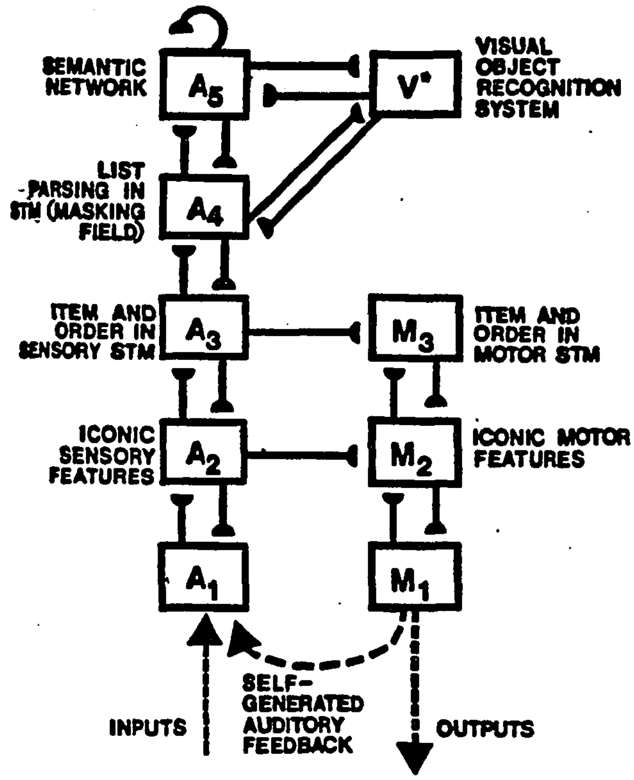



**Figure 2**

A macrocircuit according to Grossberg and Stone (1986) model of Adaptive Resonance. (From: S. Grossberg and G. Stone, 1986, Neural dynamics of word recognition and recall: Attentional priming, learning and resonance. Psychological Review, 93, p. 59. Copyright 1986 by the American Psychological Association. Reprinted by permission.)

The development of each stage builds on the development of its immediately preceding stage. All levels undergo change with development and learning through cyclical feedback mechanisms (in addition to interaction between modalities), similar to that of the microcircuit. The location in the system determines the content of the particular substage as well as the informational content encoded within the LTM filter between stages. For instance, early in development A1 encodes externally evoked auditory patterns which begin to tune LTM filters between stages A1 and A2. Stage A2 constitutes auditory "feature detectors", which are then "chunked" into featural patterns, as "item representations" in A3, (this is the level of STM or "working memory") which also encodes temporal information. These "items" of A3 are chunked and represented as "lists" in A4. Finally, stage A5 (which represents the semantic network, i.e., long term storage) groups list representations of A4 into list representations that exceed the maximal list length permitted within A4 due to the limited capacity of STM. Stage A5 binds together multisyllabic words and "supraword verbal units" (i.e., semantic information) into the most predictive subgroupings. Thus, associations among words are formed as a consequence of contiguity in the language, as well as according to semantic relations.

With the development of each substage, top down signal patterns (e.g., from A2 to A1, A5 to A4, etc...) represent learned expectancies on the basis of past experience. As such, the auditory or visual presentation of a word might match a portion of its LTM pattern at A5 which would include an entire grouping of semantic associates, thus eliciting top/down signaling not only of the particular word but of its associates as well. These would be the mechanisms underlying semantic priming.

The model also distinguishes between subliminal activation and supraliminal (conscious) attention. The theory identifies three signal sources of input which determine level and direction of attention and/or priming: 1) the bottom-up input channel, 2) the top-down template channel, and 3) the "attentional gain control" channel, controlled by an "act of will". Conscious attention is an outcome of the 2/3 rule, which states that at least two of these sources must be active in order for activity in STM to be supraliminal. Priming, however, results from either top-down or bottom-up signaling in the absence of (or prior to) input signals from the other two sources. Accordingly, semantic priming is a consequence of the following separate components of the model: the associative structure of the semantic network, the continuous top-down generation of expectancy, as it is represented in LTM, and possibly the subliminal bottom-up activation of external input signaling (i.e., during semantic priming when primes are masked). Thus, the model offers several plausible means of accounting for semantic priming in which the individual may play more or less of an active role.

In summary, various semantic priming paradigms were described in order to introduce the methodologies that have been utilized in the collection of normative data. These data have been used toward the development of theories of semantic organization and processing. Models of semantic memory represent the ways in which these data have been interpreted, and exemplify the ways in which human semantic memory is currently conceptualized. In spite of their limitations, the models provide a context in which to think about concepts of relatedness, typicality, and semantic priming. Furthermore, theories of semantic organization provide a basis from which to investigate semantic memory as it begins to break down. In the present thesis, the theoretical organization provided by these models is used in predicting the pattern of behavioral and electrophysiological responses in a population (mild Alzheimer's patients) undergoing dissolution of semantic functioning.

D. Breakdown of Semantic Organization as Predicted by Models of Semantic Memory

Most cognitive models of semantic memory do not explicitly map out the course of events and/or the nature of the semantic system as it begins to degenerate. However, given their structure and processing assumptions, reasonable inferences can be made regarding the types of changes that are likely to occur. Based on work with aphasic patients, Jackson (1874, cited in Caplan, 1987) proposed that cognitive/language functions are hierarchically organized according to complexity, and that deterioration begins with the system's most complex features, leaving the more fundamental structural properties relatively intact. In terms of the theories presented above, these features might include those that involved detailed information about concepts, which would impinge upon subtle associations between them. Thus, similar patterns of semantic disruption could be predicted. For example, in the Property Comparison model, the features of a concept might remain associated such that the concept's overall meaning is preserved, yet, the relative weightings of attributes might lose their resolution. Consequently, an item could be identified properly since its features/attributes would be recognized. However, since the relative weightings of defining and characteristic properties lacked precision, typicality and relatedness gradients might be impaired. Similarly, the finer features of network models involve quantity and strength of connections among features and concepts. Just as the relative weightings of attributes might lose their resolution in property comparison models, network models might show changes in criterialities, and/or in the capacity to represent specific types of relations among concepts and features. Again, the general meaning of a concept would remain intact since the structural connections (links) were maintained, but the finer distinctions between concepts might lose their clarity.

Behaviorally, an individual may show difficulty distinguishing between typical and atypical category members. Distinctions between related items within a category might also be problematic. For example, an apple might be called a pear since both concepts share a number of properties/links (i.e., "is a fruit", is sweet, grows on trees, etc.). Similarly, it is unlikely that unrelated items would become confused, since the attributes of a concept and overall connections would be maintained.

Since semantic priming typically involves increased activation of concepts within a general category, performance on lexical decision or naming tasks might appear normal. Assuming that the network's structural links were intact, activation would radiate from the original node, facilitating recognition or naming of related concepts. However, if the connections between items lost their criterialities, activation might not spread as a decreasing gradient for progressively less related concepts. Instead, all related items (i.e., items within a category) would be equally primed. Whereas performance on single-word presentation lexical decision tasks might appear normal, deficits would be observable when a task required differential priming for a particular item. For example, the sentence, "The sky turned dark and it started to _____.", most likely primes the concept "rain." However, if the distinguishing features between concepts are unclear, related items (e.g., snow, hail, sleet) which share many of the same attributes might all be primed to the same extent. Thus, it can be inferred from theories of semantic organization that when this system begins to degenerate, concepts within a category would lose their distinction, whereas overall meaning and superordinate category information would remain relatively intact.

E. Semantic Functioning in Normal Elderly

There is some evidence that semantic functioning undergoes subtle changes with increasing age. Relative to younger adults, older adults (typically, ages 65-80) experience more word finding difficulties in everyday speech (Bowles, 1989). Several studies have shown that older adults show difficulty retrieving word names for pictures (Albert, Heller, and Milberg, 1988; Thomas, Fozard and Waugh, 1977) and for verbal descriptions/definitions of particular words (Bowles and Poon, 1985). Although the commission of errors in fairly infrequent, the type of errors generated by normal elderly are in keeping with some of the predictions which can be inferred from the models (and are qualitatively similar to those seen in Alzheimer's patients, detailed below). Albert et al. (1988) reported an increase in semantic (e.g., semantically related associates, circumlocutions, nominalizations e.g., a word that describes a word's function such as "swinger" for "pendulum") and perceptual errors (which are unlike those predicted by semantic memory models, and not typical Alzheimer's-type errors) with advancing age, with a significant performance drop after age 70. Speed of processing may also be affected by age. Howard, Shaw and Heisey (1986) studied the effects of varying stimulus onset asynchronies (SOAs) in young and elderly adults during a lexical decision task. Whereas no age differences were shown for priming at longer SOAs (450 or 1000 msec), only younger adults showed priming effects at the shortest (150 msec) SOA. This finding was interpreted as evidence for slowing of some aspect of semantic activation; either for the initial activation of a word's representation, or, for the activation of the word's associates. In a lexical decision task using sentence contexts as primes, Madden (1989) reported no age differences in facilitation (semantic priming) at both short (100 msec) and long (1000 msec) SOAs. However, he did note age differences in inhibition (i.e., reaction time to incongruous sentence endings), although these were reduced (nonsignificant) with logarithmic transformation of reaction time.

In contrast to these studies, most of the literature dealing with semantic functioning in normal older adults suggests that semantic functioning remains intact. Unlike Albert et al. (1988) and Thomas et al. (1977), on a picture naming task, Mitchell (1989) found no differences between young and older adults on measures of picture naming accuracy or latency, Tip-of-the-tongue (TOT) responses, or naming consistency (extent to which same name was produced for an item on two occurrences). Despite their slower response times (and absence of priming at extremely short SOAs), under most conditions, normal elderly performance on lexical decision tasks is comparable to that of young adults, indicating that lexical representations are intact. Furthermore, reaction time facilitation from semantic priming during lexical decision tasks is at least as great for older adults as it is for young adults, suggesting that normal elderly maintain sensitivity to semantic relations (Cerella and Fozard, 1984; Howard, 1983).

In summary, the majority of evidence holds that semantic functioning in the elderly is preserved, although under certain conditions, relative impairments have been detected. Some of these changes are in line with models of semantic organization (described earlier), and are similar, although much less severe, to those changes that occur in Alzheimer's patients. It is not clear whether these behavioral/cognitive changes reflect similar underlying processes.

F. Breakdown of Semantic Memory in Alzheimer's Disease

Alzheimer's Type Dementia is principally associated with a decline in episodic memory functioning (i.e., poor memory for recent events; Spinnler et al., 1988). Within the past decade, however, there has been an accumulation of evidence indicating that the early stages of Probable Alzheimer's Disease (the term "Probable Alzheimer's Disease" or "PAD" is often applied because a diagnosis of Alzheimer's Disease is typically not confirmed until autopsy) can be characterized by a relatively circumscribed language impairment (Nebes, 1989). Specifically, in conversation and on a variety of behavioral tasks (e.g., object naming, recognition tasks, verbal fluency) PAD patients tend to produce semantic errors (as opposed to phonemic or literal errors), and evidence a disnomia (e.g., word finding difficulties resulting in either circumlocution or the production of a semantically related word) similar to aphasic subjects with lesions in the posterior left hemisphere (Buckingham, 1979). Although PAD is associated with a neuropathological disease process[1], for the present purposes, these neurophysiological changes are less important than the pattern of changes in language functioning that are manifest behaviorally. PAD patients are

[1]Presently, Alzheimer's Disease cannot be definitively diagnosed until autopsy. Histologically, Alzheimer's Disease has been associated with neuritic plaques and neurofibrillary tangles, found primarily in the medial temporal lobe. These histological changes are usually accompanied by marked reductions in cholinergic activity (Terry & Davies, 1983). However, similar changes have been found in individuals who did not show cognitive impairment (Terry, 1978). Thus, it is possible that neuropathologically, the differences between normal elderly and Alzheimer's patients may be a matter of degree, or, related to brain site of the plaques and tangles.

discussed (and selected as an experimental group) because they show predictable deficits in semantic processing under a variety of circumstances.

F1. <u>Word Finding</u>

Word finding difficulty in spontaneous speech and on standardized tests is apparent in the early stages of PAD. More informative than the deficit itself, is the pattern of errors generated when word retrieval is unsuccessful. Although a range of language errors are possible, e.g., a) grammatical or syntactic errors, e.g., He give me a book, b) Phonemic paraphasias e.g., "bold" instead of "gold", c) morphological errors, e.g.,"feet" instead of "foot", d) unrelated word substitutions, e.g., "house" instead of "pencil", PAD patients rarely commit these types of errors. Instead, they produce a disproportionate number of semantic paraphasias, i.e, substitution of a semantically related word for a particular word, e.g., "bank" instead of "wallet". The semantic paraphasias typically involve substitution of another member within the same category, e.g., "pear" instead of "apple", or, of the name of the superordinate category, e.g., "fruit" instead of "apple" (Nicholas et al., 1985; Ober et al., 1986).

Object naming tasks, which are considered sensitive to language impairments due to brain injury (Benson, 1979) have elicited similar patterns as those found in free speech. Although normal elderly subjects in their 70's perform poorly compared to subjects in their 30's (a 9% decrement reported by Nicholas et al., 1985), PAD patients show substantially more difficulty on this task (Albert et al., 1988). Again, PAD patients tended to produce names of other items within the same category, or the name of the superordinate category. Other semantic paraphasias included functional words, e.g., "sweep" for "broom", contextual phrases, e.g., "you find it in a jewelry store" for "watch" (Bayles and Tomoeda, 1983), and semantically related words, e.g., "socket" for "plug" (Martin and Fedio, 1983). Perceptual errors, i.e., naming an object that is visually similar to the object (e.g., calling an anchor a hammer) were much less common. Flicker et al., (1987) presented PAD and control subjects with line drawings selected from the Boston Naming Test (Kaplan et al., 1983). Subjects were instructed to select the name of the object from a list of four words presented beneath the pictures which included the name of the object and 4 of the following: a) a member of the same semantic category, b) the name of the superordinate category, c) a perceptual distractor, d) a phonological distractor, e) an unrelated item. Most errors committed by PAD patient s involved distractor items which were members of the same semantic category as the target item.

Verbal fluency is defined as the ability to retrieve members belonging to a specified category within a limited time period (Rosen, 1980), usually 60 to 90 seconds. The specified category can be either a letter category (e.g., words beginning with the letter "S") or a semantic category, (e.g., types of animals, types of fruits). On a semantic category task, Ober et al. (1986) reported that mild PAD patients generated approximately half the number of items generated by normal elderly control subjects, and that the PAD patients' rate of production reached its asymptote relatively faster. Furthermore, PAD patients produced significantly more errors (perseverations and noncategory members). Ober et al. (1986) also reported this same pattern for the letter category task. However, using these same tasks (i.e., animals and fruits for the semantic category, and F, A, S for the letter category), other investigators have reported conflicting results. For instance, although Butters et al. (1987) found similar patterns on the semantic category task, they found no differences between normal elderly and PAD patients on the letter category task. In contrast, Rosen (1980) reported the opposite pattern, in which PAD patients performed similar to normal elderly on the semantic category task, but showed impaired fluency on the letter category task. In spite of these differences, Ober et al. (1986) reported that although PAD patients produced less instances per category, their response patterns indicated that their "dominance" and frequency structure was still in place. That is, for both normal elderly and PAD patients, the typicality of the items named decreased with the amount of time on task.

The "supermarket task", a more demanding fluency test, has produced more consistent results. In this task, subjects are asked to name items found in a supermarket. Martin and Fedio (1983), Ober et al. (1986), and Troster, et al. (1989) reported similar patterns for normal elderly subjects and PAD patients. Generally, normal subjects approached the task systematically by naming several items within a particular category (e.g., fruits & vegetables, dairy products, meat, fish and poultry, etc...) before shifting to another category. In contrast, PAD patients tended to generate only a single item from a category, and were more likely to give the category name itself (e.g., fruits, meats). Normal control subjects rarely produced category names. PAD patients produced significantly less items overall, and their search was much less systematic. This pattern was regarded as particularly informative, since it could not be attributed to a general slowing, which would have resulted in a ratio of category names to category items that was similar to that of normal controls. Instead, PAD patients produced significantly less items per category, and named significantly more superordinate category terms.

F2. Attribute Knowledge

It has been theorized that the semantic deficits observed in PAD stem from a loss of detailed attribute information which comprise the meaning of a concept (Martin and Fedio, 1983; Huff et al, 1986). That is, although general semantic information is maintained, such as the item's superordinate category, there is a progressive loss of (or ability to access) more detailed information about a concept which is essential in distinguishing among related items within a category. In a study with three demented patients of unknown etiology, Warrington (1975) reported that although patients were able to correctly answer questions about an item's superordinate category, they were unable to answer detailed questions about the item's physical attributes, ( e.g, Is it bigger than a cat?; Is it made of wood?). Based on this work, Chertkow and Bub (1990) conducted a similar study in which they presented 130 pictures to PAD patients and normal elderly controls. Instead of naming the picture, subjects were instructed to answer a series of 6 forced-choice questions about each item (3 questions about perceptual attributes, e.g., for "saw": Is the tip make of metal or "wood", and 3 contextual or functional questions, e.g., Do you cut things with it or lift with it?; Is it used on a piece of wood or stone?). On this task, PAD patients made significantly more errors than normal elderly controls (PAD: 87% correct versus. Normal Elderly: 98.4% correct). Strikingly, PAD patients were markedly impaired on questions concerning biological categories (i.e., animals, fruits, vegetables), answering only 68.2% correct, compared to 98.6% correct by normal elderly controls. This finding has also been reported by other investigators (Gainotti et al., 1990). Riddoch (1988) had previously speculated that this result was due to greater conceptual similarity among members within a biological category.

Nebes and Brady (1988) conducted a study with PAD patients and normal elderly to investigate the possible loss of attribute information relative to category/associative information. Subjects made speeded yes/no decisions about whether a word (i.e., an attribute) was related to a particular target. Attributes were either the superordinate category, a general associate, a physical feature, or an action/function of the concept. It was predicted that if attribute information was lost or disrupted relative to categorical information, reaction time to the former would be longer compared to the latter. However, neither subject group was affected by the nature of the attribute. In a second part of the study, when subjects were instructed to generate an attribute of the target concepts, PAD patients were impaired (for both types of attributes), suggesting that the source of the deficit is more likely a problem in self-directed search procedures, rather than due to actual loss of information.

Grober et al. (1985) conducted a series of experiments from which they also claimed that conceptual attributes are not lost in dementia, rather, their organization is disrupted. In their first experiment, subjects were presented with a word and an 18-word check list (9 targets, 9 unrelated foils) and instructed to check off only those words which were important to the understanding of the test item. Although demented subjects responded correctly to 95% of the targets, they also responded to significantly more foils than normal elderly controls. The second experiment was a forced choice procedure in which subjects chose between one of two words (a target and a foil) that went with the test noun. Although normal elderly controls performed significantly better than the patient group, the patients were surprisingly successful (six out of 10 patients made no errors; the poorest performance of any patient was 88% correct). However, inspection of the data from the first task revealed that although some attributes are more important than others toward the understanding of a concept, omissions of essential attributes occurred as often as omissions of nonessential attributes. (Attributes were defined as individual pieces of information in a concept's representation). Thus, the third experiment was designed to probe the organization and processing of detailed attribute information. Subjects were presented with a test noun, and three related attributes, and instructed to rank the importance of attributes to the understanding of the test nouns (e.g., for TABLE: furniture- essential, wood,- intermediate and department store- nonessential). Only on this last task were the patients significantly impaired relative to normal elderly controls. Grober et al. (1985) concluded that although a substantial amount of attribute information is retained and remains connected with corresponding concepts, information regarding the relative saliency of attributes (which reflect relative importance of an attribute for the meaning of a concept) is disrupted. Thus, the findings from the majority of these studies suggest that in PAD, while categorical or more general information remains available, the representation of specific information (e.g., detailed attributes, relative importance of attributes) is somewhat impaired.

F3. Semantic Priming

Semantic priming has been assessed in PAD patients in much the same way as described above for normal young and elderly adults. As described, under most conditions, normal elderly subjects show typical semantic priming effects (Burke et al 1987; Howard, 1988). Similar findings have been reported for PAD subjects. On word naming tasks, Nebes et al. (1984) and Hartman (1987) found that although PAD patients were significantly slower than control subjects, the degree of semantic priming was similar for both groups. In another study, Nebes et al. (1986) presented auditory sentence contexts to prime visually presented words. With respect to the final word (e.g., "disease"), sentences were either: 1) congruous (e.g., "The child was born with a rare___",) 2) incongruous (e.g., "When you go to bed turn off the ____", or 3) neutral (e.g., "They said it was the _____."). Again, facilitation associated with congruous sentences was comparable for normal elderly and PAD subjects, (although inhibition from incongruous sentence contexts was greater in PAD patients). On a typical lexical decision task, Nebes, Boller and Holland (1986) actually found that PAD patients showed more facilitation by semantically related primes compared to normal elderly. The majority of evidence suggests that semantic priming effects are preserved in PAD (Nebes et al, 1989; Ober and Shenaut, 1989). However, one study (Ober and Shenaut,1988) did not find facilitation in PAD patients, but this discrepancy was probably due to differences in stimulus durations and interstimulus intervals (ISIs).

There is a growing position in the literature that the deficits shown by PAD patients on semantic tasks reflect impairments in the ability to access and use semantic information, rather than an actual loss or disorganization of semantic information (Nebes, 1989). This view is based on an accumulation of evidence that the performance of Alzheimer's patients on semantic tasks seems highly task-dependent. PAD patients seem to be most impaired on tasks requiring an intentional memory search (i.e., verbal fluency tasks, naming), which rely heavily upon attention and retrieval functioning. Under certain conditions, PAD patients are unable to produce semantic information about a particular word but appear to have access to semantic information about the same item when the task is modified. For example, PAD patients who could not describe the use of a presented object could choose the item that would be most useful for a particular chore when given multiple choice (Flicker, et al., 1987). Similar findings have been reported by Huff et al., (1988).

There is some evidence that PAD patients are more dependent upon semantic context than are normal subjects. On a Cloze procedure task (subjects completed sentences in which the final word was omitted), PAD and normal elderly performed equally well when sentence contexts were highly constrained. However, PAD patients were differentially impaired on contexts of low or medium constraint (Nebes et al., 1986). This over-dependence on semantic context implies that at some level, the semantic information is retained, yet, a number of supportive cues (although not required by normal individuals) must be present in order for PAD patients to use this information appropriately.

In an associative priming paradigm, the link between two semantically related items is primed with the expectation that when one word of the pair is presented in a free association task, the other is more likely to be given as an associate. Both Salmon et al. (1988) and Huff et al. (1988), using a semantic judgement task to induce priming, found negligible associative priming in PAD patients. Nebes (1989) attributed the discrepancy between associative priming and lexical decision priming to the active retrieval demands required in the former. It is possible that when semantic cues are present and attentional and retrieval demands are minimized, the semantic impairment in PAD is at least reduced. However, it is still uncertain whether this impairment reflects an actual loss of information, or merely a disruption in its organization and /or accessibility.

In summary, investigations of language functioning in the early stages of PAD have generated consistent patterns of disruption in semantic organization and/or processing. Specifically, superordinate category information appears to remain intact, whereas particular items within a category seem to lose their distinction. This pattern of deterioration is fairly consistent with that which was inferred from the semantic memory models described above. The types of errors made by mild PAD patients in free speech, object naming, and category fluency tasks consist of substitutions of related words and/or superordinate terms in place of particular items. When questions concerned superordinate category information, PAD patients performed normally. Yet, when questions involved information about an item's attributes and/or relative saliencies of these attributes, performance declined. Normal semantic priming in lexical decision tasks also supports the notion that superordinate category structure is maintained. It is somewhat surprising that typicality gradients do not show impairment. Hence, it is possible that this dimension is not affected until more advanced stages of the disease process. Given the behavioral data presently available, it is still unresolved whether this behavioral pattern reflects an actual loss of information, a disorganization , or an inaccessibility to this type of information.

G. Event Related Potentials (ERPs)

The scalp-recorded ERP is a segment of electroencephalographic (EEG) activity that is time locked to an event (typically a physical stimulus) and is associated with sensory and/or cognitive events. Unlike behavioral measures that are dependent upon overt responding, ERP recording provides a millisecond by millisecond readout of some of the brain's activity during information processing, and is not necessarily contingent upon overt behavior or conscious awareness. Thus, the concurrent recording of behavioral responses and ERP activity can provide important correlative information regarding the nature, sequence, and timing of perceptual and cognitive events.

ERPs recorded from the human scalp are usually too small to detect within the ongoing EEG. In order to identify the ERP elicited by a particular stimulus, a technique known as "signal averaging" is used to enhance the "signal" (the ERP) and reduce the background noise (i.e., the spontaneous EEG). The averaging of ERPs requires that the "event" be repeated. With each occurrence, the electrical activity is recorded over a period of time in synchrony with the stimulus. Since the background activity is presumably random with respect to the event, averaging causes the background noise to approach zero. However, the signal (i.e., that portion of the ERP which is related to the "event") remains relatively constant during the averaging process, and therefore becomes more prominent (Picton, 1980). The result is the ERP waveform which typically consists of a series of positive and negative peaks. These peaks or deflections are identified according to their polarity, peak latency, amplitude, duration, and scalp topography. This information is used to identify particular components that reflect the activity of functionally distinct (although not necessarily anatomically distinct) neuronal aggregates which are reliably responsive to experimental variables (Donchin, Ritter and McCallum, 1979). Analysis of the ERP involves examining changes in components along these various dimensions in response to experimental manipulations.

Components are typically classified as either exogenous or endogenous. Exogenous components (or "sensory evoked potentials") are obligatory, in that the presentation of a physical stimulus necessarily elicits the particular component, regardless of an individual's psychological state (e.g., attending or not attending). These components occur relatively early (i.e., approximately, within the first 100 msec post-stimulus), and are responsive (i.e., vary in amplitude, latency, etc...) to physical parameters of the stimulus such as intensity, duration, or frequency. In contrast, endogenous components are responsive to cognitive/psychological parameters such as attention, probability, expectancy, and/or meaning. For example, a positive deflection occuring approximately 300 msec post stimulus, (referred to as P3 or P3b)[2], is typically elicited by a task relevant stimulus (Tueting, Sutton and Zubin, 1971). Its amplitude increases as the probability of stimulus occurrence decreases and its latency varies with the perceptual difficulty of the task (Ritter, et al., 1972).

G1. <u>Background of N400</u>

Investigations in which ERPs have been recorded from subjects engaged in linguistic tasks have indicated that the N400, a negative component peaking at approximately 400 msec post stimulus is sensitive to semantic processing. Specifically, N400 amplitude varies inversely with the extent to which a word has been primed by its preceding semantic context (but see below for contrary findings).

In a series of experiments, Kutas and Hillyard (1980, 1984) recorded ERPs to individually presented words of sentences. The intention of their

---

[2]The term, "P300" was initially used to identify the late positive component, typically occuring 300 msec post stimulus onset when elicited by rare stimuli during oddball paradigms. However, "P3" or "P300" is also frequently used to refer to the "late positive component", which often occurs later than 300 milliseconds in more complex tasks (e.g., sentence processing, memory/recognition tasks). Furthermore, the "N400" typically precedes the "P300".

initial study was to use P3 as a marker of semantic deviance. They predicted that unexpected incongruous sentence endings (e.g., I take milk with cream and *dog*.) embedded in a series of meaningful sentences would elicit a relatively large amplitude P3, and that P3 latency would reflect speed of word recognition and comprehension processes. Subjects were instructed to read each sentence silently in order to complete a questionnaire later in the session. Seven-word sentences were presented one word at a time (duration 100 msec, SOA 1000 msec). In the first experiment, 75% of the sentences ended meaningfully, and 25% ended with a moderately inappropriate (but syntactically correct) word (e.g., He took a sip from the *waterfall*). In experiment 2, 25% of the sentences ended with a highly incongruous word (e.g., He took a sip from the *transmitter*). In a third experiment, 25 % of the sentences terminated in a semantically expected but physically deviant (i.e., boldface type) word (e.g., He took a sip from the **STRAW**; Kutas and Hillyard, 1980). The ERPs elicited by terminal words that were physically deviant were characterized by a large amplitude P3. However, the ERP elicited by semantically anomalous sentence endings did not show a typical P3. Instead, these ERPs were characterized by a negative component (N400) onsetting at approximately 250 msec, and peaking at about 400 msec. Furthermore, the N400 to highly anomalous endings was larger compared to the N400 elicited by moderately anomalous endings. Thus, it was inferred that N400 was not merely a general response to deviation from expectancy (since it was not elicited by unexpected boldface lettering), rather, N400 was more intimately associated with deviations related to semantic processing.

These initial findings spurred numerous investigations designed to determine the necessary conditions for the elicitation of N400 and to identify the parameters affecting its behavior. Kutas and Hillyard (1984) manipulated expectancy of congruous sentence endings. Although all sentences ended meaningfully, the predictability of the terminal word varied according to "cloze probability." The cloze probability of a word is defined as the percentage of subjects who would complete a particular sentence with that particular word. For example, 99 out of 100 people completed the sentence, "I mailed the letter without a ____", with the word "stamp" (Bloom and Fischler, 1980). Thus, for this sentence, "stamp" has a cloze probability of .99 and, the sentence fragment is considered to be highly constrained. On the other hand, "There was something wrong with the _____." is considered a low constraint sentence, since it has a number of equally acceptable endings, which, consequently would have moderate or low cloze probabilities, (e.g., "car", "knee", "food"). Subjects were presented with sentences of high, medium and low constraint, which ended with words of high, medium or low cloze probability. The ERP to high cloze probability words was characterized by a widely distributed positive component (P300). In contrast, ERPs to low cloze probability words (regardless of level of sentential constraint) showed a large N400. Although the amplitude of the N400 did not vary with contextual constraint, it did vary inversely as a function of cloze probability. Specifically, the lower the cloze probability, the larger the N400. The strength of this relationship was supported by correlations (between cloze probability and N400 amplitude) ranging from .88 to .97 (Kutas and Van Petten, 1988). In addition to establishing this relationship between N400 amplitude and expectancy in a semantic/linguistic context, it was also concluded that semantic anomaly was not a necessary condition for the elicitation of N400.

Fischler, Childers, Roucos, & Perry (1983) investigated the relationship between N400 amplitude and truth value of sentences. ERPs were recorded to final words of sentences which were similar to those used in verification tasks (described above):

| | |
|---|---|
| True affirmative | "A robin is a bird" |
| False affirmative | "A robin is a vehicle" |
| True negative | "A robin is not a vehicle" |
| False negative | "A robin is not a bird" |

Interestingly, Fischler et al. (1983) found that the presence of N400 (elicited by the second content word) depended not upon the truth value of the sentence, but on the associative relationship between the two content words of the sentence. That is, the true affirmative and the false negative sentences which both contain two related words (e.g., robin, bird) did not elicit an N400. However, the false affirmative and the true negative each contain two unrelated words (e.g., robin, vehicle) and therefore did elicit an N400. Based on these results, it was suggested that the N400 might be responsive to that aspect of semantic priming which influences reaction time as a function of semantic relatedness. Accordingly, Bentin, McCarthy and Wood (1985) recorded ERPs during a typical lexical decision paradigm. Semantic priming was indicated by the expected reaction time facilitation to primed target words preceded by semantically related words (i.e., primes). In addition, they found that although the N400 was present for all items, its amplitude to primed words was smaller relative to words that were preceded by unrelated words. Thus, the presence of N400 was not contingent upon sentential processing, and, even in the context of single words, its amplitude was responsive to semantic association.

Given the inverse relationship between N400 amplitude and cloze probability, Kutas et al. (1984) reasoned that N400 might reflect the degree of expectancy created by a semantic context. However, in lexical decision tasks using sentence contexts, reaction time is facilitated not only to the most expected word (i.e., highest cloze probability) but also to words which are *related* to the best completion (Kleiman, 1980). In this sense, an expectancy is also created for words which are semantically related to the best completion. Given these data, in combination with Fischler et al.'s (1980) data indicating that N400 is unresponsive to propositional content, Kutas et al. theorized that N400 amplitude was primarily responsive to the degree to which a word was primed by its prior semantic context. Thus, N400 amplitude should be attenuated to words which are semantically related to the particular word being primed. When ERPs to low cloze probability endings were re-averaged according to degree of relatedness to best completion (as assessed by 25 subjects in a separate rating task), N400 elicited by acceptable/related words was smaller relative to acceptable/unrelated words. (For instance, given the sentence, "He kept the dog on a ____", best completion is "leash", N400 amplitude to "chain" was smaller relative to that elicited by "diet"). Finally, in a separate experiment, ERPs were compared for terminal words which were either incongruous and related (e.g., "The game was cancelled when it started to *umbrella*. " ) or incongruous and unrelated to the sentence's best completion. Although N400 was elicited by all incongruous endings, its amplitude was significantly greater for unrelated endings. Again, N400 amplitude to final words varied inversely with the relatedness of the "target" word to the primed word (i.e., best completion). Taken together, these data provide strong evidence that N400 is intimately associated with semantic processing, and can be used as an index of semantic priming.

G2. Conditions that elicit N400

Since its initial report, N400 has been associated with controversy. For instance, it is still unclear whether the N400 response is specific to unexpected semantic deviation, or, whether other types of unexpected deviations during linguistic (or nonlinguistic, e.g., Barrett, Rugg, and Perrett, 1988; Friedman, Putnam, Ritter, Hamberger and Berman, in press) processing are sufficient for its elicitation (Kutas and Van Petten, 1988). A number of experiments support semantic specificity. As mentioned above, during a sentence reading task, surprisingly large (in physical size) but semantically expected final words did not elicit N400s (Kutas and Hillyard, 1984). Also during a reading task, Kutas and Hillyard (1983) reported that ERPs recorded to grammatical anomalies (e.g., As a turtle grows its shell *grew* too) did not show typical N400 effects (i.e., they found a small amount of negativity from 300-400 msec post stimulus at frontal electrode sites, which was unlike a typical N400 effect). Besson and Macar (1987) recorded ERPs to predictable and unpredictable stimuli which terminated in three non-linguistic and one linguistic predictable sequences. The non-linguistic conditions were : 1) well-known French melodies, 2) geometric figures increasing in size, and 3) notes of the musical scale in ascending order. The linguistic condition consisted of visually presented sentences. For each condition 25% of the stimulus sequences ended incongruously (e.g., a smaller figure in the geometric condition, an inappropriate note in the scale or song condition, etc.). The only stimuli which elicited N400 responses were the incongruous sentence completions. All other unexpected stimuli elicited P3-like positivities.

Although most reports of N400 are limited to deviations within linguistic contexts, Stuss, Sarazin, Leech and Picton (1983) recorded an "N400-like" component (called, "Ny") from subjects performing a mental rotation task. Subjects were required to decide if two figures were identical or mirror images. Stuss et al. claimed that the task itself evoked the negativity, since type of response (same or different) did not influence the component's amplitude. Although these authors suggested that the Ny/N400 might reflect a processes required in the evaluation of any complex stimulus, the distribution of Ny (frontal maximum) was unlike the N400 reported by Kutas and Hillyard (e.g., 1980, 1984) which was maximal at parieto-central sites.

Friedman et al., (in press) also reported an N400-like component (called "Neg400") elicited by the second picture in a picture-matching task. Depending upon the instruction condition, subjects judged whether two pictures were "same" or "different" based on whether they were physically identical, shared the same name (e.g., different orientations of the same item) or were members of the same category (e.g., horse and cow). Neg400 amplitude increased as the two drawings became increasingly dissimilar. Additionally, when stimuli were physically identical, N400 increased as the instruction condition became increasingly complex (i.e., N400 was largest in the category decision condition). Similar to Stuss et al.'s (1983) assertion, this finding also suggests that "task" plays an important role in the elicitation of "N400" effects. Furthermore, similar to Stuss et al.'s findings, Neg400 was most prominent at fronto-central electrode sites.

In a face matching task, Barrett, Rugg and Perrett (1987) reported an N400-like component which was largest when elicited by familiar non-matching faces. Barrett et al., speculated that N400-like components might be elicited by stimuli that are inconsistent with currently activated memory representations, although not exclusively within the semantic domain.

Also at issue is whether the N400 is a unique component, as some investigators propose that N400 might actually be a delayed N2 (Deacon, Breton, Ritter and Vaughan; in press). The N2 is an earlier component of the ERP which responds similarly to some of the same parameters that influence N400 (i.e., amplitude is greater to unexpected stimuli). N2 is typically associated with an unexpected physical change in a train of stimuli (e.g., change in intensity, frequency, etc.). However, there are reports of N2 elicitation when stimulus change was defined by an abstract rule (e.g., male versus female names-- Ritter, Simson and Vaughan,1983). It has been argued that N2 and N400 are distinct components since N2 (but not N400) is responsive to stimulus probability, and N2 and N400 show different scalp distributions (Kutas and Hillyard, 1980). However, Deacon, et al., (in press) recently demonstrated that under certain conditions, N400 is sensitive to probability. Additionally, N2 and N400 showed similar scalp distributions. Despite these similarities, they also reported that although N2 latency increased when stimulus discrimination was more difficult, N400 latency was unchanged. Hence, the distinction between N2 and N400 requires further investigation.

For present purposes, the particular classification of the N400 is less important than the fact that there is a component of the ERP that can be differentially responsive to semantic expectancies during linguistic processing. Other variables which potentially affect N400 amplitude under these circumstances are detailed below. During single word presentation (e.g., lexical decision tasks), Rugg, (1990) reported that N400 was larger to low frequency words relative to high frequency words. However, Van Petten and Kutas (1989) demonstrated that during sentential processing, this effect was attenuated with increasing word position in the sentence, and is virtually absent for the terminal word. Stuss, Picton and Cerri (1988) reported that N400 was larger for atypical words (individually presented) compared to names of typical category exemplars, although these stimuli showed a potential confound with word frequency.

Each of these variables, however, is a correlate of expectation. In the context of single words, there is a greater expectation for high frequency words and typical category exemplars relative to low frequency words and atypical exemplars. In normal sentences, each additional word (in sequence) provides increasing constraint on the possibilities for the subsequent word, which increases expectation for a particular word or class of words (e.g., the final word of a 7-word sentence will be more predictable/expected than the second or third word of that sentence).

G3. Effect of Task Demands

Another unresolved issue concerns the extent to which task demands contribute to N400 effects. Is N400 is an automatic brain response, or, is it contingent upon a particular type of attentionally directed processing? Several investigators assert that regardless of instructional set, when linguistic stimuli are employed, comparison of ERPs to related and unrelated stimuli (with respect to their prior semantic contexts) do show N400 effects. However, the magnitude of this effect appears to be influenced by task demands. Kutas (1981, 1985) and Kutas and Hillyard (1989) conducted a series of experiments which addressed this issue. In a task which did not require semantic analysis, subjects were presented with two words and a single letter in sequence. Subjects were to determine whether or not the single letter was present in either of the preceding words, and they were not informed that 75% of the word pairs were semantically related. Although semantic association was only incidental to the assigned task, comparison of ERPs to the second word showed that N400 was larger to words that were preceded by unrelated words. In another paradigm, subjects' attention was drawn to the semantic association of word pairs by the instruction to rate the strength of the association between the words. As expected, N400 amplitude was larger for unrelated words. Furthermore, this effect was larger than that found in the letter search task, but smaller when compared to N400 effects found during sentence priming tasks (although this comparison is based on two different experiments with different subjects, Kutas and Hillyard, 1980,1984).

Several interpretations have been proposed to account for this pattern. On the one hand, the magnitude of the N400 effect might reflect the degree to which the subject's attention is directed toward semantic analysis. Thus, smaller effects were found in the letter search task which did not explicitly require subjects to access word meanings. On the other hand, the N400 effect may reflect the extent to which a particular word is primed. Kutas and Van Petten (1988) noted that sentence fragments provide a more restrictive context than does a single word, and that N400 effects elicited by low cloze probability sentence completions are more similar to those elicited by words pairs. These same authors discuss that most likely, there are both automatic and attentional components of the N400 effect, and that the presence of N400 effects even when semantic processing is not task relevant may indicate that semantic analysis may be a default mode of processing when dealing with word stimuli.

In contrast to this position, other investigators have suggested that semantic analysis is a necessary condition for the elicitation of N400 effects. Hamberger and Friedman (submitted) and Deacon, et al. (in press) demonstrated an absence of N400 effects using word stimuli when the task required an orthographic (i.e., size) rather than a semantic discrimination. Although these findings appear at odds with those outlined above, Deacon et al. (in press) offered an explanation which may account for these discrepancies. These authors proposed that the presence of N400 effects is not determined by task relevance *per se*. Rather, regardless of task relevance, N400 effects occur when items are processed for meaning. If task performance is made more efficient by semantic analysis, e.g., as in the sequential letter-search task previously described (Kutas and Hillyard, 1989), N400 effects occur. However, in a task requiring a size discrimination, semantic analysis provides no advantage (and perhaps interferes with task performance), and therefore, N400 effects are not observed.

H. Rationale

In spite of the controversies concerning N400, it is apparent that under certain conditions, N400 provides a graded index of semantic priming. As indicated by performance on several behavioral measures, PAD patients evidence a degeneration in this gradient, in which semantically related items seem to lose their distinction while superordinate category (i.e., more general) information is relatively preserved. The present study was designed to determine whether the collapsing of the semantic gradient in this population would also be observed in the N400 component of the ERP.

Most studies of PAD patients engaged in semantic processing have relied upon overt behavioral responding which occurs long after stimulus presentation. Hence, it is unknown where in the course of processing the disruption occurs. For instance, an impairment could reflect a disruption: 1) in the structural organization of the semantic system 2) in accessing particular concepts or types of concepts, and/or 3) in the production of a response. Since the N400 occurs prior to (and is not contingent upon) behavioral responding, the collection of ERPs affords the ability to observe activity related to processing prior to the overt response. The correspondence (or dissociation) between ERPs and behavioral measures in this population could thus be used to further understand the nature of semantic processing as it begins to unravel. This, in turn, could provide information about the organization of the system when it is intact.

A pilot study was conducted in order to create an appropriate task, to collect normative data that would provide a basis for the selection of stimuli, and to determine whether the task and stimuli would elicit the desired N400 gradient in normal young adults.

# II. PILOT STUDY

## A. Part 1: Development of the Stimuli

Before beginning the full investigation, it was necessary to create a stimulus set and task that would produce a reliable gradient (as a function of semantic relatedness) in both the behavioral and electrophysiological responses of normal subjects. First, it was critical that on any given trial, one particular word within the semantic network would be most activated (i.e., primed) relative to its semantic associates. It was decided that sentence contexts (as opposed to single words) would most effectively provide this strong contextual constraint. Second, it was necessary to construct a task that would induce semantic processing of the stimuli (see Introduction: Effect of task demands). A sense/nonsense decision task was chosen in which for each sentence, this decision was contingent upon the meaning of the final word. Thus, adequate performance would provide reasonable evidence that the target word was processed for meaning.

The pilot study consisted of two phases. The initial phase involved the development of the stimuli. Four sets of sentences were needed in which the final word of each sentence would be differentially primed according to its set or "Sentence Type". A series of norming procedures were conducted toward this end. First, a large number of sentences were designed to be highly constrained. Cloze norms were collected for the final words of these sentences via a sentence completion task. Only sentences which elicited a Cloze probability of at least 84% were retained for the experiment. For present purposes, the final words of these highly constrained sentences (i.e., those with the highest Cloze probability) will be referred to as "best completions", and the portion of the sentence prior to the final word will be called the "sentence context" (e.g., Sentence context: "She put lemon and sugar in her___"; best completion: "tea"). A sentence context would provide the strongest activation for its best completion, and less activation for a word which was related to its best completion. The next step involved the construction of two sets of sentences in which the final word of each sentence would be semantically related to the sentence's best completion. An effort was made to construct one set in which sentences were clearly sensical and another in which sentences were clearly nonsensical. In order to minimize ambiguity, a norming study was conducted in which subjects judged whether each sentence "made sense" or "did not make sense". Only those sentences for which there was consensus or near consensus were retained. Forty sentences (contexts and best completions) were left intact. The remainder of the sentence contexts and their best completions were mismatched such that the terminal words of these new sentences were unrelated to the sentence context's best completion, and these sentences were nonsensical. These two sets of stimuli were also subject to sense/nonsense norming procedures. These norming studies generated four sets of forty sentences, each defined by: 1) the relationship of the sentence's final word to the best completion for that sentence, and 2) whether or not the sentence made sense. The names of the four Sentence Types with examples of each are shown in table 1. Throughout the text, Sentence Types will be referred to as "Best Completion" or "BC", "Related/Sense" or "RS", "Related Nonsense", or "RN" and "Unrelated/Nonsense" or "UN". It was expected that these two constraints (i.e., relatedness and sensibility) conferred an ordinality such that terminal words of the first Sentence Type (Best Completion), would be maximally primed, terminal words of the second (Related/Sense) and third (Related/Nonsense) Sentence Types would be moderately primed, and those of the fourth Sentence Type (Unrelated/Nonsense) would not be primed. Thus, the ascending order of N400 amplitude would be: 1 (BC), 2 (RS), 3 (RN), 4 (UN).

Table 1
Sentence Types

| Sentence Type | Example |
|---|---|
| 1 "Best Completion"<br>Best Completion/Sense | The guard sounded the alarm. |
| 2 "Related/Sense"<br>Related to Best Completion/Sense | The guard sounded the bell. |
| 3 "Related/Nonsense"<br>Related to Best Completion/Nonsense | The guard sounded the lock. |
| 4 "Unrelated/Nonsense"<br>Unrelated to Best Completion/Nonsense | The guard sounded the molars. |

Norming procedures are detailed below.

A1. Cloze Norms

In addition to the requirement that sentence contexts be of high constraint, in order to minimize stimulus variability, it was decided that all final words be limited to nouns. Although a number of stimuli were taken from Bloom and Fischler's (1980) sentence completion norms, only a subset of their sentences met these criteria. Furthermore, the population sampled by Bloom and Fischler included only college-aged students. Since the population to be tested in the present study included elderly as well as young adults, subjects over the age of 65 were included in the present norming study in order to test the generalizability to older subjects of Cloze norms produced by a younger population.

A1a. Method

A1a1. Subjects

One hundred "young" volunteers, (72 female, mean age 29.9 years old, SD = 8.3, age range = 18 - 56 years), participated in the sentence completion task. The mean number of years of formal education (ED) was 16.21, (SD = 3.0 years, range = 9-24 years). Mean rating of Socio-economic Status (SES) according to the Hollingshead index (Hollingshead and Redlich, 1958) was 58.17, (SD=14.8, range = 20 - 82, higher score= lower SES). The "older" sample included 30 elderly volunteers, (mean age = 71.3 years, SD = 4.5 years, age range: 65 - 82 years). Mean ED was 16.3 (SD = 11.0, range = 8-24 years), and mean SES rating was 56.9 (SD = 21.6, range = 20 - 91). All subjects were native English speakers, or had learned English prior to age four. There were no significant differences between age groups in SES or in years of formal education.

A1a2. Construction of sentence contexts

A total of 198 sentence contexts were created by the experimenter and colleagues. An attempt was made to construct sentence contexts that would elicit the same final word in most subjects. The following guidelines, as described by Bloom and Fischler (1980), were applied: 1) the addition of a single word would make each context a grammatically acceptable English sentence, 2) obvious cliches were avoided, and 3) sentences were no longer than 10 words.

A1a3. Procedure

All subjects were given a 9-page booklet. The cover page provided a space for demographic information (i.e., age, sex, occupation, and education background), and the instructions, which read as follows:

> On the following pages are a large number of sentences, each with the final word left blank. Your task is simply to read each sentence at your normal rate, and write down the word that first occurs to you as a likely end of that sentence. For example, if the sentence "frame" were: "The party did not end until _____", possible responses might include "dawn", "three", "late", "midnight" and so forth. Don't try to be either unique or average; just be natural. You should keep within the following bounds, however: (1) Only one response word per sentence; (2) The word should "make sense" of the sentence and be from an appropriate class of words (nouns, verbs, adjectives, etc.); (3) English words only; (4) No proper names, hyphenated or contracted words; (5) Try to avoid repetitions. For some of the sentences the response will seem obvious; for others, several words may seem possible. Please print clearly.

The succeeding pages each contained 30 sentence contexts followed by a solid line (with the exception of the final page which contained the last 18 sentence contexts). Sentence order was the identical for all subjects.

A1b. Results

The probability of each response was calculated for each sentence separately for the two age groups. Appendix G lists the sentence contexts and the frequencies of their eliciting responses. The sentence contexts are presented in the order of decreasing contextual constraint. Similarly, the response words elicited by each sentence are listed in descending order of frequency of response. A frequency score was obtained for each subject indicating the number of sentence contexts (out of a possible 198) which were completed with best completions. This score, referred to as the best completion (BC) score, was the measure by which groups were compared.

Within the "younger" group, t-tests were performed to determine if differences in Age, Sex, SES, or ED showed any systematic effect on BC. The mean BC score was 178.1 (SD = 9.4, range = 151-192). Median scores were used to divide the sample into two groups for the comparison of BC scores according to high and low SES and ED. Age was also examined by comparing BC scores for the 20 oldest and 20 youngest subjects. In all cases, comparisons of BC scores along these four dimensions were not significant.

For the Young group, Pearson correlations of BC with Age, SES and ED are shown in table 2.

Table 2
Pearson correlations of BC with Age, SES, and ED and corresponding probability values for the younger group (N = 100).

| | Age | SES | ED |
|---|---|---|---|
| BC r= | .230 | -.290 | .290 |
| p= | .020 | .003 | .004 |

Although these correlations were significant, it should be noted that each accounted for less than 8% of the variance. Age did not correlate with SES or ED (the correlation of SES with ED (r = -.67) was significant (p = .001), which was expected since ED values are used in the calculation of SES).

Average BC score for the elderly group was 181 (SD = 5.9, range = 165-190). The difference between the young and elderly group on this measure was nonsignificant.

The Pearson correlation of Age with BC for the full sample (N = 130) was weak (r=.18). Although significant (p<.05), Age only accounted for approximately 3% of the variance.

The results of this norming study confirmed that the majority of sentence contexts were of high contextual constraint. Bloom and Fischler (1980) warned that the use of their sentence completion norms with populations other than college students should be done with caution. However, the current findings indicated that at least for highly constrained sentence contexts, older adults are as likely, or more likely, to complete a sentence with a Best Completion.

A2. Judgement Norms

A2a. Set 2 and Set 3

According to the experimental design, the terminal words of sentences in Set 2 (Related/Sense, "RS") and Set 3 (Related/Nonsense, "RN") were to be semantically related to the sentence contexts' best completions. Ideally, the selection of these words would have been based on the results of relatedness ratings, with relatedness values balanced between the two sets. Given a number of practical constraints, these procedures were not feasible. Instead, semantically related words were selected by the experimenter and colleagues, and a group of sentences terminating with semantically related words were created. Approximately half of these sentences were considered to "make sense" (RS), and half were considered to "not make sense" (RN). Since it was important for the experimental stimuli to be categorized quickly and easily as either "sense" or "nonsense", a norming study was conducted in order to eliminate sentences for which this classification would have been questionable.

A2a1. Method

A2a1.1. Subjects

Twenty seven volunteers participated in this norming procedure (mean age = 31.3, SD = 5.72, age range 22 - 40 years, mean ED = 16.6 years, SD = 3.4 years, ED range = 12 - 23 years, 10 male).

A2a1.2 Procedure

All subjects were given a 5-page booklet. The cover page provided a space for demographic information (i.e., age, sex, education background), and the instructions, which read as follows:

> Attached you will find a list of sentences. Please read each of the following sentences and determine whether they do or do not make sense. For the purpose of this list of sentences, we define "makes sense" as an event, image or situation which could occur in everyday life. For example, the sentence: Grandmother baked some bread., clearly makes sense. All of the following sentences will tend to make sense up until the final word. The sentence: Grandmother baked some cheese., does not end as most people would expect it to end, but it still does "make sense", that is, it could occur in everyday life. However, the sentence: Grandmother baked some socks., does not make sense. Although it is possible to *imagine* a situation in which the sentence might make sense, "imaginable" does not fit the criteria of "makes sense" for this list of sentences. Again, for this list, "makes sense" is defined as an event, image or situation which could occur in everyday life.

Sentences were presented in pseudorandom order such that no more than 5 sentences of one type occurred in sequence. However, this order was identical for all subjects. To the right of each sentence were two blank spaces in which subjects indicated their response (yes or no). All subjects completed 152 sentences.

A2a2. Results

The frequency of "yes" and "no" responses was tallied for each sentence. A total of 40 sentences was used for each condition. Ideally, only those sentences which elicited complete agreement (i.e., 100% either "yes" or "no") would have been chosen as experimental stimuli. However, since fewer than 80 sentences met this criterion, the best 40 sentences per condition were selected. These sentences with their respective judgement frequencies are shown in tables A2 (set 2: Related/Sense) and A3 (set 3: Related/Nonsense) of Appendix A.

A2b. Set 1 and Set 4

A similar procedure was employed for the construction of set 4 (Unrelated/Nonsense, "UN"). Nonsense sentences were constructed by assigning a best completion word from one sentence to a different (unrelated) sentence context. These and the best-completion stimuli (already established by the Cloze norm procedure) were combined in a list similar to that of set 2 and set 3. Subjects read each sentence and indicated whether the sentence made sense (yes) or did not make sense (no).

A2b1. Subjects

Ten volunteers (mean age 27.9, age range 22 - 40 years, mean ED = 16.6 years, SD = 3.4 years, ED range = 12 - 24 years, 6 male) participated in this phase of the norming study.

A2b2. Procedure

All subjects were given a 4-page booklet. The cover page provided a space for demographic information (i.e., age, sex, education background), and the instructions, which read as follows:

> On the following pages are a large number of sentences. Some of the sentences clearly make sense, and some clearly do not. Please read each sentence carefully. If you think the sentence makes sense, check the space marked "yes" to the the right of the sentence. If you think the sentence does not make sense, check the space marked "no". Please respond to all sentences.

The frequency of "yes" and "no" responses were tallied for each sentence. A total of 40 sentences were needed for each condition. The best 40 sentences (i.e., highest percentage agreement) per sentence type were selected. These sentences and their respective judgement frequencies are shown in table A1 (set 1: Best Completion/Sense) and table A4 (set 4: Unrelated/Nonsense) of Appendix A.

A3. Final Words

Average word length of final words in each of the four stimulus sets are shown in table 3.

Table 3
Average word length of final words by Sentence Type

| Sentence Type | Mean (SD) |
|---|---|
| Set 1: Best Completion | 5.450 (1.63) |
| Set 2: Related/Sense | 5.175 (1.69) |
| Set 3: Related/Nonsense | 5.450 (1.43) |
| Set 4: Unrelated/Nonsense | 5.250 (1.66) |

The results of a one-way ANOVA indicated that there were no significant differences in mean number of letters per final word among the four sets.

Final words were also assessed for differences in word frequency. These values, obtained from Francis and Kucera's (1982) word frequency norms are listed in table 4.

Table 4
Mean word frequency of final words by Sentence Type

| Sentence Type | Mean (SD) |
|---|---|
| Set 1: Best Completion | 96.3 (85.5) |
| Set 2: Related/Sense | 68.2 (77.6) |
| Set 3: Related/Nonsense | 60.5 (86.9) |
| Set 4: Unrelated/Nonsense | 70.3 (70.7) |

The results of a one-way ANOVA indicated that differences in average word frequency among sets were nonsignificant.

B. Part 2

The norming procedures detailed above resulted in four sets of highly constrained stimuli. However, the development of these stimuli was based exclusively on behavioral response patterns. Therefore, the objective of the second phase of the pilot study was to ascertain whether the stimuli and task conditions would generate the desired ERP gradients (i.e., least amount of negativity elicited by Best Completion sentences, greatest amount of negativity elicited by Unrelated/Nonsense sentences).

B1. Session 1

B1a. Method

B1a1. Subjects

Five right-handed young healthy subjects (4 female, 1 male, mean age = 25.6 years, range = 21 - 30 years old; mean ED = 16.8 years, range = 13 - 22; mean

SES 48, range = 20 - 88) with normal (20/20) or corrected to normal vision served as paid volunteers.

B1a2. Stimulus Presentation

Sentence frames and final words were presented horizontally (white on black) on a Macintosh Plus computer screen, subtending a vertical visual angle of .45 degrees, and horizontal angles ranging from .9 to 3.8 degrees. Each trial consisted the following sequence: 1) presentation of the sentence frame (e.g., "He mailed the letter without a"), 2) a 500 msec blank screen, and 3) the final word and a period to mark the end of the sentence (e.g., "stamp."; see figure 3). The duration of the sentence frame ranged from 1000-2800 msec, depending upon the reading rate of individual subjects (determined during a practice block). The sentence frame duration was made flexible to ensure that all subjects would see the final word after having read the sentence frame. In this way, slower readers would have sufficient time to read the entire sentence, and faster readers would not have time to become distracted. Once established, the sentence context duration was held constant across all experimental blocks. The interval between sentence onsets ranged from 3800 msec (when sentence duration was 1000 msec) to 4800 msec (when sentence duration was 2000 msec). The final word (to which the ERP was recorded) was presented for 300 msec for all subjects.

Four blocks of 40 sentences were presented (40 sentences per Sentence Type, 10 of each Sentence Type per block). Sentences were presented in a quasi-random order, such that no more that three sentences of the same Sentence Type occurred in sequence. All sentences and final words were presented only once during a session. All subjects viewed all 160 sentences.

## Stimulus Presentation

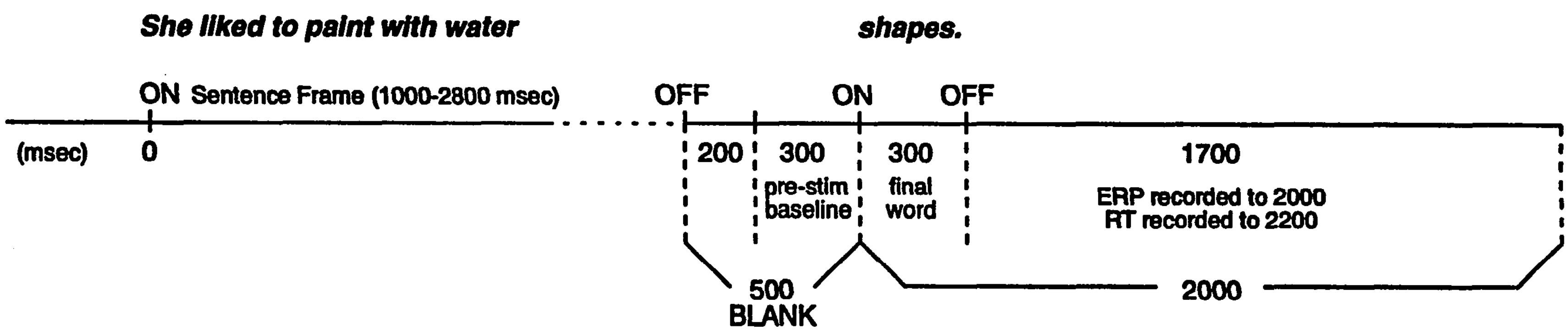



**Figure 3**

Time course of Stimulus Presentation and recording of response measures.

B1a3. Procedure

After electrode application, subjects were seated in a comfortable chair 30 inches away from the Macintosh plus video screen. Two response buttons were given to the subject (one in each hand), and the following instructions were read aloud:

> On the screen in front of you, you will see a series of sentences. Each sentence will be presented separately, and each sentence will be divided into two parts. First, you will see the beginning of the sentence, and then you will see the last word followed by a period. Read each sentence silently to yourself. Your task during the last word, is to decide whether or not the sentence makes sense. If the sentence does make sense, press the (right/left) button, and if the sentence does not make sense, press the (right/left) button. Try to respond as quickly and as accurately as possible.
>
> It is also important that you sit still and refrain from facial and eye movements while you ar doing the task. There will be two thin lines on the screen to help you focus. The sentence and the final words will always appear in between these two lines. If you must blink, try to blink only during the beginning of the sentence. First we'll run through a practice block. Do you have any questions?

Following the instructions, but prior to practice, subjects viewed 5-10 sentences on the screen (from the practice block) in order to determine the fastest comfortable rate of stimulus presentation. Subsequently, the experimenter left the room, and white noise was presented over a loud speaker to dampen extraneous noise from the outer room. Subjects initiated each block of trials (for both practice and test) by pressing both buttons simultaneously. The practice block consisted of 40 trials (20-sense, 20-nonsense, although these particular sentences did not meet all criteria for inclusion in experimental blocks). In order to proceed past the practice block, subjects were required to make no more than 3 errors each for "sense" and "nonsense" stimuli. After the practice block, the experimenter queried the subject to ensure that the rate of stimulus presentation was appropriate. When necessary, final adjustments were made. Time between blocks was approximately 2 minutes.

The hands assigned to the two responses were counterbalanced across subjects. Only reaction times between 200 and 2200 msec post stimulus were considered in the analysis.

B1a4. EEG Recording Procedures

EEG was recorded from midline electrodes at Fz, Cz, Pz, with lateral chains at F3, C3, P3, T5 and O1 on the left, and homologous locations on the scalp overlying the right hemisphere. Vertical (bipolar) EOG was recorded from electrodes located on the supra- and infra-orbital ridges of the right eye, and horizontal EOG was recorded from electrodes placed at the outer canthi of both eyes (bipolar recording). Except for the EOG electrodes, all leads were referred to nosetip. EEG and EOG were recorded with a bandpass of of .01 to 30 hz (5.3 sec time constant), with a Grass Neurodata system. Data acquisition and stimulus presentation were controlled by a PDP 11/34 computer, which digitized the EEG and EOG to the final word at 10-msec intervals for a 300 msec pre- and 1700 msec post stimulus onset period separately for each final-word presentation, and stored the digitized records, choice responses, and reaction times on 9-track digital tape. EEG was corrected for intrusion of EOG artifact using linear regression procedures (Gratton et al., 1983; Verleger et al., 1982). Blink and vertical EOG artifact were corrected first. In a second pass, horizontal EOG artifact was corrected following removal of vertical artifact from the horizontal EOG channel. The vertical and horizontal transmission coefficients were computed separately for each trial. Prior to computing averaged voltage measures of ERP activity, the individual averages were digitally filtered off-line at 15 hz using the low-pass filter described by Ruchkin and Glaser (1978).

ERP deflections were measured as averaged voltages falling within a specified time window with respect to the voltage of the 300 msec prestimulus baseline. Time windows were selected by visual inspection of within group averaged ERPs for each Sentence Type. This procedure is detailed in Methods section.

Behavioral measures analyzed included reaction time and error scores.

B1a5. Data Analysis

Effects of experimental variables were assessed for behavioral and ERP data by ANOVAs using the BMDP4V (Dixon, 1983) program. Where appropriate, F ratios were tested using degrees of freedom computed from an application of the Geisser- Greenhouse correction (included in BMDP4V program) for heterogeneity of variance-covariance matrices associated with repeated measures according to the methods of Jennings and Wood (1976). Corrected degrees of freedom were truncated to the nearest integer. Unless otherwise specified, homogeneity of variance was tested by Levene's test (BMDP7D) which consists of an ANOVA of the standard deviations of the different levels (e.g., sentence types) of a variable. Planned comparison t-tests with Bonferroni's correction (BMDP7D) were used to test a priori hypotheses (i.e., the effect of Sentence Type). Newman-Keuls post-hoc tests were used to make comparisons not specified prior to data collection. When significant interactions occurred, simple effects procedures (Winer, 1971), with corrected degrees of freedom where appropriate, were applied to further explore the interactions.

One-way ANOVAs were used to assess the effect of Sentence Type on N400 amplitude and behavioral measures (e.g., reaction time and error scores.

Analyses of within-subject Standard Deviations of reaction time are shown in Appendix E; Analyses of Homogeneity of variance are detailed in Appendix F.)

B1b. Results

B1b1. Behavioral Data

B1b1.1. Reaction Time

Only correct responses were included in reaction time (RT) analyses. Figure 4 shows mean RT as a function of Sentence Type separately for each subject. Mean RT across subjects for each Sentence Type are presented in table 5.

Table 5
Mean RT by Sentence Type

| | BC | RS | RN | UN |
|---|---|---|---|---|
| RT (msec) | 770 | 1132 | 955 | 840 |
| (SD) | (101) | (139) | (141) | (82) |

Since Levene's test indicated equality of variances among Sentence Types, the RT data were not transformed. RT was ordered in the following manner: Best Completion (1) < Unrelated/Nonsense (4) < Related/Nonsense (3) < Related/Sense (2). The effect of Sentence Type was significant ($F(3/16) = 8.91$, $p<.001$). Pairwise comparisons indicated that RT to Related/Sense was significantly delayed relative to all other Sentence Types ($p<.03$), which were not statistically differentiated from one another.

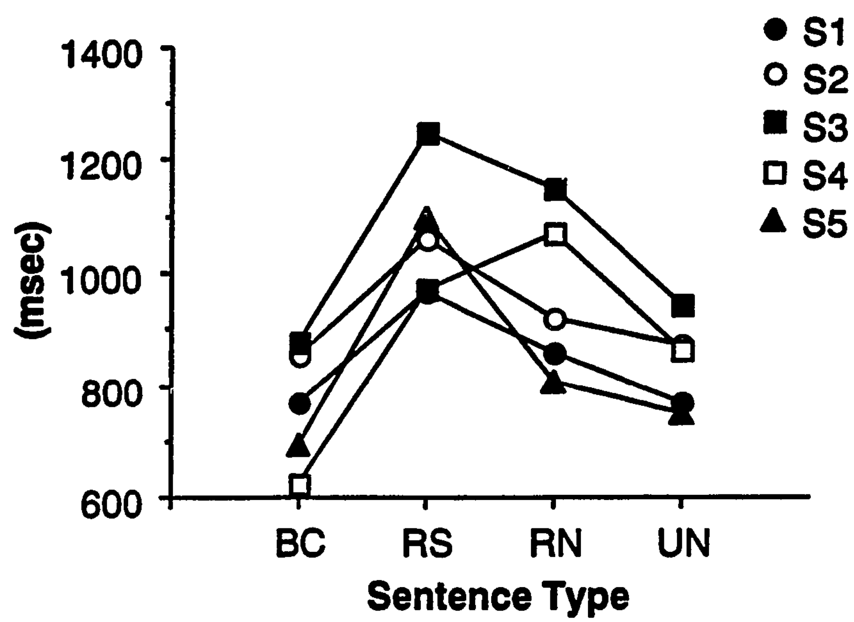



**Figure 4**

Mean individual Reaction Time: Pilot Subjects, Session 1, by Sentance Type

B1b1.3. Accuracy

Table 6 lists the number of errors per Sentence Type for each subject. In each subject, Related/Sense clearly generated the greatest number of errors. This result was unexpected since the majority of sentences chosen for the experiment were those that caused minimal ambiguity regarding a sense/nonsense decision (i.e., sentences which generated consensus or near consensus during norming procedures). However, experimental conditions, i.e., time pressure to respond within 2200 milliseconds, differed from those during the norming procedure, in which subjects worked at their own pace.

Table 6
Error scores by Sentence Type

| Subject | BC | RS | RN | UN |
|---|---|---|---|---|
| S1 | 0 | 4 | 0 | 0 |
| S2 | 0 | 13 | 1 | 1 |
| S3 | 1 | 16 | 1 | 3 |
| S4 | 2 | 9 | 2 | 1 |
| S5 | 0 | 12 | 0 | 0 |
| Total | 3 | 54 | 4 | 5 |
| Mean | 0.6 | 10.8 | 0.8 | 0.8 |

As evident from table 6, comparisons of means indicated that there were more errors committed to Related/Sense relative to all other Sentence Types ($p<.001$). Best completion, Related/Nonsense and Unrelated/Nonsense were not differentiated on this measure.

B1b2. ERPs

Averaged ERP waveforms for midline and lateral electrodes are shown in figure 5. Each waveform represents the averaged ERP elicited by the last word of each of the four Sentence Types. As can be seen in figure 5, the ERPs were characterized by an early negativity, peaking at approximately 200 msec (N1: 150-250 msec), maximal at lateral posterior electrodes. For Related/Sense, Related/Nonsense and Unrelated/Nonsense, a second negative deflection (identified as N400) was present at all electrodes, onsetting at approximately 250 msec, peaking at approximately 400 msec at the midline, and at 450 msec at the lateral posterior electrodes (T5, T6, O1, O2). Since the principal hypothesis involved specific predictions regarding the differentiation of N400 as a function of Sentence Type, N400 was subdivided into three segments, (N4a, N4b, N4c) in order to maximize the probability of detecting the window during which the Sentence Types were highly differentiated.

N4b encompassed the middle portion (350-450 msec) of the negative deflection observed at the midline. N4a (250-350 msec) and N4c (450-550 msec) included the early and late portion of the total N400 interval (N4tot; 250-550 msec). For Related/Sense, the ERP remained negative (relative to waveforms for the other three Sentence Types) for several hundred msec. For Best completion, Related/Nonsense, and Unrelated/Nonsense, the N400 region was followed by a broadly distributed positive deflection at approximately 600 msec (P600; i.e., classical P300), which was largest to Unrelated/Nonsense. It was not assumed that the positivity elicited by Best Completion sentences during the N400 interval is the classical P300. Rather, it is most likely similar to the P380 (Ritter, Simson, and Vaughan, 1988), or the P341 (observed in children; Friedman, Vaughan and Ehrlenmeyer-Kimling, 1978), a large positive potential associated with visual stimuli whose

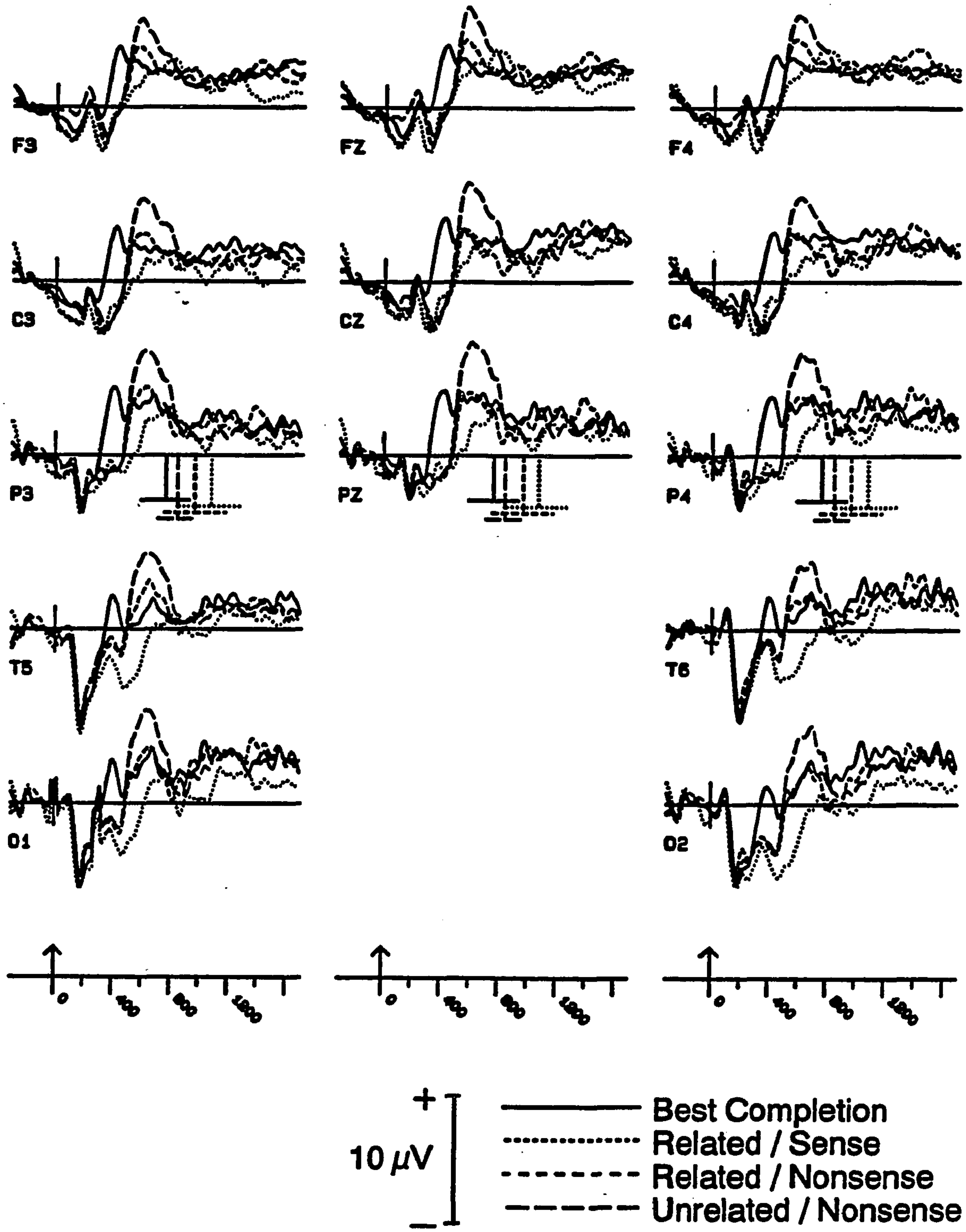



**Figure 5**

Grand mean ERPs by Sentance Type
Pilot Study, Session 1

latency and amplitude are not affected by manipulations that typically modulate classical P300. The P380 was not observed in the ERPs of the other Sentence Types probably due to the overlapping N400. Since N400 is typically elicited by unprimed words, it was not anticipated that Best Completions would elicit an N400.

Of primary interest in the present study was the relative degree of negativity elicited by each Sentence Type. Therefore, for simplicity, this ERP interval will be referred to as N400, although this does not imply that there was an N400 elicited by Best Completions. The latter portion of the ERPs for all Sentence Types was characterized by a long duration, positive slow wave.

Because the specific aim of the pilot study was to assess whether the task and stimuli would elicit an N400 amplitude gradient as a function of semantic relatedness, only analysis of the N400 component of the ERP will be detailed. The N400 region was quantified as the averaged voltage from 250-500 msec post stimulus onset.

It was anticipated that N400 amplitude would be smallest (or absent) to Best completions (1), moderate to Related/Sense (2) and Related/Nonsense (3), and largest to Unrelated/Nonsense (4). As can be seen in figure 5, this gradient was not observed. Although the effect of Sentence Type across all electrodes was significant for N4b ($F(1/6)= 5.83$, $p<.02$) and approached significance for N4c ($F(1/6)= 5.95$, $p<.055$), the expected gradient was not obtained[3]. N400 (N4b) amplitudes for each Sentence Type are shown in table 7.

---

[3] Levene's test indicated inequality of variances across Sentence Types for N400 amplitude, even after square root and log transformations. This was most likely due to the small sample (N=5) employed in the pilot study.

Table 7
Mean N400 amplitude across all electrodes in microvolts by Sentence Type.

| | BC | RS | RN | UN |
|---|---|---|---|---|
| N4b | 1.15 | -2.5 | -2.2 | -2.1 |

Planned comparisons indicated that although Best Completion was statistically distinct from the other Sentence Types ($p < .05$), Related/Sense, Related/Nonsense, and Unrelated/Nonsense were not differentiated from one another.

Visual inspection of the data from individual subjects indicated that 1 of the 5 subjects did show the expected N400 gradient. However, these data patterns were not reflected in the group averaged ERPs, and it remained unclear as to why the manipulation of Sentence Type did not produce the predicted N400 gradient in most subjects. In an effort to understand why, the data were re-averaged and analyzed in the following ways.

B1c. RT quartiles

Inspection of the behavioral data together with ERP data revealed that Best completion elicited the most rapid RT, fewest number of errors, and smallest N400 amplitude, whereas Related/Sense elicited the slowest RT, greatest number of errors and largest N400 amplitude compared to the other three Sentence Types. Given this pattern, it seemed plausible that N400 amplitude might vary directly with degree of difficulty. It was also important to note that N400 and P3 often occur in overlapping latency intervals. Under these circumstances, it can be difficult to tease apart the relative contributions of each component. It was possible that the delay in RT, most evident for

Related/Sense was accompanied by a delay in P3 (P3 latency and RT are correlated under certain conditions; Ritter et al., 1979), and that N400 elicited by Related/Sense only appeared larger due to a delayed P3. Consequently, the observed N400 amplitudes could be fully accounted for by degree of difficulty as reflected by RT, which might have been accompanied by corresponding alterations in the latency of P3. This possibility was explored by re-averaging the ERPs into quartiles as a function of RT, regardless of Sentence Type (see figure 6 As expected, the first quartile is disproportionately weighted with ERPs from Best Completion (i.e., shortest RTs) and ERPs from Related/Sense are clustered in the last quartile. However, as can be seen in table 8, responses from each Sentence Type are represented in each quartile.

Table 8
Decomposition of RT quartiles: Number of sentences from each Sentence Type per quartile, mean RT, and N400 amplitude for each quartile.

| RT Quartile | 1st | 2nd | 3rd | 4th |
|---|---|---|---|---|
| Best Completion | 80 | 35 | 25 | 17 |
| Related/Sense | 4 | 16 | 32 | 58 |
| Related/Nonsense | 19 | 43 | 42 | 52 |
| Unrelated/Nonsense | 27 | 60 | 54 | 14 |
| Mean RT (msec) | 657 | 792 | 918 | 1214 |
| N400 (μV) | 1.9 | -1.6 | -1.4 | -1.6 |

The N400 latency interval used for these analyses was 250-450 msec. As can be seen in table 8 and figure 6, RT increased substantially as a function of quartile whereas N400 amplitudes for the last three quartiles were quite similar. Accordingly, although the effect of Quartile was significant for both

RT (F3/56)=52.71, p<.001) and N400 amplitude (F3/56)=4.62, p<.01), post hoc tests indicated that mean RTs for each quartile were statistically distinct from one another (p<.001), whereas N400 amplitudes only differentiated the first and latter three quartiles (p<.05). These analyses did not support the position that N400 amplitude was fully attributable to changes in P300 latency associated with level of difficulty as reflected by RT.

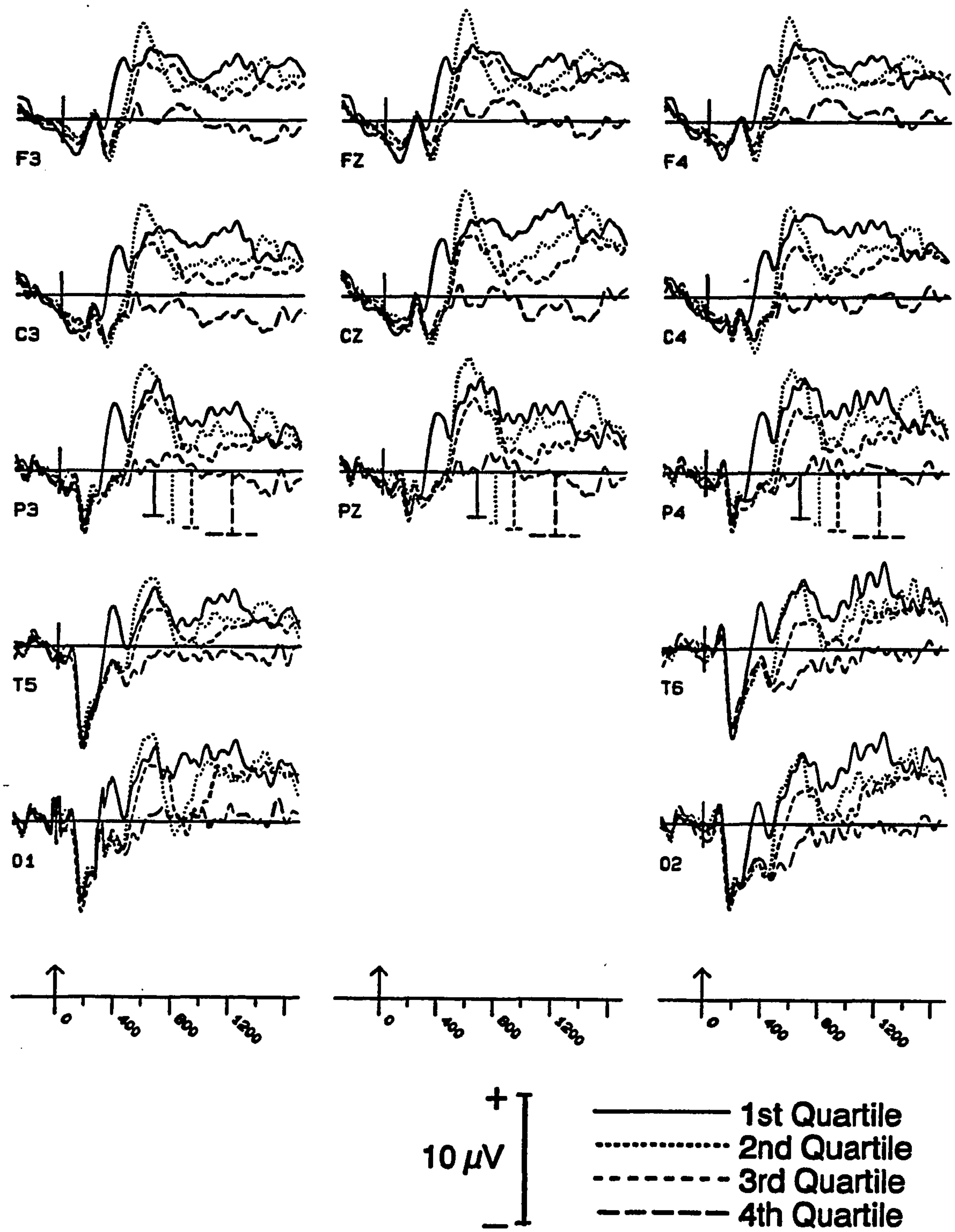



**Figure 6**

Grand mean ERPs by Reaction Time Quartiles regardless of Sentence Type, Pilot Study, Session 1

B1d. First Half versus Second Half

It was conceivable that during the initial part of the experiment response patterns may have differed in comparison to those of the latter part of the experiment. Practice can alter the processes by which the subject performs the task (Shoben, 1982), e.g., increased experience with the task may lead to a refinement of strategy and/or improved efficiency in responding. For instance, at the beginning of the experiment, the instructions created a clear expectation for Best completion and Unrelated/Nonsense (i.e., "sense", "nonsense"). By contrast, Related/Sense and Related/Nonsense were not directly addressed in the instructions, and therefore, their appropriate categorization was cognitively more demanding. It is possible that only after experience with the task, subjects developed both an expectation for all Sentence Types, and a "strategy" for categorizing Related/Sense and Related/Nonsense (This is discussed in greater depth below). These changes which may have occurred during the course of the experiment could have influenced N400 amplitude. To explore this possibility, behavioral and ERP responses were re-averaged separately for the first and last two blocks of the experiment.

Based on visual inspection of these data, it appeared that both behavioral and ERP responses changed during the course of the experiment. To quantify these observations, the data were subjected to statistical analyses. However, given the small sample size (N = 5) together with the reduced number of stimuli per Sentence Type (i.e., only 20 trials per Sentence Type within each half of the experiment), it was anticipated that the experimental effects might not show statistical significance.

B1d1. Reaction Time

As can be seen in table 9, RT was more rapid during the second half of the experiment for all Sentence Types except Related/Nonsense.

Table 9
RT by Sentence Type for the first and second half of the experiment.

| | First Half | Second Half | RT Difference |
|---|---|---|---|
| BC | 785 (171) | 745 (173) | -40 |
| RS | 1076 (225) | 1065 (200) | -11 |
| RN | 960 (200) | 966 (196) | +6 |
| UN | 835 (102) | 845 (133) | +10 |

Reaction Time in msec; Difference in RT between first and second half shown in right column; SD in parentheses.

Although the main effect of Sentence Type on RT was significant, ($F(2/8)= 17.93$, $p<.001$) the effect of Half and the interaction of these two variables was not. Nonetheless, this pattern of change (i.e., faster "sense" responses, slower "nonsense" responses) suggest a shift toward stricter criteria for "nonsense" decisions.

B1d2. ERPs

Figure 7a and 7b show mean ERPs for the first two blocks and the last two blocks of trials, respectively. By visual inspection, it appeared that by the second half of the task, the increased negativity elicited by Related/Sense (and to a lesser extent by Related/Nonsense) was diminished, and the expected gradient had emerged. This pattern is also represented in the difference waveforms ([Related/Sense - Best completion], [Related/Nonsense - Best completion], and [Unrelated/Nonsense - Best completion]; see figure 8a and

8b). N400 amplitude values for N4b by Sentence Type for each Half are shown in table 10.

Table 10
N400 amplitude (N4b) in µV by Sentence Type for the first and second half of experiment

| Sentence Type | First Half | Second Half |
|---|---|---|
| Best Completion | 3.4 | 3.7 |
| Related/Sense | - 3.5 | 0.1 |
| Related/Nonsense | -2.7 | -0.9 |
| Unrelated/Nonsense | -1.1 | -2.3 |

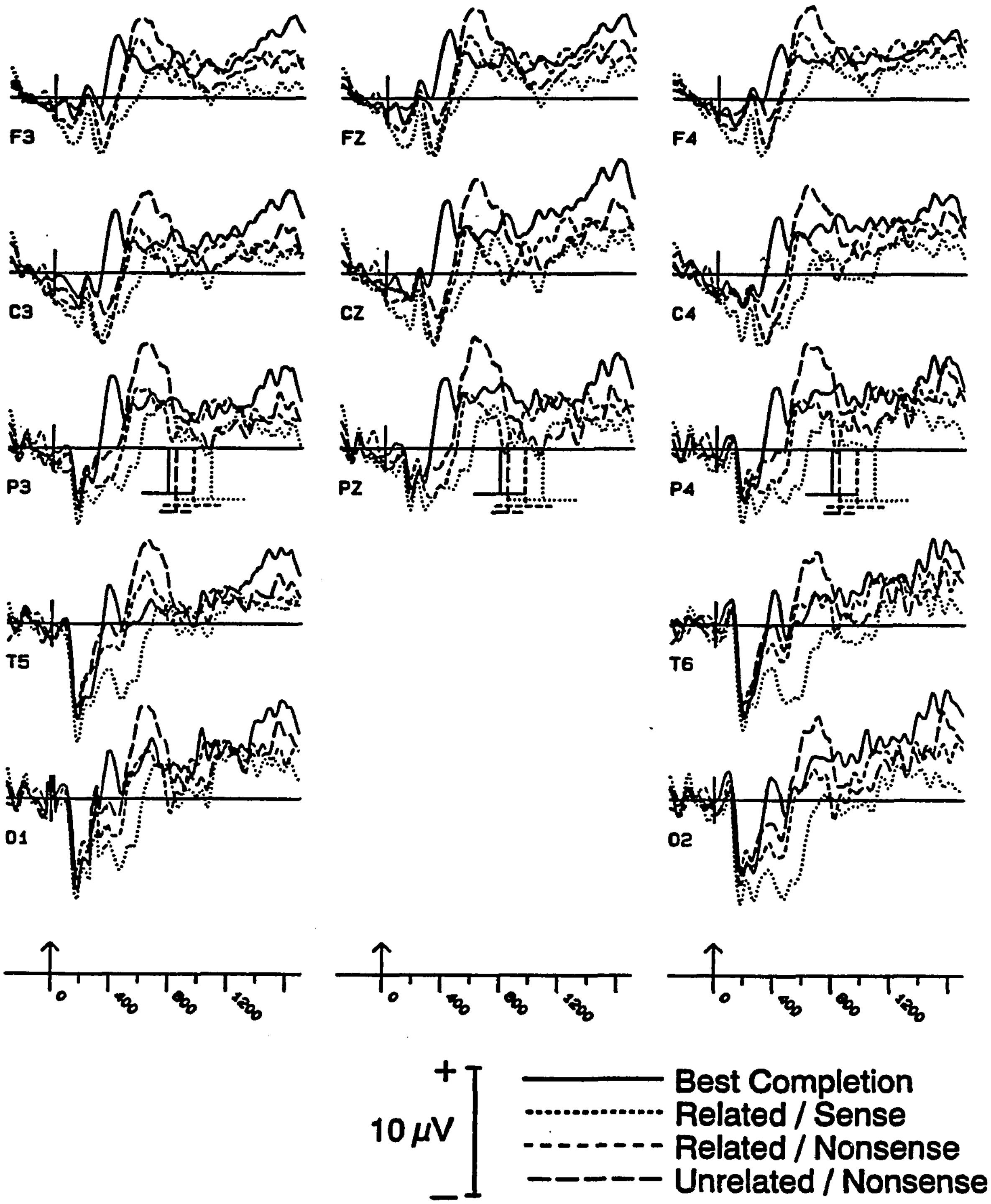



**Figure 7a**

Grand mean ERPs by Sentence Type, Pilot Study
(a) Session 1, First Half

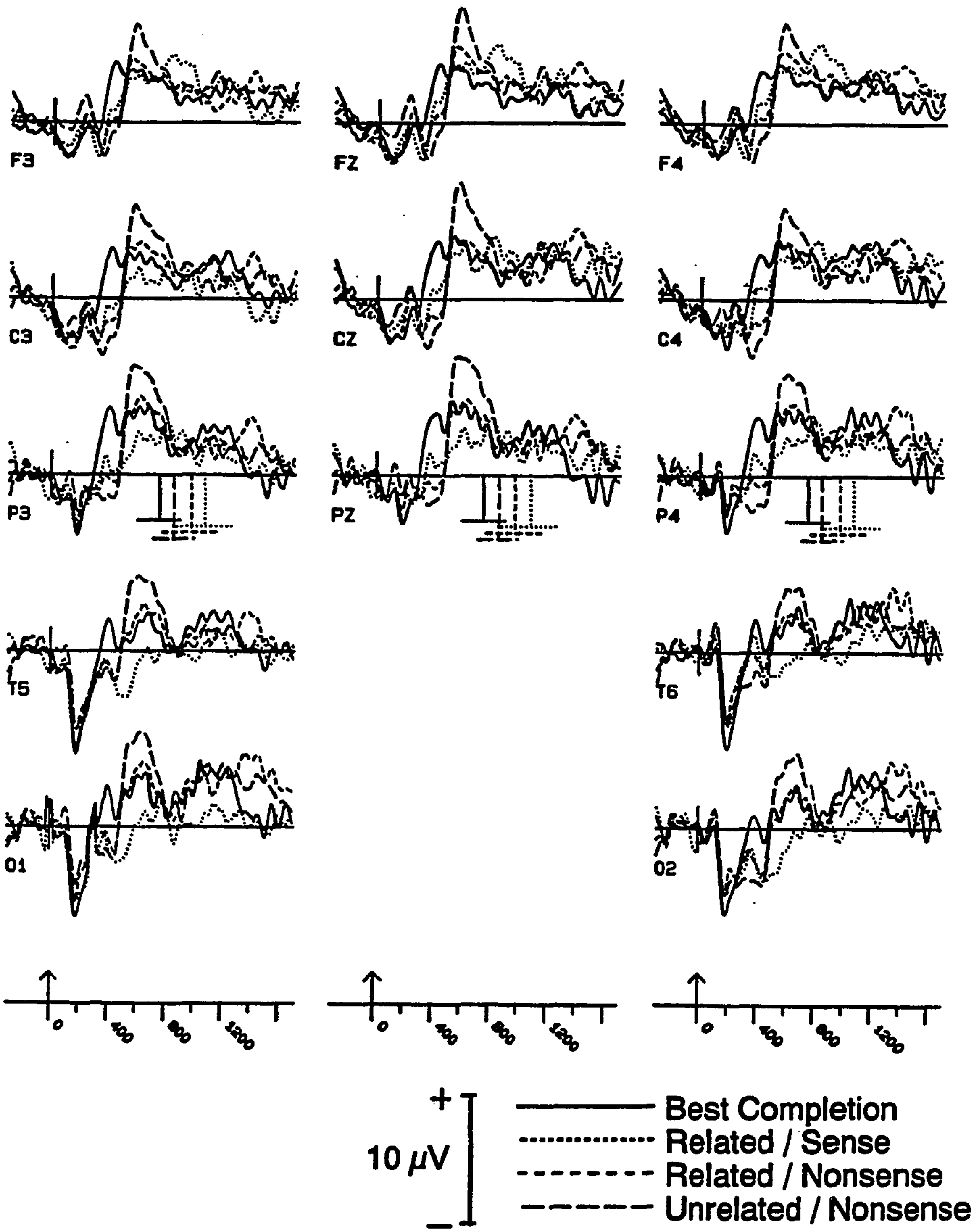



**Figure 7b**

Grand mean ERPs by Sentence Type, Pilot Study

(b) Session 1, Second Half

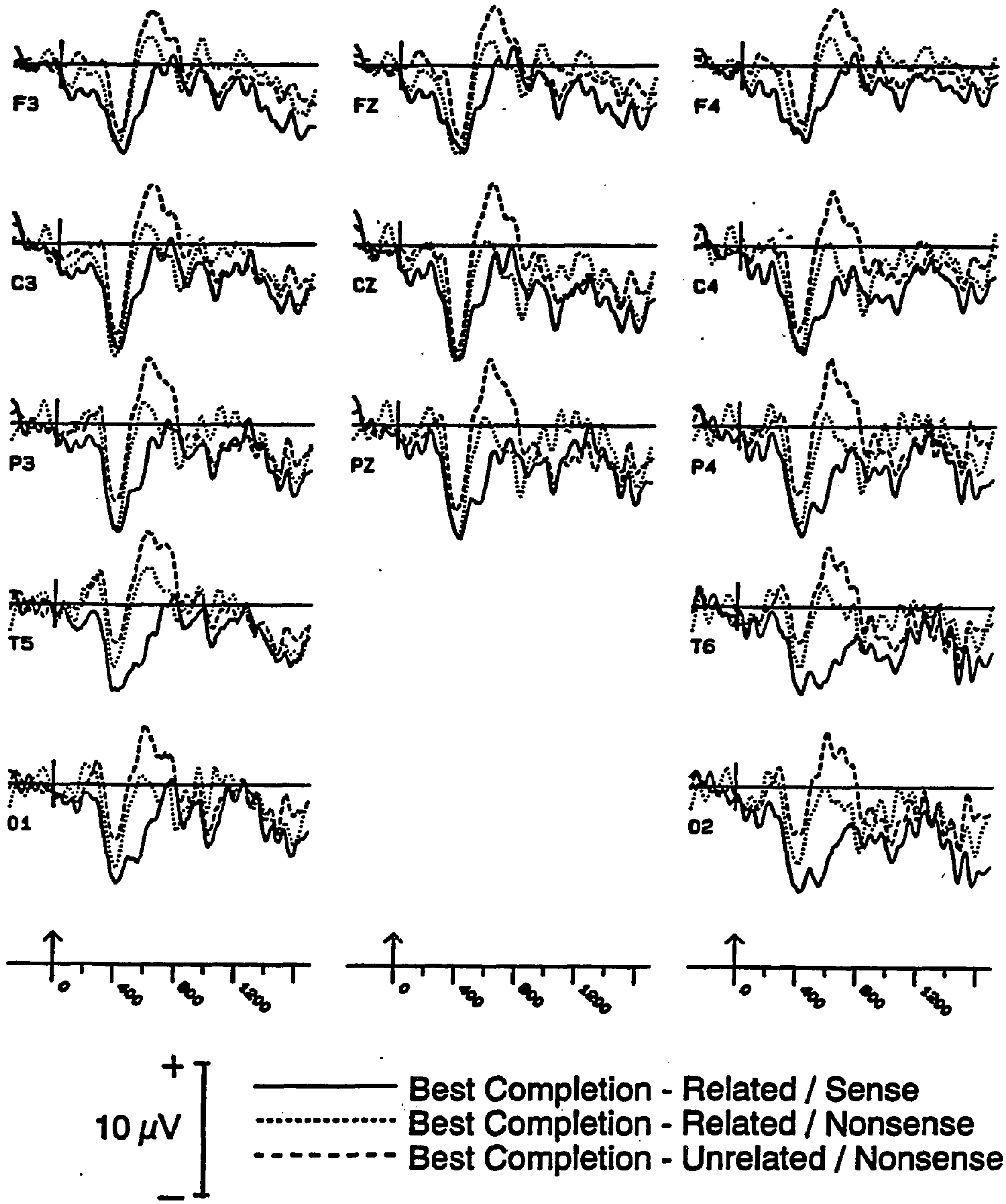



**Figure 8a**

Difference waveforms, Pilot Study,
(a) Session 1, First Half

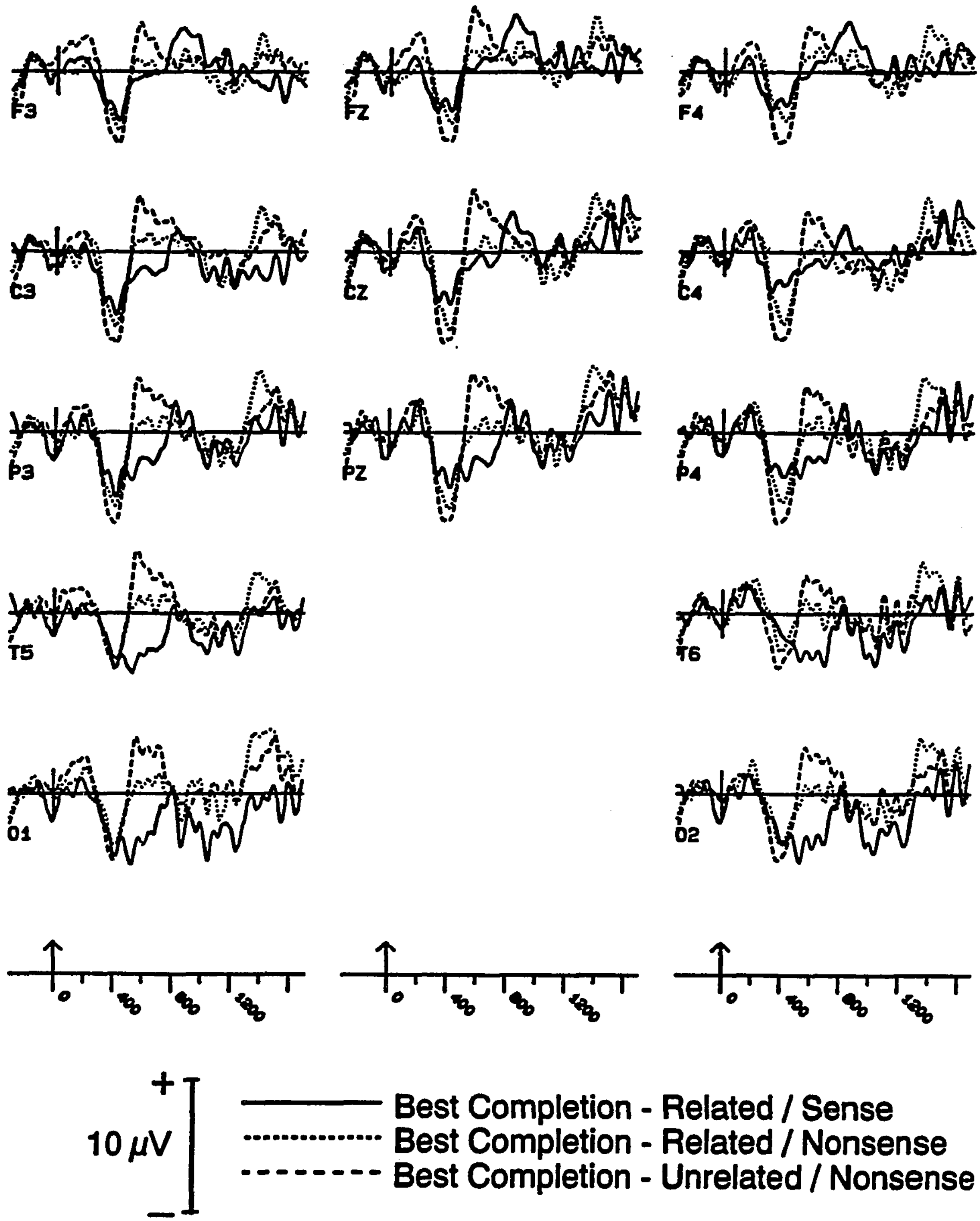



**Figure 8b**

Difference waveforms, Pilot Study,

(b) Session 1, Second Half

Across halves, the main effect of Sentence Type was significant only for the N4b segment (latency interval identified as 350-450 msec; $F(3/12)= 5.5$, $p<.02$). By visual inspection, the ordering of Sentence Types clearly shifted by the second half of the experiment. However, most likely due to the explanations offered above, the interaction of Half and Sentence Type did not reach significance.

Although the cause of enhanced negativity to Related/Sense in the beginning of the task was not clear, it appeared that it was no longer operative by the second half of the session. It is possible that one practice block did not provide sufficient experience with the task. Furthermore, inspection of the practice stimuli indicated that the majority of these sentences resembled Best completion and Unrelated/Nonsense stimuli. Thus, the practice block may have created a strong expectation for these particular Sentence Types. After two experimental blocks which included Related/Sense and Related/Nonsense, subjects may have become familiar with and developed an expectation for these other sentences. Consequently, subjects may have adopted a different strategy by the second half of the experiment.

Although other investigators have shown N400 gradients in response to semantic relatedness, these particular stimuli in combination with the response requirements produced a unique set of cognitive demands. It is possible that the expected "small" N400 for Related/Sense was overlapped by an additional (presently unknown) strategy-related negative ERP component, which was no longer active after sufficient practice.

B2. <u>Session 2</u>

Regardless of the cause, the results of the pilot study suggested that the predicted N400 amplitude gradient could be obtained after subjects had acquired additional experience with the task. Prior to the last 2 blocks of stimuli, subjects had already participated in 1 practice and 2 experimental blocks. This implied that three (120 trials) blocks would provide subjects with sufficient practice such that N400 amplitude would show the expected gradient. To test this hypothesis, the experiment was rerun. The procedure was identical to the original session, with the exception that 2 additional practice blocks were administered. An effort was made to use sentences that more closely resembled those used in the experimental blocks (i.e., all four Sentence Types were included, see Appendix B). In order to proceed past the practice blocks, subjects had to respond correctly on at least 8 out of 10 trials for Best completion and Unrelated/Nonsense, and at least 7 out of 10 trials for Related/Sense and Related/Nonsense.

All 5 subjects from the original pilot study were retested within 8-12 weeks of their initial testing. Since these data would be confounded by prior experience with the same task and stimuli, six naive right-handed subjects were also recruited (mean age = 26 years old, age range = 23 - 29 years; mean ED = 18.7 years, range = 16 - 22; mean SES 45.8, range = 20 - 51). The data from these two groups were analyzed separately.

B2a. Results

Figure 9 shows the averaged ERPs for "Session 2" of the original 5 subjects (N400 latency interval identified as 300 - 450 msec: N4a 300 - 350 msec, N4b 350 - 400 msec, N4c 400 - 450 msec). By visual inspection, the N400 amplitude gradient was obtained (see N4b amplitude values, table 11). Accordingly, the effect of Sentence Type was significant for N4b ($F(2/7) = 4.91$, $p<.05$) and N4c ($F(2/6) = 4.98$, $p<.05$). Planned comparison t-tests indicated that N400 elicited by Best Completion and Related/Sense were distinct from

Related/Nonsense and Unrelated/Nonsense. Thus, unlike the first half of session 1, N400 amplitudes were ordered as expected.

Table 11
N4b amplitude in μV across all electrodes by Sentence Type
Session 2

| Sentence Type | N4b Amplitude |
|---|---|
| Best Completion | 0.6 |
| Related/Sense | -0.5 |
| Related/Nonsense | -2.3 |
| Unrelated/Nonsense | -3.9 |

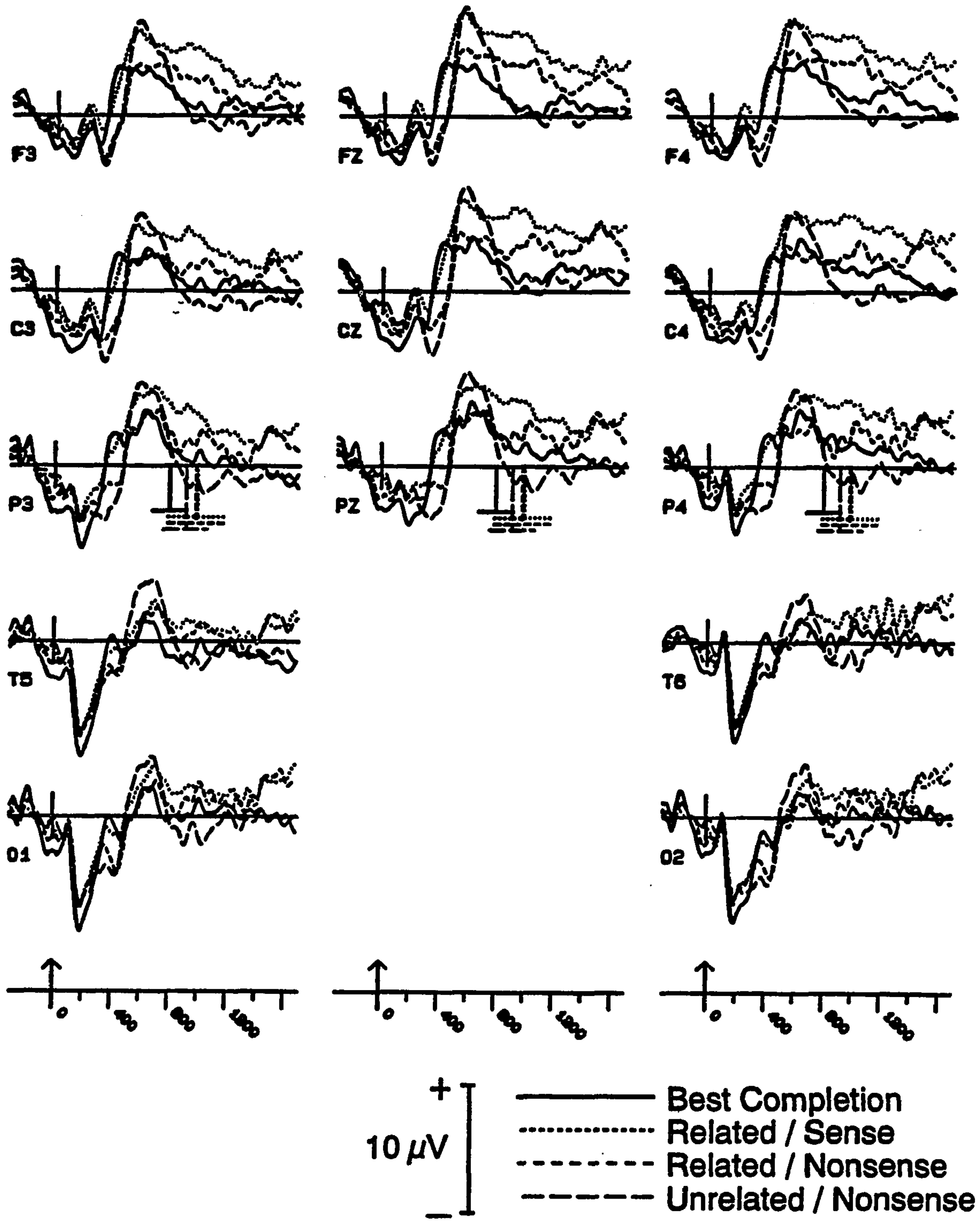



**Figure 9**

Grand mean ERPs by Sentence Type
Pilot Study, Session 2

B2b. Session 1 versus Session 2

The behavioral and ERP changes which occurred from Session 1 to Session 2 were examined via 2-way ANOVAs (Session by Sentence Type).

B2b1. Behavioral data

Reaction time data as a function of Session and Sentence Type are shown in table 12.

Table 12
Mean RTs as a function of Session and Sentence Type

| | BC | RS | RN | UN |
|---|---|---|---|---|
| Session 1 | 770 (196) | 1132 (303) | 955 (249) | 840 (156) |
| Session 2 | 794 (209) | 996 (309) | 999 (269) | 917 (224) |

Although it appeared that RT patterns changed from Session 1 to Session 2, only the main effect of Sentence Type was significant ($F(2/7) = 19.31$, $p<.001$). Neither the effect of Session nor the interaction of Session and Sentence Type reached significance. However, in comparisons of each Sentence Type across Sessions (matched sample t-test) the change in RT for Related/Sense approached significance ($p<.06$). Although the statistical evidence is only suggestive, these data seem to reflect a strategy change associated with Related/Sense from Session 1 to Session 2.

As can be seen in table 13, error patterns changed markedly across Sessions. Compared to Session 1, Related/Sense elicited fewer errors and Related/Nonsense elicited a greater number of errors. The interaction of Session and Sentence Type was significant ($F(3/14) = 8.59$, $p<.001$), and comparisons of means confirmed that error scores for Related/Sense and Related/Nonsense changed significantly ($p<.05$) from Session 1 to Session 2. These differences together with the changes in RT most likely reflect a shift in strategy.

Table 13
Error patterns in Sessions 1 and 2

| | Session 1 | | | | Session 2 | | | |
|---|---|---|---|---|---|---|---|---|
| Subject | BC | RS | RN | UN | BC | RS | RN | UN |
| S1 | 0 | 4 | 0 | 0 | 0 | 5 | 11 | 0 |
| S2 | 0 | 13 | 1 | 1 | 1 | 8 | 1 | 0 |
| S3 | 1 | 16 | 1 | 3 | 3 | 5 | 6 | 1 |
| S4 | 2 | 9 | 2 | 1 | 3 | 0 | 4 | 0 |
| S5 | 0 | 12 | 0 | 0 | 0 | 18 | 3 | 0 |
| Total | 3 | 54 | 4 | 5 | 7 | 18 | 25 | 1 |
| Mean | 0.6 | 10.8 | 0.8 | 0.8 | 1.4 | 3.6 | 5.0 | 0.2 |

Session 1 (left) and Session 2 (right): Number of errors per Sentence Type by subject, total, and mean across subjects.

B3. Session 3

Grand mean ERPs for the 6 new subjects are shown in figure 10. As shown, N400 shows a progressive increase from Best completion to Unrelated/Nonsense. Analysis of these data also revealed similar patterns to those of "Session 2"; The effect of Sentence Type on N400 across all electrodes was significant ($F(3/68) = 5.3$, $p<.002$) and comparisons among Sentence Types indicated that Best completion and Related/Sense were significantly more positive relative to Related/Nonsense and Unrelated/Nonsense (further distinctions were not found, most likely due to insufficient power as a consequence of the small sample size). These data are elaborated further in the Results section, since these subjects were included in the larger experiment.

C. Summary

Since other investigators have shown N400 gradient effects in young normal subjects under conditions similar (but not identical) to those proposed here, it was necessary to determine whether the present particular combination of stimuli and task requirements would also elicit this ERP pattern. The initial session of the pilot study resulted in an unexpected ERP pattern in which final words related to the "target" word (Related/Sense) elicited the largest negativity in the N400 region. Further analysis suggested that this unexpected ERP pattern was strategy related. For instance, when the ERP data were examined separately for the first and second half of the experiment, the desired N400 gradient was observable for the second half. Reductions in RT for three Sentence Types suggested increased efficiency during the course of the experiment. Further evidence of a strategy shift involved marked changes in error patterns when the initial pilot subjects were retested with the additional 80 practice trials. With this additional practice, both the retested group of subjects and a group of new subjects showed the hypothesized N400 gradient.

Having determined the set of conditions that would reliably elicit a systematic gradient in N400 amplitude as a function of semantic relatedness (to a primed word) in young normal subjects, these could then be applied to other experimental groups with reasonable confidence that any deviations from this ERP pattern would be attributable to differences in the cognitive system engaged in the task.

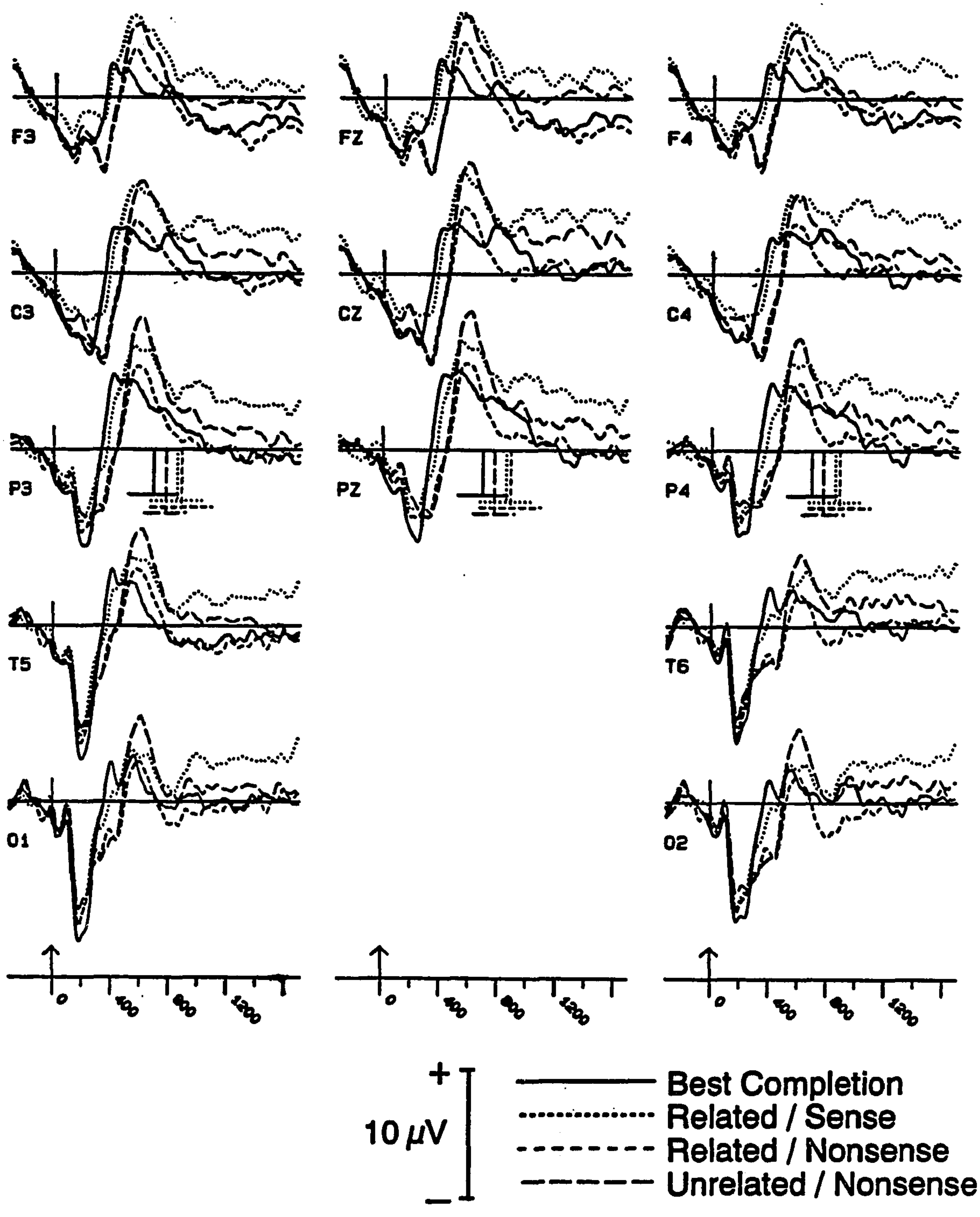



**Figure 10**

Grand mean ERPs by Sentence Type
Pilot Study, New subjects

# III. HYPOTHESES

Evidence from the literature, in combination with the results from the pilot study support the notion that semantic memory is functionally organized according to type and strength of semantic relationships. It has been reliably demonstrated (e.g, lexical decision tasks, sentence verification tasks, production frequencies, etc..) that the presentation of a word/concept elicits activation of its semantic associates. When the semantic context is sufficiently constrained, a particular target word becomes activated with the greatest strength relative to other words within the network. This activation moves out along the network in a decreasing gradient as a function of semantic relatedness, (i.e., a close semantic associate. e.g., "bird" for "robin", will be activated with greater strength relative to a more distant associate, e.g., "animal" for "robin"). This graded response is reflected in both response time and N400 amplitude.

If:

1) semantic memory is organized according to semantic relationships,
2) activation spreads outward along the network in a decreasing gradient,
3) decomposition of the network (as in PAD) proceeds in a predictable fashion, and
4) the sentence contexts developed in the pilot study serve as a relatively invariant prime of a particular target.

then the following predictions can be made.

There should be a main effect of Sentence Type on measures of N400 amplitude, RT and Percent Correct. Based on most reports in the literature (e.g., Nebes, 1989) it is expected that RT will be slowest and error rates will be highest for the PAD group relative to both Young Normal (YN) and Normal Elderly (NE) groups. However, this effect should be modified by a significant Group by Sentence Type interaction, indicating a different ordering of Sentence Types within groups. Specifically, response patterns in Young Normal and Normal Elderly groups should be similar to each other, yet distinct from those of the PAD group.

A. Young Normal and Normal Elderly

1) YN and NE should show similar behavioral and electrophysiological orderings of Sentence Types, reflecting a functionally intact semantic network:

a) N400 Amplitude:

Planned comparisons between adjacent Sentence Types should reveal a pattern in which N400 amplitude is smallest to Best Completion, intermediate to Related/Sense and Related/Nonsense (semantically related to the best completion), and largest to Unrelated/nonsense (unrelated to the best completion).

b) Behavioral:

In typical semantic priming paradigms (e.g., lexical decision tasks, word naming tasks), RT facilitation is expected for semantic associates of target words. However, the sense/nonsense decision requirement of the present task, superimposed upon the presentation of differentially primed words requires a different set of predictions. The present sense/nonsense decision is somewhat similar to the true/false decision of sentence verification tasks (see

Introduction: Semantic Memory). This analogy, in combination with the results of the pilot study suggest the following set of predictions:

1) RT:

Planned comparisons should indicate that RT is most rapid to Best Completion (similar to "fast true" effect), followed by Unrelated/Nonsense (similar to "unrelated-false" effect), since "yes" responses are typically faster than "no" responses. RT to Related/Sense (similar to "atypical-true" effect) and Related/Nonsense (similar to "related-false" effect) should be delayed relative to Best Completion and Unrelated/Nonsense.

2) Accuracy:

In accord with RT, planned comparisons should indicate a greater percentage of accurate responses to Best Completion and Unrelated/Nonsense compared to that of Related/Sense and Related/Nonsense.

B. PAD

1) Response patterns in PAD subjects should reflect the decomposition of semantic organization.

a) N400 amplitude:

1) Similar to YN and NE subjects, planned comparisons should indicate larger N400 amplitude to Unrelated/Nonsense relative to that of Best Completion. However, if N400 amplitude reflects the

loss of discriminability between semantically related items, then N400 elicited by Related/Sense and Related/Nonsense (both semantically related to the Best Completion) should be indistinguishable from that elicited by Best Completion.

2) Alternatively, based on the position that semantic organization is intact in PAD although access and use are impaired, ERP patterns may show a dissociation from behavioral responses, such that the N400 gradient will appear similar to that of normal controls, but RT and and accuracy would be dissimilar.

b) RT:

1) As with N400, RT patterns of PAD patients for Best Completion and Unrelated/Nonsense should be similar to that predicted for YN and NE groups. However, if semantically related items lose their discriminability, Related/Sense and Related/Nonsense items might rapidly be mistaken as equivalent to Best Completion items. In this case, there may not be substantial differentiation among Sentence Types as reflected by RT.

2) Alternatively, since Related/nonsense is related to the Best Completion but does not make sense, Related/nonsense may cause confusion and consequently delay RT. In this case, RT to Related/Nonsense will be significantly slower relative to that of other Sentence Types.

c) Accuracy:

If Related/Sense stimuli are regarded as equivalent to Best Completion, this would not be revealed in terms of accuracy since

Best Completion and Related/Sense require an identical response (i.e., "sense"). However, since Related/Nonsense items are also related to the sentences' best completions but require a "nonsense" response, it is likely that PAD patients will commit significantly more errors to Related/Nonsense compared to other Sentence Types.

# IV. METHOD

## A. Subjects

All subjects were administered the Modified Mini Mental Status Exam (MMS; Mayeux, et al.,1981; maximum = 57), the Vocabulary subtest of the Weschsler Adult Intelligence Scale-Revised (WAIS-R), and the Boston Naming Test. Mean test scores and demographic data for all three groups (Young Normals, Normal Elderly, and PAD) are shown in table 14. All subjects were native English speakers.

Table 14

Demographic characteristics and test scores of the 3 samples

| | YN | NE | PAD |
|---|---|---|---|
| N = | 10 | 10 | 6 |
| AGE (mean) | 26.1 | 67.0 | 67.1 |
| (SD) | 3.0 | 4.7 | 3.4 |
| EDUC (mean) | 18.1 | 14.1 | 16.2 |
| (SD) | 2.1 | 2.2 | 2.8 |
| MMS (mean) | 56.8 | 54.6 | 39.7 |
| (SD) | 0.7 | 1.3 | 6.2 |
| SES (mean) | 49.1 | 57.3 | 38.7 |
| (SD) | 11.3 | 18.8 | 21.9 |
| DF (mean) | 7.6 | 6.7 | 5.7 |
| (SD) | 0.4 | 1.1 | 1.0 |
| DB (mean) | 6.1 | 4.7 | 3.8 |
| (SD) | 0.7 | 1.4 | 1.7 |
| Boston | 15.0 | 14.4 | 13.8 |
| (SD) | 0.0 | 1.3* | 2.2 |
| WAIS-R VOC | 58.6 | 55.9 | NA |
| (SD) | 6.6 | 5.5 | |

SES= Socio-economic status (higher score = lower SES)
MMS= Modified Mini-mental status exam (highest score is 57)
DF = digits forward from MMS; DB = digits backward from MMS
Educ = number of years of education
Boston= 15 selected items from Boston Naming Test
WAIS-R VOC= Raw Score WAIS-R Vocabulary subtest
NA = not available

*Mean based on 7 subjects

A1. Normal Controls

Ten Young Normal subjects ("YN"; 5 male, 5 female, mean age = 26.1 years old) and 12 Normal Elderly subjects ("NE"; 7 male, 3 female, mean age = 67.0 years old) participated in the present experiment. YN and NE subjects were recruited from the community via local newspaper advertisements and posted notices in the Columbia Presbyterian Hospital vicinity. For inclusion in the study, subjects achieved a score of at least 53 on the MMS, and scored within normal limits on the WAIS-R Vocabulary subtest and Boston Naming tests, according to appropriate age norms. In addition to these criteria, subjects in the NE group were free of dementia, depression, and were not limited in the activities of daily living as determined by the Short CARE (Gurland et al., 1984), an extensive structured interview. All normal control subjects received payment for their participation.

Due to excessive artifact during the experimental procedure, data from two NE subjects could not be used.

A2. Probable Alzheimer's Disease patients

Eleven Probable Alzheimer's Disease (PAD) patients were recruited from the Alzheimer's Disease Research Center (ADRC) of Neurological Institute at the Columbia Presbyterian Medical Center. In addition to the tests listed above, PAD patients were administered a larger neuropsychological battery as part of their participation at the ADRC during a previous session (although within five months prior to the present experiment), which included the Similarities subtest of the WAIS-R, Buschke's Selective Reminding Test (SRT; Buschke,1973), Selected subtests of the Wechsler Memory Scale (WMS, Wechsler, 1981), Verbal Fluency tests (CFL, Animal Naming), and the Bentin Visual Retention Test (BVRT, Bentin,1955). Other causes of cognitive decline were ruled out by normal results on: computed tomography, complete blood count, serum B-12, hepatic renal and thyroid function tests, tests for syphilis and folic acid levels. The diagnosis of PAD, determined by the Neurologist and Neuropsychologist on staff at the ADRC, was based on a combination of the NINCDS ADRDA criteria (McKhann et al., 1984), neuropsychological test results, and the patient's level of functioning in daily activities. Only patients who were characterized as "mild" according to the following operationalized criteria were chosen for participation in the study: MMS scores were required to be no lower than 34 and no higher than 49, duration of illness between 6 months and 3 years, and Dementia Rating (Hughes et al., 1982) of "mild."

Due to excessive artifact during the experimental procedure, data from five PAD patients could not be used.

B. Criteria for exclusion of subjects

The acquisition of artifact-free ERP measurements requires that subjects reduce muscle tension and refrain from facial, eye and bodily movements. Consequently, even in the best of circumstances, it is typical for a portion of the acquired data to be eliminated from analyses due to artifactual contamination. Although the eye-blink correction procedure employed in the present study reduces this data loss, subjects were eliminated from the study if on more than approximately 70% of the total number of trials a) eye blinks were present and/or b) artifactual movements were excessively large and led to blocking of the amplifiers.

B1. PAD subjects: Additional considerations

Given the methodology's sensitivity to artifact, combined with the limited ability of many PAD patients to restrict movements, a disproportionate number of patients were either discontinued during the experiment, or excluded from data analysis. Thus, excessive artifact was one factor contributing to the difficulty in acquiring a sufficiently large sample for the PAD group. Additionally, the task was cognitively demanding. Although designed to examine semantic processing, the task also involved memory (e.g., retention of task instructions, right/left button press, etc.) and attentional skills (e.g., ability to sustain attention and concentration) which are impaired in Alzheimer's Disease (Joynt and Shoulson, 1985). In addition to the processing load of the task itself, it was necessary for subjects to inhibit otherwise natural responses such as blinking, speaking, and moving. Although a total of 11 PAD patients participated in the protocol, the data from five patients were excluded due to excessive movement artifact. Thus, only six PAD patients were able to successfully complete the experimental procedure.

Within the PAD group, five subjects had received a firm diagnosis of PAD. However, at the time of testing, one subject, "JM", scored higher on a number of neuropsychological tests (e.g., Full Scale IQ = 126, "Above Average", as assessed by the Wechsler Adult Intelligence Scale- Revised; Memory Quotient = 116, "Above Average" as assessed by the Wechsler Memory Scale) relative to the other patients, and did not meet all standard criteria for "dementia". Yet, performance on other tests was indicative of cognitive decline (e.g. impaired delayed visual memory as assessed by the WMS, poor delayed verbal memory as assessed by Buschke Selective Reminding Test, impaired verbal fluency, difficulty with diadikokinetic movements and competing programs, body part naming, and right/left orientation) which was consistent with the family's report of functional decline in daily activities. Although patients are not typically seen for evaluation this early in the dementing process, JM's family was particularly attuned to subtle indications of dementia since a close family member had recently been diagnosed with Probable Alzheimer's Disease. Thus, although not technically "demented", the consensus of the clinical team was that JM was in the early stages of Alzheimer's-type Dementia, and that within the year (when he would return for reevaluation) it was extremely likely that he would meet standard criteria for dementia.

Consequently, the inclusion of JM's data in the analyses was questionable. In the Results section , the behavioral data and ERP waveforms are presented both with (PAD, N=6) and without (PAD, N=5) this subject. However, the analyses detailed in the Results section include all 6 subjects in the PAD group. Analyses excluding JM are presented in Appendix B. As can be seen in both the behavioral data (figures 11 a and 11b) and ERP waveforms (figures 12 and 15), the patterns were relatively unaffected by the exclusion of the JM's data. However, in the ERP analyses, the inclusion of these data increased statistical power, enabling a number of experimental effects to reach significance.

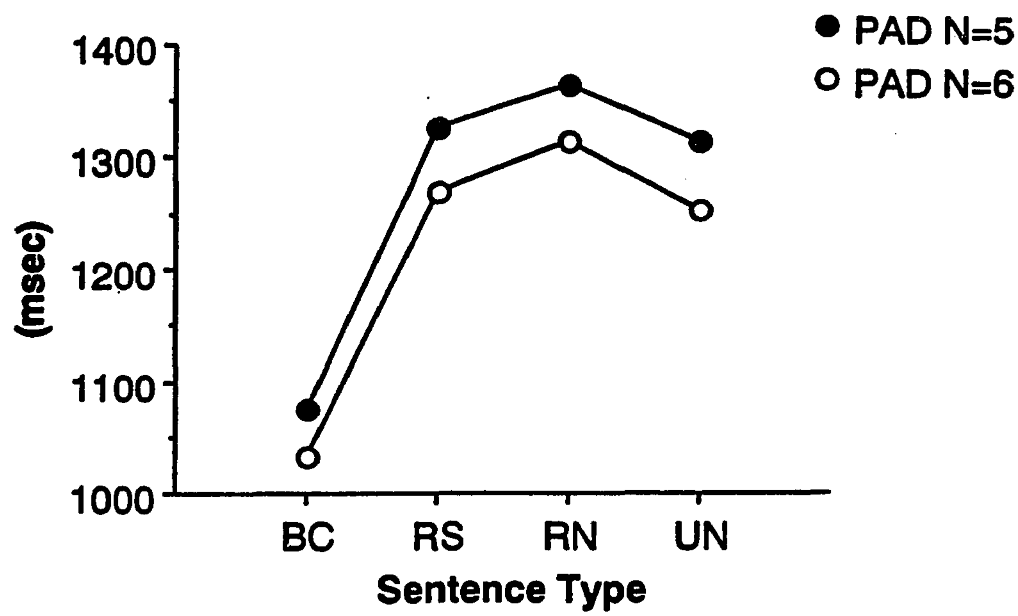



**Figure 11a**
PAD, N=5 versus N=6, Mean RT by Sentence Type

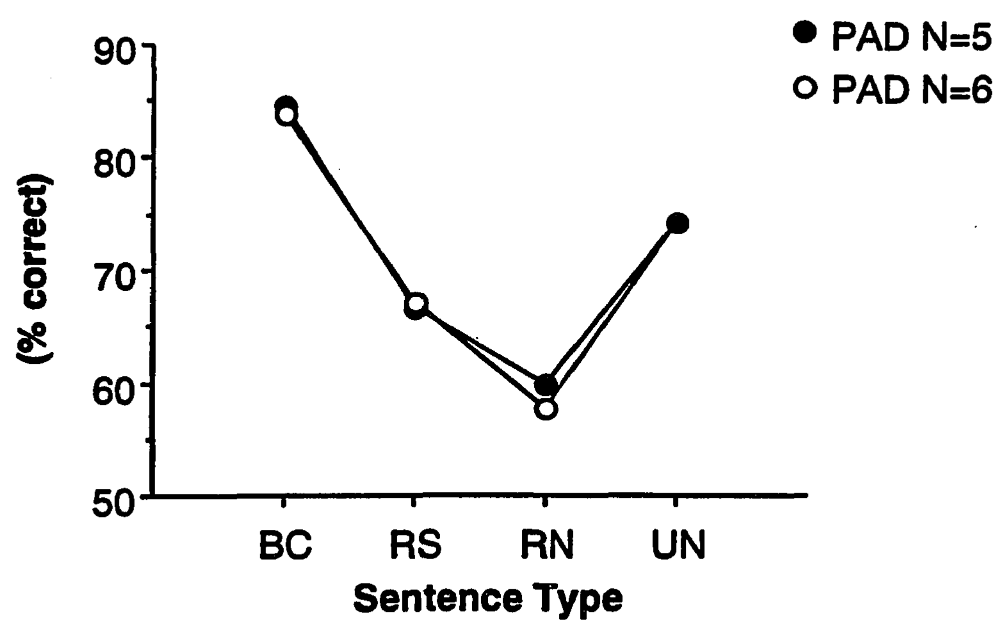



**Figure 11b**
PAD, N=5 versus N=6, Mean Percent Correct by Sentence Type

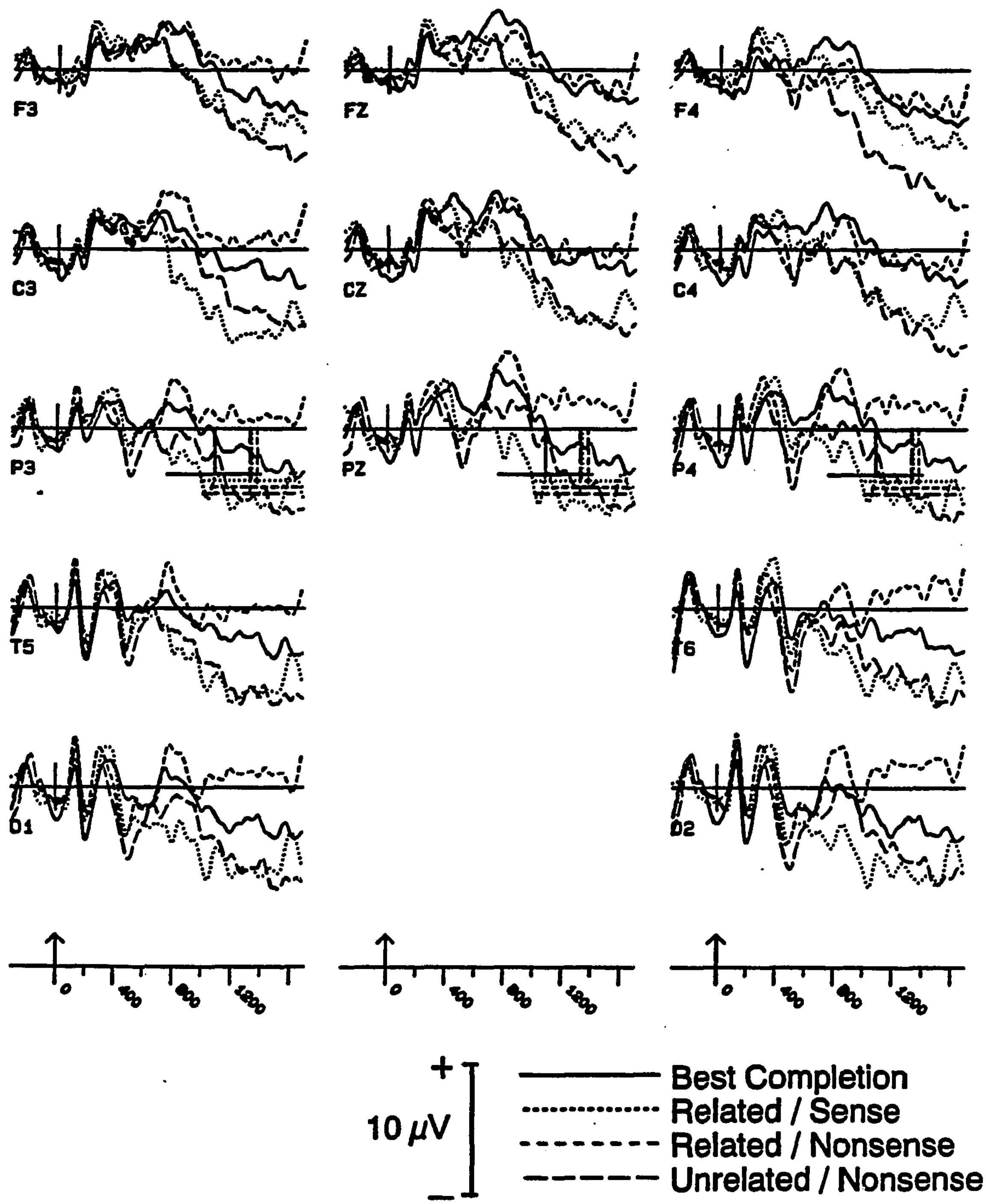



**Figure 12**

Grand mean ERPs by Sentence Type, PAD group, n=5

C. Stimuli

The sentences used in the study were identical to those used in the Session 2 of the pilot study.

D. Stimulus presentation

Stimulus presentation was identical to that described for the pilot study. However, since PAD patients are cognitively impaired, the maximum duration that could occur for sentence-context presentation was increased to 3000 milliseconds. During practice trials, this duration was titrated for each subject to the shortest duration possible. Once established, the particular sentence-context duration was maintained throughout the experimental procedure.

E. Procedure

For screening purposes, the MMS was administered to YN and NE subjects prior to the experimental procedure. However, the WAIS-R Vocabulary subtest and the Boston Naming test were administered subsequent to the experiment, in order to avoid potential confounds such as word repetition effects and/or interference from words semantically related to critical target words included in the experimental stimuli. PAD subjects had already undergone all neuropsychological testing at Neurological Institute prior to the experiment (detailed above).

For YN and NE subjects, the procedure was identical to that described for Session 2 of the pilot study. However, PAD patients required additional personal attention throughout the experiment. At the beginning of each experimental block, it was necessary to remind all PAD subjects of the assignment of response buttons (i.e., sense, nonsense). Instructions were read several times, and practice blocks were repeated until artifactual movement was minimal (i.e., eyeblinks were not time locked to the target words or button presses, or excessively large, which would block amplifiers) and response accuracy was at least 70% for Best Completion and Unrelated/Nonsense, and 60% for Related/Sense and Related/Nonsense. Most PAD subjects required 4-5 practice blocks. Although additional practice may have further improved performance, these less stringent criteria were accepted in order to reduce the risk of inducing frustration and fatigue in PAD subjects.

All YN and NE subjects were left alone in the subject room during each block of trials. However, the experimenter remained with all PAD subjects for the duration of the task, in order to repeat task instructions and to remind patients to remain still when necessary.

F. EEG Recording Procedures

These procedures were identical to those described in the Pilot study.

G. Criteria for Exclusion of data

The present task and stimuli were carefully constructed in order to induce a particular type of processing for each stimulus type (i.e., sentence type). Although the particular items within a set are unique, the responses elicited by stimuli within a set were averaged together based on the assumption that the stimuli within each set elicited similar cognitive operations. Generally, when RTs are either excessively long or responses are incorrect, it is less likely that the processes underlying these responses meet these assumptions. Therefore, trials which elicited incorrect responses, or

RTs prior to 200 msec or beyond 2200 msec post stimulus were excluded from RT and ERP analyses and ERP waveforms.

H. Data Analysis

N400 effects are typically reported to be largest at midline electrodes (Fz, Cz, Pz) and lateral electrodes located at T5 and T6. Generally, N400 effects during sentence priming tasks have been reported to be larger over the right hemisphere (Kutas and Hillyard, 1988). Although there were no particular expectations of electrode effects or hemispheric differences with respect to the current hypotheses, these variables were included for comparison with previous literature.

Other than these exceptions, the Results section includes only the behavioral and ERP (i.e., N400) analyses which were relevant to specific hypotheses. All other analyses (i.e., N1, P300, SW1, SW2, Hemisphere analyses) are reported in Appendix C.

I. Component indentification

ERP deflections were measured as averaged voltages falling within a specified time window with respect to the voltage of the 300 msec prestimulus baseline. Time windows were selected separately for each group by visual inspection of averaged ERPs for each Sentence Type.
N1 was identified by its topographic distribution (largest posteriorly) and early peak latency. For N400, it was anticipated that in all groups, Unrelated/Nonsense would elicit the greatest negativity in the N400 region relative to other Sentence Types. Thus, N400 was selected by identiftying the most negative deflection (associated with Unrelated/Nonsense) at Cz and Pz during the interval between 250 and 550 msec. As described in the Pilot study,

this total interval was defined as N4tot, the middle portion, N4b, and the early and late segments as N4a and N4c respectively. Since the size of the interval for N4a and N4c depended upon the of the rise and fall times of the deflections, these segments did not necessarily span equal time intervals. For instance, if the deflection showed a rapid onset but a slow return to baseline, N4a was shorter than N4c. P600 was defined as the region of positivity subsequent to the N400. The latency at which this positivity showed a slight negative trough defined the onset of the Slow Wave interval, which was divided into two equal segments (SW1 and SW2).

J. Analyses

Procedures for evaluating effects of experimental variables are described in the Methods section of the Pilot Study. Planned analyses were as follows:

1. Four three-way ANOVAs, i.e., Group (YN, NE, PAD) by Sentence Type (ST1, ST2, ST3, ST4) by Electrode (Fz,Cz, Pz, F3, C3, P3, T5, O1,F4, C4, P4, T6, O2), were performed for amplitude measures of each segment of N400 (N4tot, N4a, N4b, N4c):

2. Three two-way ANOVAs, i.e., Group by Sentence Type, were performed for each of the behavioral variables: RT and Percent Correct.

K. Scaling Procedures

When assessing group or condition differences in the scalp distribution of ERPs, an ANOVA could yield a significant interaction (e.g., Group/Condition by Electrode Location) which is due to genuine differences in amplitude at different scalp locations. However, in assessing differences in shape or distribution across electrodes, this interaction would be misleading, (McCarthy and Wood, 1985). In the present study, when examining shape differences irrespective of differences in magnitude, the data were scaled. This involved a computation in which the root mean square (RMS) amplitudes of the ERP averaged voltages from each group (or condition) averaged across subjects were made the same (i.e., magnitude differences were removed), (McCarthy and Wood, 1985). An ANOVA was then performed on these scaled data. Although this procedure is typically used in comparing topographic distributions across the scalp, it can also be used in assessing group/condition differences in shape of other functions as well, e.g., shape of RT or N400 amplitude across Sentence Types.

# V. RESULTS

Demographic data and test scores are shown in table 14. There were no significant differences in SES between groups. As expected, MMS scores were significantly different across groups ($F (2/20) = 61.68, p<.0001$). Post hoc procedures indicated that Young Normals and Normal Elderly did not differ from one another, although both groups scored significantly higher compared to PAD patients ($p<.05$).

## A. Event Related Potential Waveforms

Averaged ERP waveforms from the full electrode montage are shown in figures 13,14, and 15 for the YN, NE, and PAD groups, respectively. Although the waveforms showed differences as a function of Group and Sentence Type, their general morphologies were similar. As can be seen, the waveforms were characterized by an early positivity (P1) largest posteriorly (most clearly visible at O1/O2), followed by a negative deflection peaking at about 200 msec (N1), also largest posteriorly (most clearly visible at T5/T6, O1/O2), followed by a second negative deflection (N400) that was broadly distributed across the scalp. This deflection was not apparent for Best Completion, but was evident for other Sentence Types. As in the Pilot Study, the peak latency was earlier at the midline (approximately 350 msec) compared to lateral posterior sites (approximately 425 msec) in the Young Normal group. However, this did not characterize the Normal Elderly and PAD waveforms, as the negativity at midline and lateral posterior sites showed similar latency. As detailed in the pilot study, the N400 region was divided into three segments. Again, the positivity associated with Best Completion was not considered to be the P3 or the N400, but for simplicity, this region will be referred to as the N400. The N400 was followed by a large-amplitude, positivity (P600) with a widespread scalp distribution, peaking at approximately 600-700 msec. The remaining portion of the waveform was characterized by a long duration slow wave, (positive relative to baseline in the YN and NE groups, but negative relative to baseline in the PAD group) that extended to the end of the recording epoch. For measurement purposes, this activity was partitioned into Slow Wave 1 (SW1) and Slow Wave 2 (SW2). Table 15 lists the latency intervals determined for each component within each group.

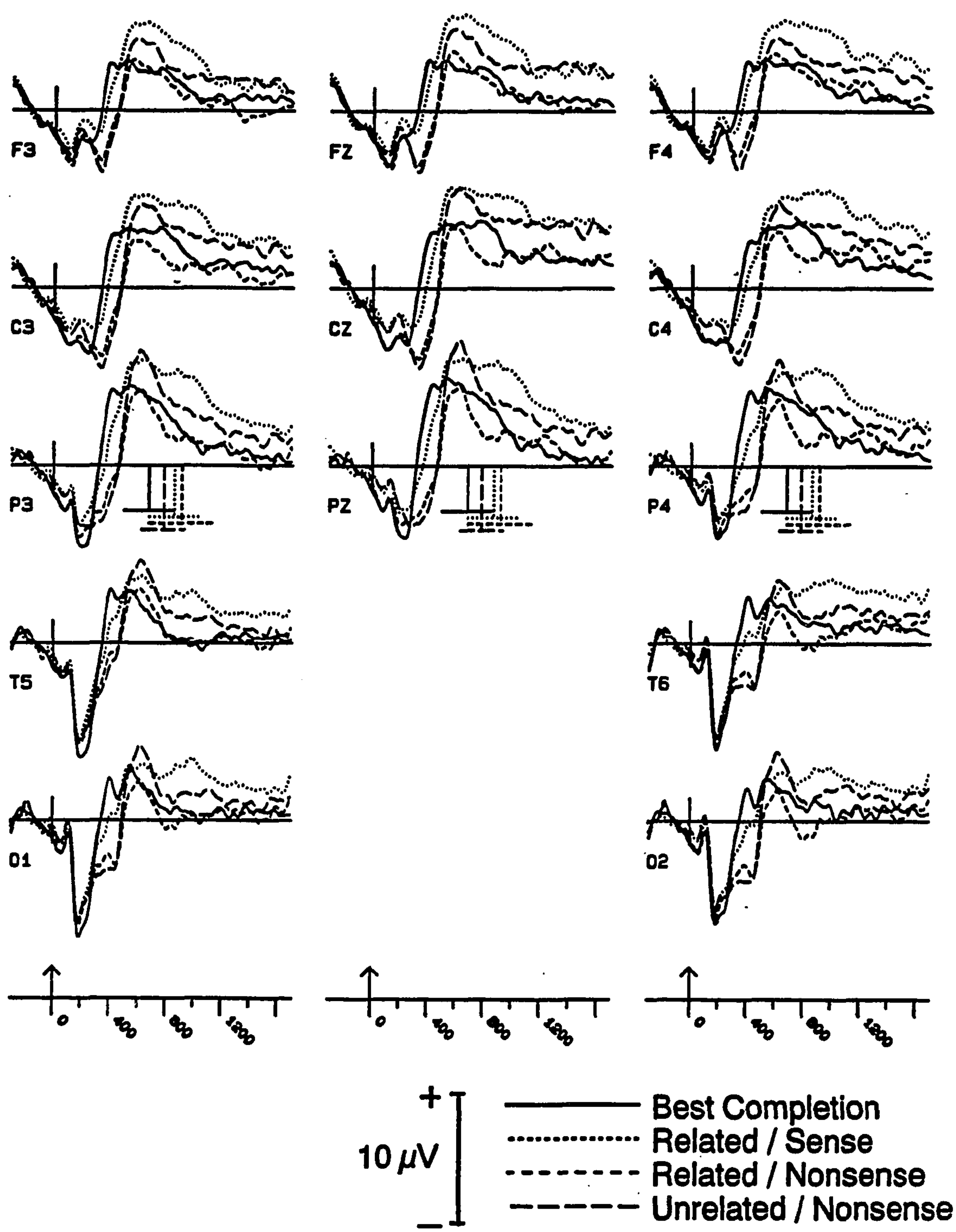



**Figure 13**

Grand mean ERPs by Sentence Type,
Young Normal group

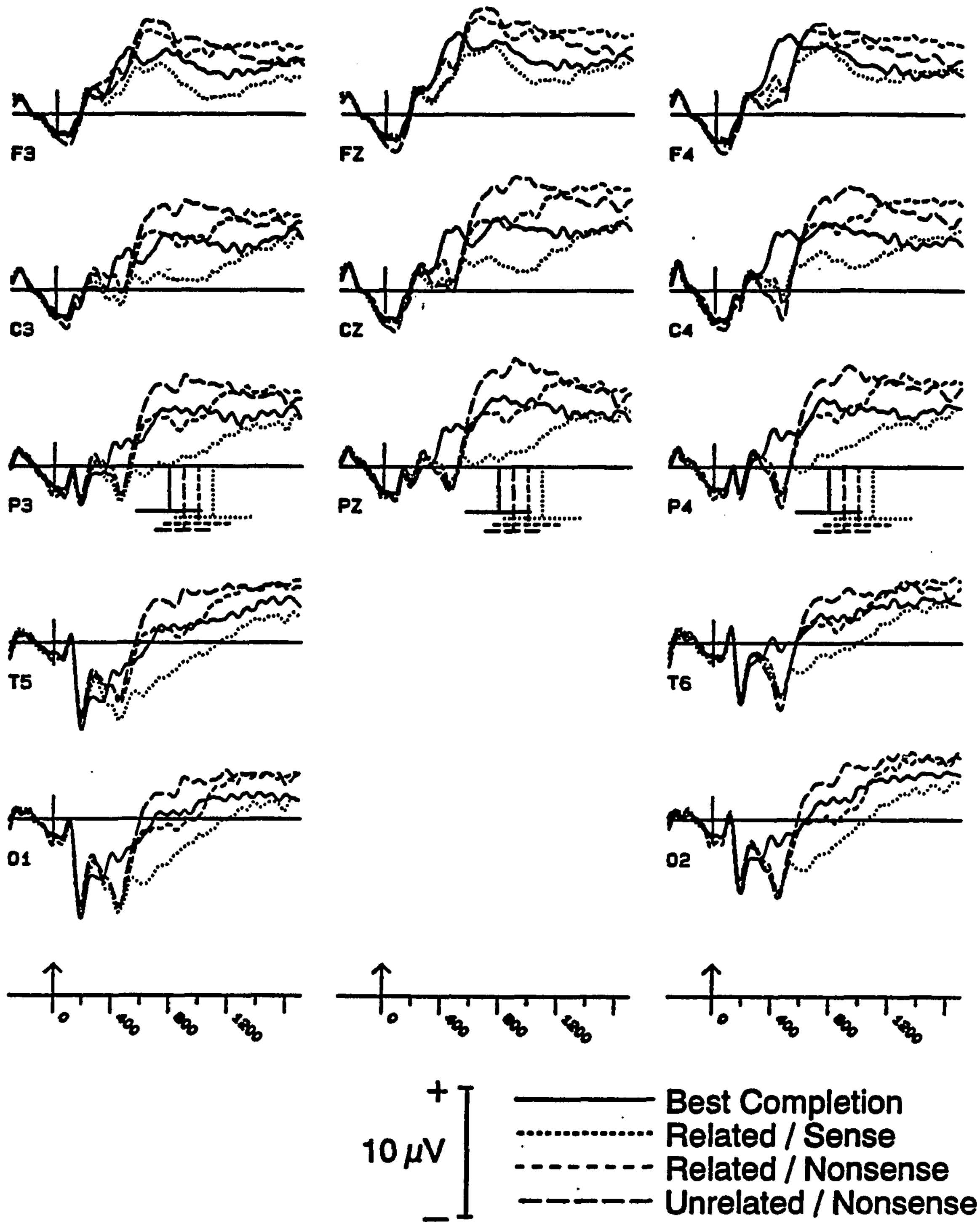



**Figure 14**

Grand mean ERPs by Sentence Type, Normal Elderly group

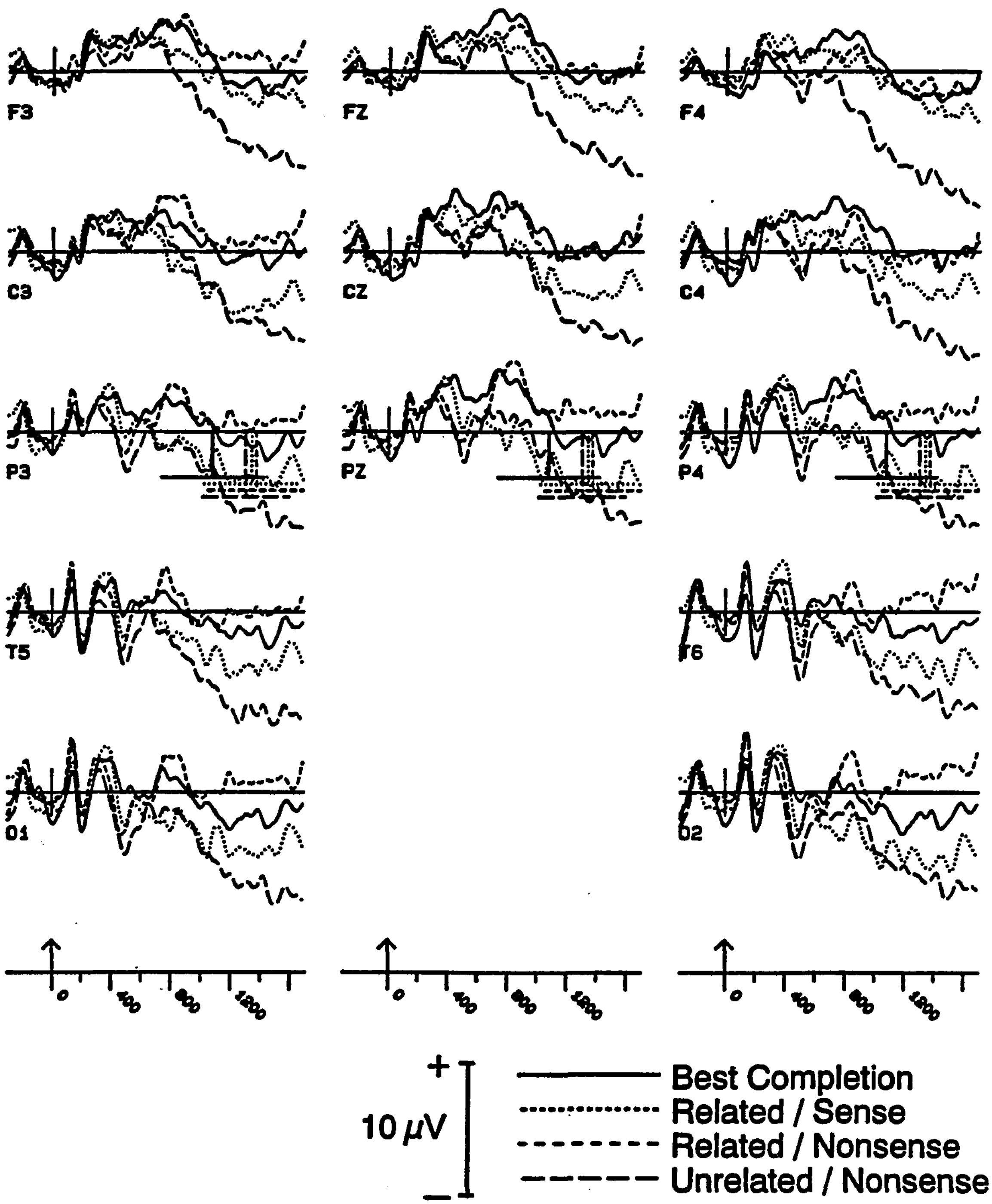



**Figure 15**

Grand mean ERPs by Sentence Type
PAD group, n=6

Table 15
Component latency intervals

| | Young Normal | Norm Elderly | PAD* |
|---|---|---|---|
| N4a | 250-290 | 300-400 | 350-425 |
| N4b | 290-360 | 400-500 | 425-525 |
| N4c | 360-440 | 500-550 | 525-600 |
| N4tot | 250-440 | 300-550 | 350-600 |
| P600 | 450-700 | 550-800 | 600-900 |
| SW1 | 700-1000 | 800-1100 | 900-1200 |
| SW2 | 1000-1300 | 1100-1300 | 1200-1500 |

*Same latency intervals for PAD (N=5)
Component latency intervals (in milliseconds post stimulus) within each group.

B. Behavioral and ERP Analyses

B1. Homogeneity of Variance

Both the behavioral and ERP data were analyzed for effect of Group, the interaction of Group with Sentence Type, and the effect of Sentence Type within each group. Before proceeding with the analyses, these data were examined for homogeneity of variance between groups, and among Sentence Types within each group. These analyses are detailed in Appendix F. Assumptions of homogeneity of variance were met for 1) behavioral and N400 measures among groups, collapsed across Sentence Types; 2) for behavioral measures among Sentence Types within each group; and 3) for N400 among Sentence Types in the NE and PAD groups. However, in the YN group no portion of the N400 showed homogeneity of variance among Sentence Types. Therefore, the YN N400 data were subjected to logarithmic transformation prior to analyses for the effect of Sentence Type. As a result of these transformations, homogeneity was indicated for log(N4b) and log(N4c). Although analyses were performed on the transformed data, raw amplitude measures are presented in the text.

B2. Group Differences

It was expected that overall (i.e., regardless of Sentence Type), PAD subjects would be slower and less accurate relative to both YN and NE controls. Accordingly, the main effect of Group was significant for RT ($F_{(2/13)} = 13.46$, $p<.001$), Within-subject SD ($F_{(2/23)} = 10.07$, $p<.001$), and Percent Correct ($F_{(2/23)} = 16.76$, $p<.0001$). As shown in table 16, the PAD group demonstrated slowest RTs and lowest Percent Correct. Pairwise comparisons confirmed that across Sentence Types, YN and NE were similar to each other, but distinct from the PAD group on measures of RT and Percent Correct ($p<.001$).

Table 16
Behavioral Data collapsed across Sentence Types within each group.

| Group | RT (msec) | % Correct |
|---|---|---|
| YN | 806 (187) | 94 (6.4) |
| NE | 941 (219) | 89 (12.1) |
| PAD (6) | 1271(214) | 73 (16.9) |
| PAD (5) | 1271(226) | 71 (17.0) |

C. Effects of Experimental Variables on N400

The principal hypothesis of the study involved a Group by Sentence Type interaction in which both N400 amplitude and behavioral patterns of YN and NE subjects would be similar to one another, yet distinct from that of the PAD group. It was predicted that in YN and NE subjects, N400 amplitude would show a gradation as a function of Sentence Type, whereas N400 amplitude patterns in PAD subjects would reflect the collapsing of this gradient.

These questions were explored by assessing the effect of Sentence Type, the interaction of Group with Sentence Type and by examining response patterns as a function of Sentence Type within each group for N400 and behavioral variables.

C1. Effect of Sentence Type

C1a. N400 amplitude (All electrodes)

As can be seen in figures 13, 14, and 15, and as assessed by a Group by Sentence Type ANOVA, Sentence Type appeared to affect N400 amplitude in each group, leading to a main effect of Sentence Type (N4a ($F (3/63)=4.1$, $p<.01$), N4b ($F (2/56) = 15.13$, $p<.001$), N4c ($F (2/56)=13.92$, $p<.001$) and N4tot $F (2/59) = 15.32$, $p<.001$).

The main effect of Electrode (N4b: $F (1/13)=11.38$, $p<.01$) and the interaction of Electrode and Sentence Type (N4b: $F (3/30) = 2.89$, $p<.05$) reached significance only in the NE group. Further analysis indicated that the effect of Sentence Type was significant at each scalp electrode ($p <.05$).

In addition, there were significant Sentence Type by Group interactions (N4c: $F (5/56) = 4 .53$, $p<.001$; N4tot: ($F (3/63) = 4.1$, $p<.01$). These interactions were further explored via within group analyses detailed below. As stated earlier, the particular segment of the N400 (i.e., N4b, N4c...) showing the differentiation among Sentence Types was less important that the fact that any portion of the waveform could be identified that would demonstrate this pattern. Each portion that was responsive to Sentence Type within each group (and also demonstrated homogeneity of variance) was further examined via pairwise comparisons.

D. Within Group Analyses

N400 amplitude measures across all scalp electrodes for each group as a function of Sentence Type are shown in table 17.

D1. N400

In order to obtain both homogeneity of variance and sufficient power to demonstrate distinctions among Sentence Type within each group, within-group analyses included amplitude measures from all scalp electrodes. As can be seen in the ERPs (figures 13-15), the effect of Sentence Type appeared widespread across the scalp.

Table 17
N400 amplitude by Sentence Type in each group

| Sentence Type | YN (N4c) | NE (N4b) | PAD (N4b) |
|---|---|---|---|
| Best Completion | 1.9 | 2.2 | 2.7 |
| Related/Sense | 0.5 | -1.4 | -0.02 |
| Related/Nonsense | -0.9 | -0.7 | -2.7 |
| Unrelated/Nonsense | -0.07 | -1.5 | -4.2 |

N400 amplitude (segment with both homogeneity of variance and greatest differentiation among Sentence Types).

D1a. Young Normals

Within-group analysis indicated that the effect of Sentence Type was significant for (log)N4b ($F(2/16) = 5.4$, $p<.02$) and (log)N4c ($F(3/22) = 21.64$, $p<.001$). Paired comparisons indicated that (log)N4c showed the greatest degree of differentiation among Sentence Types. (* indicates $p<.05$, ** $p<.01$)

BC <* RS <** RN, UN

Although (log)N4b demonstrated the same ordering of Sentence Types, BC and RS were not statistically distinct (BC, RS <** RN,UN).

D1b. Normal Elderly

In the NE group, N4b ($F (2/20) = 7.12$, $p<.003$), N4c ($F (3/23) = 3.25$, $p<.05$), and N4tot ($F (2/19) = 4.76$, $p<.02$), showed effects of Sentence Type. Paired comparisons indicated that for N4b and N4c, Best Completion was differentiable from the other Sentence Types (RS, RN, UN; $p< .01$) but that the latter three were not reliably different from one another.

N4b: BC <* RN ,RS, UN

N4c: BC < UN, RN, RS

N4tot: BC < RN ,UN , RS

D1c. PAD

In the PAD Group, within-group analyses indicated that the effect of Sentence Type was significant for N4b ($F (2/8) = 5.84$, $p<.03$). Interestingly, post hoc tests revealed an ordering among Sentence Types identical to that of the YN Group. Amplitude patterns were as follows:

BC *< RS *<* RN, UN

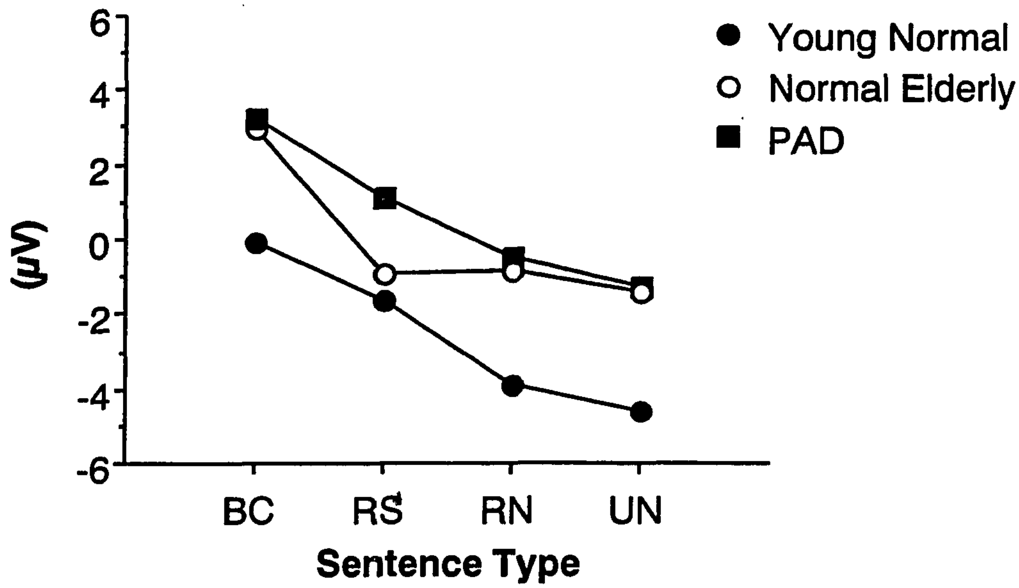



**Figure 16**
N400 amplitude by Sentence Type

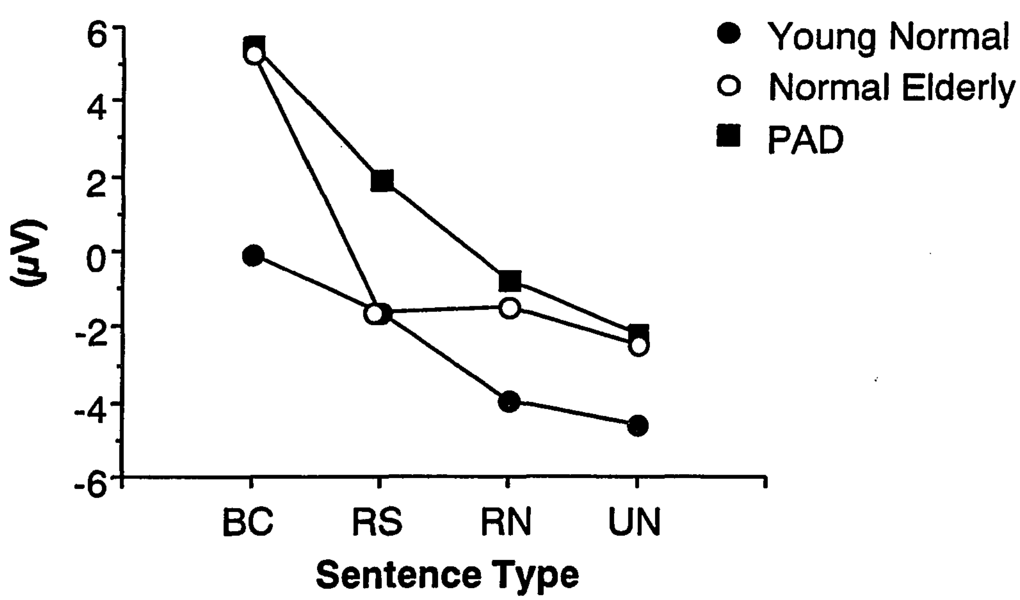



**Figure 17**
Scaled data. Mean N400 amplitude by Sentence Type

Figure 16 depicts for each Group, N400 (N4b) amplitude at Pz (i.e., the midline electrode at which Sentence Types showed the greatest degree of differentiation in N400 amplitude). As can be seen, although there were generalized amplitude differences between groups, both the YN and PAD groups evidenced a similar "shape" , i.e., a progressive increase in negativity from Best Completion to Unrelated/Nonsense. By contrast, the NE Group lacked this progression, instead showing more of a dichotomy between Best Completion and the other Sentence Types. This figure illustrates the patterns revealed in the analyses reported above. However, by viewing the data in this manner, it is also observable that the absolute amplitude measures of the NE and PAD Groups are virtually indistinguishable for Best Completion, Related/Nonsense, and Unrelated/Nonsense (confirmed by Newman-Keuls post-hoc tests, $p>.05$), and are therefore probably different from those of the YN group due to age-related factors. Thus, the greatest discrepancy appears to be that contributing to the amplitude of N400 elicited by Related/Sense in the NE Group and is therefore more likely due to differences in processing as opposed to nonspecific population differences.

The data ordered in this fashion were observed at both midline and lateral (elaborated under Hemisphere analyses) electrode sites. By visual inspection, it appeared that the YN and PAD Groups showed a similar N400 amplitude Sentence Type function , which was discrepant from that of the the NE group. To eliminate differences in magnitude, the N400 of the NE and PAD groups were scaled (at Pz) to that of the YN Group (McCarthy and Wood, 1985). As can be seen (figure 17), the general pattern among Sentence Types was maintained. However, analyses of these data revealed that at Pz, the Group by Sentence Type interaction was not reliable. Thus, the shape of

N400 amplitude as a function of Sentence Type did not differ among groups at Pz.

E. Behavioral Variables

E1. Reaction Time

Mean RTs for each group as a function of Sentence Type are shown in table 18 and figure 18. As assessed by a Group by Sentence Type ANOVA, the main effect of Sentence Type ($F(2/50) = 62.96$, $p<.001$) was modulated by the interaction of Group and Sentence Type ($F(4/51) = 3.48$, $p<.01$).

Within-group analyses indicated that the effect of Sentence Type was significant within each group (YN: $F(3/36) = 3.25$, $p<.03$; NE: $F(3/36) = 4.5$, $p<.008$; PAD: $F(3/20) = 2.94$, $p<.05$).

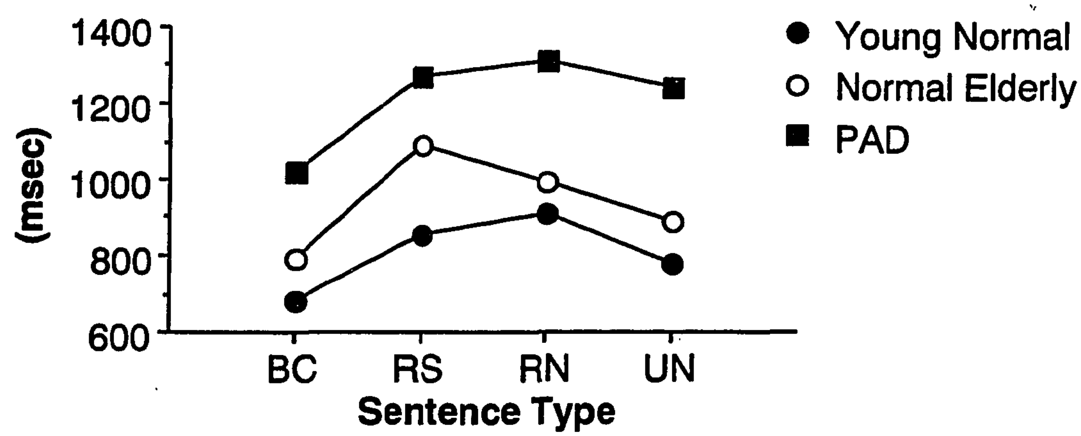



**Figure 18**
Mean RT by Sentence Type

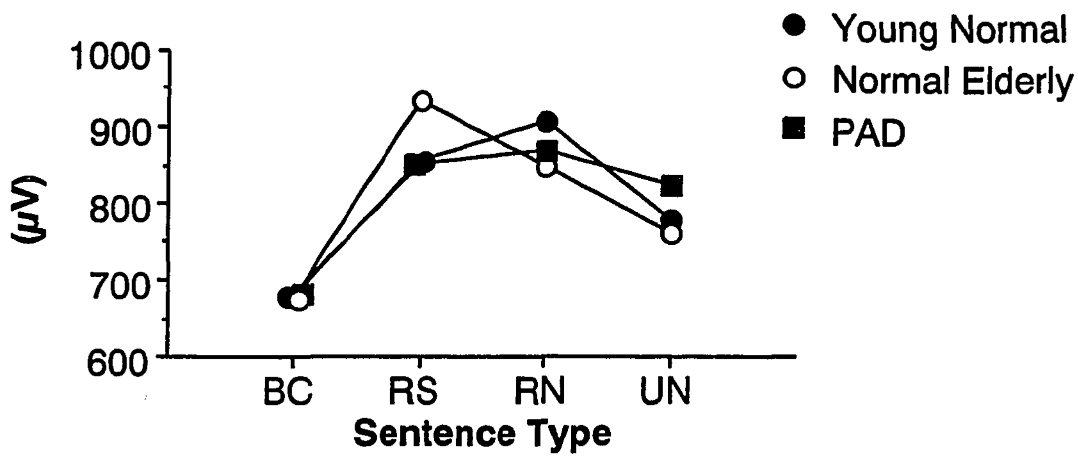



**Figure 19**
Scaled data. Mean RT by Sentence Type

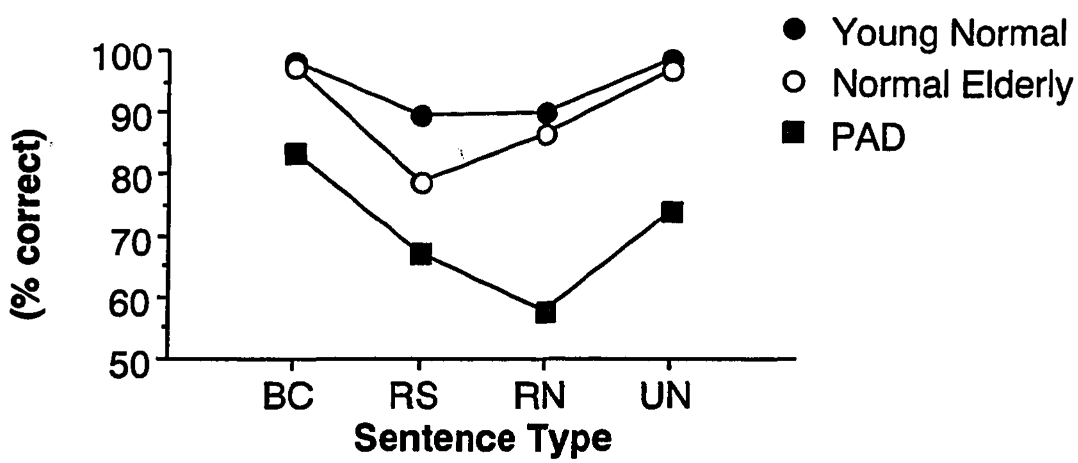



**Figure 20**
Mean Percent Correct by Sentence Type

Table 18
Mean reaction time by Sentence Type

| ST | YN | NE | PAD (N=6) | PAD (N=5) |
|---|---|---|---|---|
| 1 BC | 680 (187) | 789 (185) | 1075 (212) | 1075 (238) |
| 2 RS | 855 (177) | 1092 (195) | 1343 (179) | 1326 (195) |
| 3 RN | 907 (184) | 992 (200) | 1369 (179) | 1365 (200) |
| 4 UN | 780 (134) | 892 (195) | 1296 (193) | 1316 (209) |

Within each group, paired comparisons indicated the following patterns in RT:

YN: BC,UN *< RS, RN ( p<.02)

NE: BC,UN, RN **< RS ( p<.01)

PAD: BC < UN, RS, RN (p<.05)

Inspection of figure 18 suggested that the distribution of RT across Sentence Types in the PAD and YN Groups was relatively similar, but different from that of the NE Group. The RT data were subjected to further analysis in order to determine the source of the Group by Sentence Type interaction. As with the N400 data, NE and PAD groups were scaled to the YN group. As can be seen (figure 19), although group differences in magnitude were reduced, the interaction of Group and Sentence Type was significant ($F(4/51)=3.20$, $p<.02$). Newman-Keuls post hoc comparisons indicated the following patterns among Groups for each Sentence Type ("*"=p<.05):

Best Completion: No differences among Groups
Related/Sense: YN, PAD *< NE
Related/Nonsense: PAD, NE *< YN
Unrelated/Nonsense: NE, YN *< PAD

Although there were similarities between groups within Sentence Types, no two groups shared an identical pattern across all Sentence Types.

E2. Percent Correct

Mean values for Percent Correct as a function of Group and Sentence Type are shown below in table 19 and figure 20. The main effect of Sentence Type ($F (2/37) = 23.35$, $p<.0001$) was significant. Across Groups, responses to Best Completion and Unrelated/Nonsense were more accurate relative to those elicited by Related/Sense and Related/Nonsense ($p<.05$). The interaction of Group and Sentence Type approached but did not reach significance ($p<.06$).

Table 19
Mean Percent Correct by Sentence Type

| Sentence Type | Young Normal | Normal Elderly | PAD (N=6) | PAD (N=5) |
|---|---|---|---|---|
| 1 BC | 98.4 (3) | 97.3 (4) | 83.8 (15) | 84.6 (15) |
| 2 RS | 89.7 (6) | 78.9 (11) | 67.2 (13) | 66.8 (14) |
| 3 RN | 90.3 (7) | 86.5 (15) | 57.7 (14) | 59.8 (13) |
| 4 UN | 98.7 (2) | 96.8 (5) | 74.3 (18) | 74.2 (19) |

Within-group analyses confirmed that the effect of Sentence Type on Percent Correct was significant in YN ($F(2/14) = 9.45$, $p<.03$), NE ($F(1/13) = 10.07$, $p<.001$), and PAD ($F(2/9) = 10.16$, $p<.006$) groups. Paired comparisons indicated the following patterns (*= $p<.05$):

YN: UN,BC >* RN,RS
NE: BC,UN >* RN,RS
PAD: BC >* UN,RS >* RN

In summary, the YN group showed the expected N400 amplitude gradient as a function of semantic relatedness to a primed word. However, the N400 patterns generated by both the NE and PAD groups were unanticipated. In the NE group, N400 amplitudes to all unexpected final words were similar, regardless of their semantic relationship to the primed word (i.e., the best completion). By contrast, the PAD group, in spite of their relatively impaired behavioral performance, demonstrated an orderly N400 gradient, similar to that exhibited by the YN group. Across groups, Best Completion and Unrelated/Nonsense Sentence Types elicited relatively quick and accurate responses, whereas the ambiguity imposed by Related/Sense and Related/Nonsense Sentence Types generated slower, more variable (see Appendix E) and fewer accurate responses.

As expected, the YN group generated the most rapid and most accurate behavioral responses in contrast to the PAD group, whose responses were slowest and least accurate. Also as expected, PAD subjects committed more errors to Related/Nonsense) than any other Sentence Type, demonstrating difficulty discriminating among semantically related items. An unexpected finding was the fact that the NE group evidenced the most difficulty with Related/Sense relative to the other Sentence Types.

F. Additional Analyses

F1. N400 at T5/T6

Although not central to the present thesis, the Sentence Type effect was examined at T5 and T6 electrode sites, because some investigators have found similar N400 effects to be prominent at these locations relative to other lateral sites (e.g., Kutas et al., 1988). Results from analyses inclusive of all lateral electrodes are presented in Appendix C. The data presented here will include analyses of N4b (middle portion of overall N400) at T5 and T6

N4b amplitude measures, as a function of Hemisphere (T5/T6) and Sentence Type, separately for each Group are shown in figure 21. There were no main effects of Group or Hemisphere on N4b. However, the main effect of Sentence Type ($F(2/51) = 10.56$, $p<.001$) was significant.

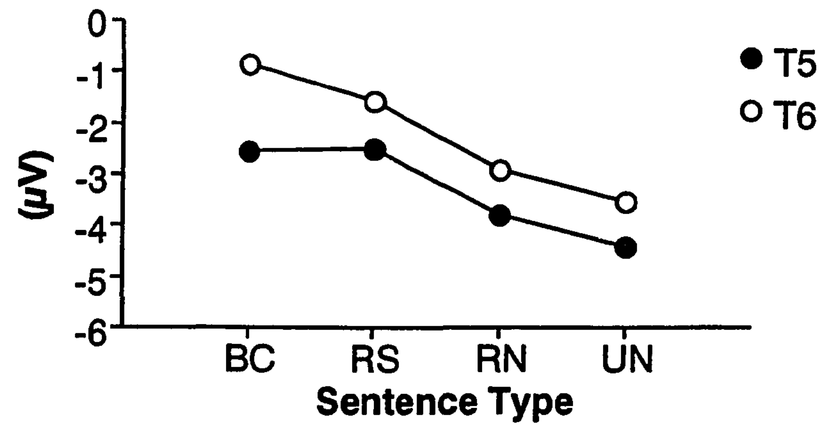



**Figure 21a**

Young Normal. Mean amplitude at T5 and T6.

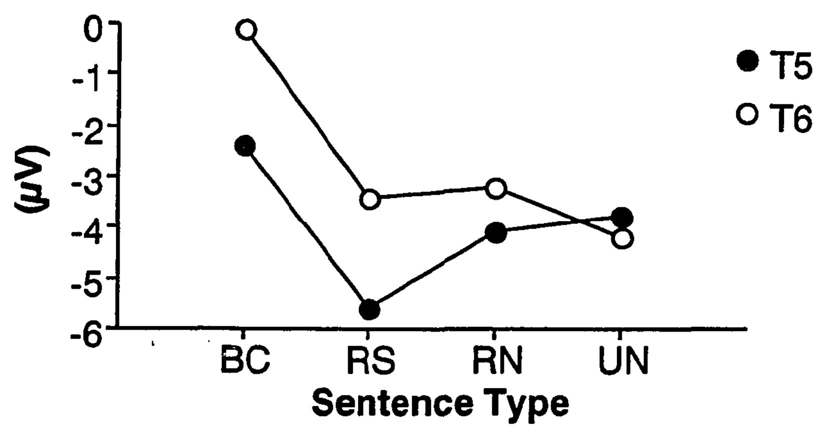



**Figure 21b**

Normal Elderly. Mean amplitude at T5 and T6.

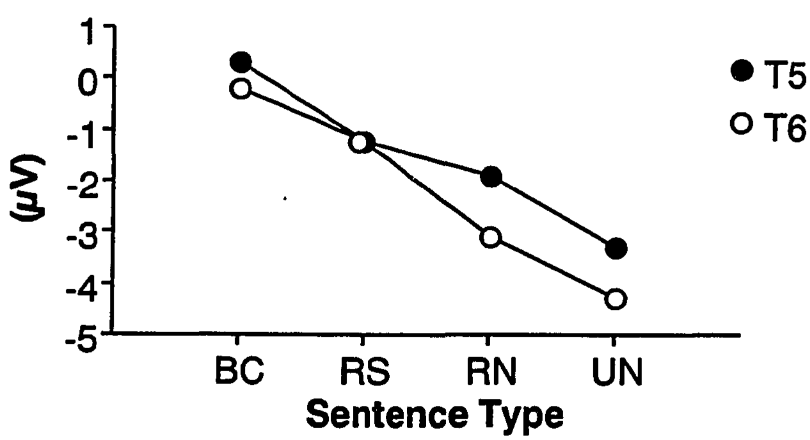



**Figure 21c**

PAD. Mean amplitude at T5 and T6.

The interaction of Hemisphere and Sentence Type (F (2/51) = 7.05, p<.001) was modified by the 3-way interaction of Group, Hemisphere and Sentence Type (F (4/47) = 2.79, p<.04). Further analyses indicated that although Sentence Type was significant in each Group (YN: F (2/65) = 13.06, p<.001; NE: F (2/51) = 5.55, p<.001); PAD: F (2/51) = 6.00, p<.01) the interaction of Hemisphere and Sentence Type (F (2/51) = 10.84, p<.001) was significant only in the NE Group. From the patterns observed in figure 21, this result was most likely due to the markedly enhanced negativity elicited by Related/Sense at T5 (similar to, yet larger than that observed at Pz), which was markedly distinct from the patterns shown for the other two groups.

To further examine this effect, N4b amplitude measures at T5 in the NE and PAD groups were scaled to that of the YN group. These scaled data are shown in figure 23. Unlike the analysis of the scaled data at Pz, the interaction of Group and Sentence Type on the scaled data at T5 was significant (F (5/56)=3.52, p<.01). Post Hoc tests for the effect of Group at Related/Sense confirmed that the N400 of the NE Group was larger (p<.01) compared to YN and PAD groups, which did not differ. Results of post hoc tests for the remaining Sentence Types were as follows:

Best Completion: YN , NE >* PAD
Related/Nonsense: no differences among groups
Unrelated/Nonsense: PAD, YN, NE ( PAD >* NE)

Thus, although it seemed that the different pattern of Sentence Type across groups (e.g., the Group by Sentence Type interaction) could be accounted for by the deviation of the NE group for Related/Sense, post hoc testing revealed differences at other Sentence Types as well.

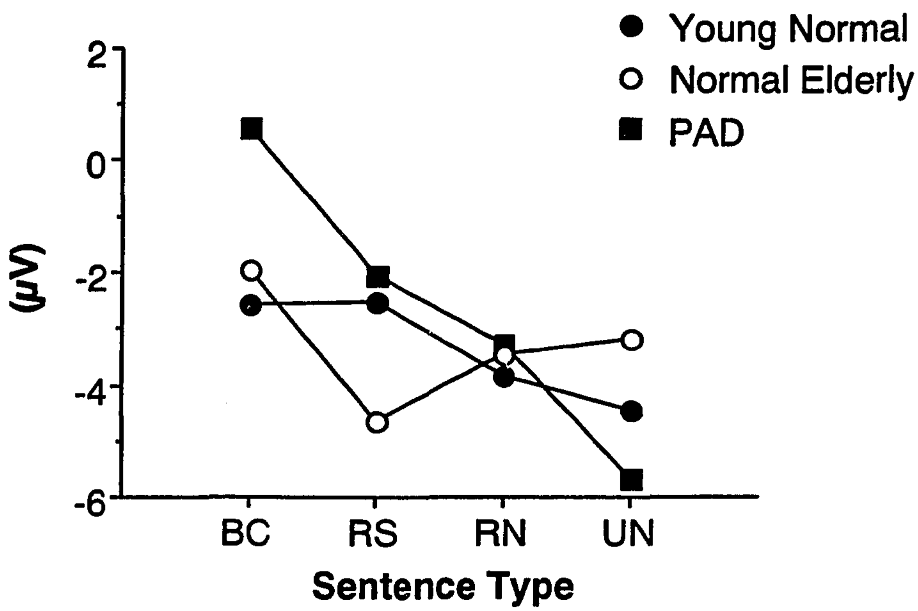



**Figure 22**
Scaled data. N400 amplitude at T5

F2. Late Positive Components

There were no main effects of Group, Sentence Type or Electrode on P600, SW1 or SW2. Interactions of these variables however, are detailed in Appendix C.

F3. Sweep Count

The data presented above included trials to which responses were correct, and were either artifact-free or corrected for eye artifact. The mean number of trials contributing to each Sentence Type average within each Group are shown in table 20.

Table 20
Mean number of trials per Sentence Type

| ST | YN | NE | PAD (N=6) |
|---|---|---|---|
| 1 BC | 39.1 (1.1) | 38.8 (0.4) | 33.0 (6.1) |
| 2 RS | 35.9 (2.5) | 31.5 (4.4) | 25.0 (4.2) |
| 3 RN | 36.1 (2.6) | 34.6 (5.8) | 23.8 (5.9) |
| 4 UN | 39.5 (0.8) | 38.4 (2.0) | 28.2 (7.8) |

Clearly, the ERPs of the PAD group were based on fewer trials compared to that of the other two groups. Accordingly, the main effect of Group was significant F (2/23 = 23.02, p<.001) as was the effect of Sentence Type (F (3/40) = 24.24 , p<.001). The interaction of these variables approached but did not reach significance (p<.07). Despite the reduced number of sweeps contributing to the average conditions in the PAD group, it apparently was sufficient in that differences among Sentence Types were demonstrated. However, it is unknown what other differences might have emerged had the number of sweeps been comparable to that of the other groups.

G. Informal Observations

During the course of data collection, a number of informal observations were noted. Since the behavior and comments by subjects within a group were similar, these observations are reported as a function of Group.

The data from the 3 practice blocks were not analyzed since these stimuli were not as carefully screened as the experimental stimuli via norming procedures. However, during these blocks, many subjects asked questions and/or reported that they altered their strategies. During practice, most YN subjects committed a disproportionate number of errors on Related/Sense stimuli. Some reported that if the final word was not the word they expected, they classified the sentence as "nonsense." Some subjects reported that they felt confused, since "although the word was not expected, it could possibly make sense," and asked how these stimuli should be categorized. After hearing the instructions a second time, several subjects commented that they would "broaden the category" for "sense." Following this clarification, RT and errors to Related/Sense stimuli were reduced. At the end of the experiment, most YN subjects reported that they shifted their strategy once the instructions were made clear.

Similar to the YN subjects, subjects in the NE group also evidenced the most difficulty with Related/Sense sentences during the practice blocks. Approximately half of these subjects reported that the classification of these stimuli was confusing. Those who did not ask had assumed that sentences which were completed with unexpected words did not make sense. Although instructions were repeated, and subjects were able to paraphrase the instructions, demonstrating that they understood, subjects reported having difficulty with these sentences. Consequently, (as detailed above) RT remained relatively prolonged.

Unlike subjects in the YN and NE groups, subjects in the PAD group rarely initiated questions about the task. Only one subject in the PAD group inquired about Related/Sense sentences. Also, PAD subjects experienced more than 3 practice blocks, in which most effort was directed at reducing artifactual movements. Most PAD patients required at least four practice blocks before they were ready to begin the experiment. As previously stated, although additional practice might have benefitted performance, the cost would have been increased frustration and fatigue.

# VI. DISCUSSION

## A. Major Findings

Behavioral data from various research domains have provided evidence that semantic memory is functionally organized along gradients of semantic relatedness. There is also evidence that Alzheimer's patients show deficits when task performance is contingent upon the integrity of this organization. This investigation attempted to demonstrate the semantic gradient in both the behavior and ERP of normal subjects, and to examine the correspondence between behavioral and electrophysiological response patterns in mild Alzheimer's patients.

The ordering of N400 amplitudes in young normal subjects clearly demonstrated the ERP gradient. As predicted, N400 amplitude progressively increased as target stimuli were decreasingly primed by the preceding semantic context. In addition to the predicted ordering of N400, mean amplitude measures for the four sentence types were statistically distinct at three of these data points (i.e., (1) Best Completion < (2) Related/Sense < (3) Related/Nonsense, (4) Unrelated/Nonsense). The behavioral response pattern, although not as systematic, also reflected the semantic gradient. Response times were most rapid to Best Completion followed by Unrelated/Nonsense, and were statistically distinct from those of Related/Sense and Related/Nonsense, which were relatively delayed. Since the sense/nonsense decision task was not a task that is typically used to assess semantic priming (as is lexical decision), it was not expected that the presentation of a semantically primed word would necessarily promote response facilitation. To the contrary, to the extent that Related/Sense was analogous to "atypical true" sentences and Related/Nonsense was analogous to "related false" sentences of sentence verification tasks, it was expected that final words of these sentences would impede rather than facilitate behavioral responding. This prediction is based on a model in which activation spreads outward from a word to its semantic associates. For instance, the model predicts that a final word of Related/Nonsense would be activated due to its semantic relationship to the best completion. As a consequence of its activated state, it was relatively difficult to distinguish the Related/Nonsense stimulus from the expected word (which would require a "sense" response), and therefore RT was delayed, i.e., these words generated a conflict in response choice due to their elevated state of activation. Otherwise, these RTs would have been rapid, i.e., similar to those of Unrelated/Nonsense which were not activated by the sentence context. Similarly, final words of Related/Sense were also activated. Although these stimuli required the same response as the best completion (i.e., sense), confusion was created by the strong expectation for one particular word together with the presentation of a related (and therefore activated) word. Unlike Related/Nonsense, these stimuli required the same response as the best completion. Therefore, although RTs were delayed, they were more rapid (although not statistically) than RTs to Related/Nonsense. Thus, the gradient was demonstrated in both the ERP and behavioral response patterns of young normal subjects.

Another major finding was that the error pattern in the PAD group reflected the systematic decomposition of the semantic gradient. The basis of the gradient is the information or processes that distinguish semantically related items from one another. As these distinctions become less clear, related items would seem equivalent. As described, this could account for an Alzheimer's patient naming an "apple", a "pear" (i.e., semantic paraphasia). Furthermore, if the gradient had collapsed, there would no longer be any means of differentially priming related words, i.e., all related items would be equally activated. In the present task, based on behavioral indices (Percent Correct), Related/Nonsense played a critical role in demonstrating the collapse of the semantic gradient. Despite the fact that a final word of Related/Nonsense caused the sentence to be nonsensical, its level of activation may have been indistinguishable from that of the best completion. Since the best completion required a "sense" response, many of the Related/Nonsense sentences also received a "sense" response. Within the PAD group, a disproportionate number of errors were committed to Related/Nonsense.

Perhaps the most interesting result of this investigation was that similar to young normals, N400 amplitude in the PAD group varied with the extent to which a word was was primed by its preceding semantic context. In addition to the similar ordering of N400 amplitudes by sentence type, the statistical pattern among sentence types in the PAD group was identical to that demonstrated in the young normal group (i.e., Best Completion < Related/Sense < Related/Nonsense, Unrelated/Nonsense). This ERP pattern in the PAD group was discrepant with their accuracy performance. As stated above, the behavioral responses (i.e., Percent Correct) of the PAD group was indicative of a collapsed semantic gradient. Although performance was impaired, the integrity of the ERP gradient provides evidence that the information distinguishing semantically related items may be retained at some (more fundamental) level.

As described earlier, the deficits in performance on semantic tasks might reflect impairments in the ability to access and use semantic information, rather than an actual loss or disorganization of the semantic information (see Introduction). Perhaps, the intact N400 gradient in PAD patients reflects the integrity of the semantic network, whereas their poor behavioral performance reflects the difficulties with its utilization. For instance, one possibility is that the critical information is inaccessible to consciousness, and therefore cannot be applied appropriately. It is also possible that this information does influence responding, although PAD patients lack conscious awareness of this information. Perhaps the uncoupling of conscious and unconscious processing resulted in poor performance associated with difficult items (i.e., Related/Nonsense). Another possibility is that only partial information reaches consciousness, and that it is this information that is available to consciousness that affects behavior. For example, on some trials, partial information may have caused confusion and delayed RT (but responses were accurate), whereas on other trials, insufficient information led to an incorrect response. Alternatively, it is possible that the necessary information does in fact reach consciousness, but that Alzheimer's patients lack the attentional resources to appropriately use the information.

The present task was highly demanding. As stated earlier, in addition to semantic processing, attentional and memory systems were involved, and possibly the time pressure to respond quickly induced some level of anxiety in some subjects. Thus, if performance is actually task dependent (as several investigators presently theorize, described in the Introduction) poor performance would be expected on this task. It would be interesting to observe ERP patterns in PAD patients on other tasks in which their behavioral performance is impaired (e.g., modified verbal fluency tasks, object naming tasks, etc.).

B. PAD Response Patterns and Models of Semantic Memory

Earlier, semantic memory models were used to predict the nature of the degeneration of the semantic system. The combination of behavioral and ERP patterns observed in the PAD group can be described in terms of the structure and processing assumptions incorporated in theories of semantic organization. It was hypothesized that the structural properties of the semantic network which encode relatively subtle differences between related items might begin to disintegrate. Again, this would explain the types of semantic errors observed in PAD patients. However, if the intact N400 gradient in PAD patients is taken as evidence for the integrity of this organization, this hypothesis was violated. In retrospect, using Grossberg and Stone's (1986) theory[4], this outcome (i.e., intact semantic network but impaired task performance) is viable. According to the model, development of semantic organization begins early, and follows a gradual sequence of events. Once encoded in LTM, information is biased toward stability; changes in LTM require repeated exposure of conflicting information (see Introduction). However, during the course of a lifetime, the information encoded within a semantic/language system is largely reinforced by experience within a given cultural community. Thus, meaning/language is overlearned, and is highly resistant to disruption. However, the preservation of this organization in LTM does not ensure its appropriate utilization. Although Adaptive Resonance Theory is best applied as a model of word recognition, and thus, the present task is beyond its scope, its application to the present data can be discussed up to the point of word recognition (i.e., without the subsequent judgement task). The presentation of a highly

---

[4] Although the other theories could also accommodate the results, Grossberg and Stone's (1986) model is used here because it provides a more detailed description of structure and processing assumptions.

constrained sentence context to an intact semantic system would elicit top/down activation of the primed (best completion) word. Related words from level A5 (the semantic network) would also be primed, but to a lesser extent. An intact system would detect this difference in activation, and thereby distinguish the best-primed word from the others. In PAD, it is possible that although the LTM (semantic) organization is intact, the processes by which this information reaches consciousness and/or affects behavioral responding are impaired. Using Grossberg and Stone's (1986) theory, this deficit could occur at several different points. In a healthy system, the distinguishing feature of a best completion word might be the degree of resonance between the best-completion input stimulus and the highly activated word from LTM. In PAD, it is possible that there is an inability to discriminate different levels of activation among related items because either a) semantically related words elicit equivalent levels of resonance, or b) although resonance activity will differentiate related items, PAD patients are unable to detect these differences. Another possibility could be that these processes function normally as long as they are not directed by conscious, controlled attention. Again, if the N400 gradient reflects the integrity of semantic organization, then it can be assumed that the processes from incoming input to activation of the semantic network function normally, and that the deficit exists at some point between this activation and the emission of a response.

Similar to reports of the performance of PAD patients on semantic tasks, several investigators report that, compared to both normal and Broca's aphasia control subjects, Wernicke's and Anomic aphasics show impaired ability to use semantic relations (e.g., Zurif, Caramazza, Myerson and Galvin, 1974; Goodglass and Baker, 1976; Whitehouse, Caramazza and Zurif, 1978;

cited in Milberg and Blumstein, 1981). However, Milberg and Blumstein (1981) reported that the same subjects who performed poorly on language tasks that involved controlled cognitive processing demonstrated intact semantic priming effects. Similar to some of the interpretations regarding the pattern of performance of PAD patients on semantic tasks, these findings were interpreted as evidence for the retention of semantic information which is inaccessible to intentional semantic decisions.

Neuroanatomically, lesions in temporo-parietal or temporo-occipital regions give rise to disturbances in word-meaning and word-finding (Luria, 1975). Although the neuropathology most commonly associated with early Alzhermer's Disease involves the medial temporal region, there is evidence of abnormalities in these areas as well. Especially in patients whose initial presentation involves a severe but gradual language deficit, PET studies have indicated reduced glucose utilization, and autopsies have revealed a substantial density of placques and tangles in these areas (Albert, Duffy, and McAnulty, 1990; Cappa, Covalloti, and Vignoto, 1981 cited in Martin, 1987; Martin, 1987). It would be interesting to determine empirically whether all patients who show this particular combination of intact semantic structure with impaired utilization also show a high concentration of pathology in the same brain region(s).

C. Response Patterns in the Normal Elderly

The expected N400 gradient was not demonstrated in normal elderly subjects. In the NE group, Related/Sense elicited the largest N400, although statistically, N400 amplitudes elicited by Related/Sense, Related/Nonsense and Unrelated/Nonsense were not differentiable.

Several lines of evidence suggest that the N400 pattern in the NE group reflects strategy related factors, as opposed to a disruption in semantic organization. First, task performance was adequate. Second, this pattern was markedly similar to that demonstrated by younger subjects during the initial pilot experiment. Although three practice blocks were sufficient for the induction of a strategy change in young adults, it might have been insufficient for the elderly adults. In retrospect, it was an unwarranted assumption that the experimental parameters suited for the young adults would be equally appropriate for the elderly subjects. However, the decision to norm task parameters (i.e., number of practice blocks) on young adults was based on the limited availability of normal elderly subjects, who were needed for the main experiment. Also consistent with this line of reasoning were several indications that the NE group maintained a strategy in which any unexpected sentence completions were apt to be classified as nonsensical. The NE were the only group to show the most difficulty with Related/Sense. Unlike the YN or PAD groups, reaction times were slowest and most variable (see Appendix F) and more errors were committed to these stimuli compared to other stimulus types. Furthermore, many of the NE subjects commented that their first impulse was to classify Related/Sense sentences as "nonsense" because these words were "not the words that should" complete the sentence.

It is not difficult to conceptualize how the strong expectation created by the sentence context for a particular word made it difficult to accept an alternative word as sensible. This explanation addresses the behavioral data. However, it is more difficult to speculate which cognitive processes were responsible for the enhanced negativity in the ERP to Related/Sense. It has already been discussed that this negativity was dissociated from RT (i.e., as evident from quartile analyses), and therefore was not directly related to level of difficulty or effort. Although only speculative at this point, there is some evidence from the behavioral literature that suggests that this negativity might reflect inhibitory processes associated with semantically related words within the semantic network. Brown (1979) unexpectedly found delayed response times and more errors associated with semantically related primes compared to unrelated primes. In this study, subjects (young normal adults) were shown a word prime which they read aloud, followed by a definition. On a given trial, the word prime could be related to the definition in one of the following ways: 1: (C) the correct word, 2: (S) a word semantically related to the correct word, 3: (O) a word orthographically related to the correct word, or 4: (U) a word unrelated to the correct word. For example:

Definition: a mythical animal with a horn at the center of its head.
C: unicorn
S: dragon
O: uniform
U: chair

The subjects' task was to name the word appropriate to the definition (after hearing the "prime" followed by the definition). Subjects showed the most difficulty when primes were semantically related to the correct word. Brown interpreted this result as a spread of inhibition affecting related items within a category and compared this finding to the tip-of-the-tongue (TOT) phenomenon. Rather than aid in the recall of a target word, the related word in fact inhibits recall due to output interference (Brown and McNeill, 1966; Brown 1979). However, using a similar procedure, Roediger, Neely and Blaxton (1983) found that by removing the the "Correct" primes from the stimulus list, the "Semanticlally Related" primes no longer elicited prolonged

RTs or greater error rates. They interpreted their findings as evidence against Brown's theory of "spreading inhibition" and asserted that the inhibition to related primes was due to strategic factors. When correct primes were present, subjects would evaluate the prime to determine if it was the correct word. If correct primes were absent, this strategy was no longer necessary, and consequently, the inhibition to related primes disappeared.

When these procedures were conducted with elderly subjects, an interesting set of results emerged which are relevant to the present findings. Bowles and Poon (1985) tested normal young and elderly subjects on vocabulary, a standard lexical decision task, and in a procedure similar to that described by Brown (1979). Although the elderly generated higher vocabulary scores and performed equivalently on the lexical decision task, their word retrieval performance to semantically related primes was impaired relative to that of younger adults. In a subsequent investigation (Bowles, 1989), when "Correct" primes were not included in the stimulus set, young and elderly subjects performed comparably. Perhaps Related/Sense stimuli (which were closely related to the primed or "correct"word), elicited a similar sequence of cognitive events as did the related primes in the word retrieval task. Although the tasks were very different, they both involve the priming of one particular word (unlike typical priming tasks, e.g., lexical decision, in which general categories are primed), and consequently, the presentation of a semantically related word impaired performance, most markedly in older subjects.

The speculation offered here is a combination of Brown's (1979) spreading inhibition interpretation and Roediger et al.s' (1983) strategy-related explanation. Perhaps the enhanced negativity elicited initially in the YN group during the pilot study, and persistently in the NE group reflects a strategy-related inhibitory process. It is also assumed here that this negativity, and the inhibitory response pattern were superimposed upon a functionally intact semantic system in the NE group.

It is reasonable to question why this inhibition was not apparent in the behavioral performance or ERPs of the PAD patients, who were of comparable age to the normal elderly. One possibility is that PAD subjects lacked the attentional ability to implement a strategy. In the absence of a cognitive strategy, the default mode of operating might be 'semantic relatedness'. Also, despite their intact N400 gradient, it is possible that at the level of consciousness, the boundaries between related items are blurred. If this indeed is their experience, then a strategy (such as that adopted by the NE subjects, i.e., all unexpected words were likely to be considered "nonsensical") would not seem necessary. This might explain why unlike the YN or NE subjects, the PAD subjects seemed unperturbed by Related/Sense endings.

D. N400 in Normal Young Adults during the Pilot Study

Since the N400 gradient had already been reported in normal young adults (Kutas and Hillyard, 1984), it was anticipated that the results of the pilot study would replicate this ERP pattern. However, instead of the expected inverse relationship between N400 amplitude and semantic relatedness (i.e., N400 smallest to semantically related words, largest to unrelated words), N400 was larger to stimuli which were semantically related to the primed words (Related/Sense), compared to those which were unrelated (Unrelated/Nonsense). This ERP outcome was paralleled by the behavioral data in which RT was longest and error rates were highest in response to these Related/Sense stimuli. Although a reasonable hypothesis was that N400 amplitude might be directly related to RT, the results of several analyses demonstrated that N400 amplitude and RT were clearly dissociated (see Pilot Study: Quartile analyses). Thus, level of difficulty or effort as reflected by RT could not account for the enhanced negativity elicited by Related/Sense. Further examination of the data revealed that when the first and last two blocks of trials were averaged separately, the negativity elicited by Related/Sense was attenuated.[5] Visual inspection of the ERPs (across all 4 experimental blocks) of individual subjects indicated that one of the five pilot subjects did actually show the expected N400 gradient. Thus it appeared that while one practice block was sufficient for one individual, most subjects required up to 3 blocks of experience (i.e., one practice block plus two experimental blocks) with the task in order for the superimposed negativity to attenuate (see Pilot Study). It was therefore speculated that this negativity was strategy-related. To test this hypothesis, the original 5 subjects were retested, and an additional 6 naive subjects were tested using a total of 3 practice blocks (40 stimuli per block) prior to the four experimental blocks. As reported, the expected N400 gradient was obtained in both of these groups.

The results of the initial pilot study were at odds with most reports of N400 amplitude during semantic processing tasks. To date, there have been no reports of greater N400 amplitude elicited by words which were related to their preceding semantic context in comparison to N400 elicited by unrelated words. However, there are some important differences between the present task and the more typical semantic priming tasks which could account for this discrepancy. For instance, most reports of N400 have involved single-word

---

[5] Although N400 to RN also showed some change in the same direction, N400 to RS showed the greatest change with respect to its position among the other sentence types. With 10 subjects (main experiment), N400 to RS was statistically more positive relative to that of RN and UN, whereas N400 to RN remained statistically undifferentiated from that of UN. For this reason, the discussion of the shift in N400 amplitude focuses on RS.

presentation lexical decision (Bentin et al., 1985, Rugg, 1990) and category decision (e.g., Polich, 1985; Hamberger and Friedman, 1990) tasks. The cognitive operations required for these tasks (i.e. word recognition in the Lexical Decision task, and category membership in the category decision task) are much less demanding than the sensibility judgements of relatively complex statements in the present task. In the latter, the decision component of the task was superimposed upon word recognition, possibly occuring in parallel and/or overlapping in time. The current study also differed from N400 studies which have employed sentence contexts in several ways. For instance, although Kutas and Hillyard (1980, 1984) presented contexts similar to those used here, they did not employ a sense/nonsense judgement. Stimulus presentation also differed, in that they presented each word separately (stimulus duration approximately 130 msec, ISI = 700 msec). In the current experiment, the sentence context was presented in full, and the final word appeared 500 milliseconds following sentence context offset. This created a situation in which subjects were probably more prone to generate predictions for the final word. Fischler and Bloom's (1985) true/false decision task was also unlike the present task in that the sentence structure of each trial was identical (e.g., An "x" is a "y"). Although stimuli were presented as sentences, the task could be performed at the level of a category decision task. Thus, the combination of stimuli, presentation method and task demands in the current study created a unique set experimental conditions.

Another possibility is that the enlarged N400 to Related/Sense could be attributed to the composition of the stimulus list (i.e., "context effects"). The four types of stimuli may have interacted in ways not previously encountered in N400 experiments. Strong context effects have been demonstrated for RT during sentence verification tasks (Chang, 1986). For example, the inclusion of related-false statements (e.g., All bats are birds) in a list of stimuli differentially delays RT to atypical-true statements (e.g., All chickens are birds). Earlier, it was speculated that Related/Sense stimuli were similar to atypical-true statements, and Related/Nonsense stimuli were similar to related-false statements. (Response times to both of these stimuli were prolonged relative to RTs to Best Completion and Unrelated/Nonsense.) Thus, if the context of a stimulus list can influence RT, it is plausible that context effects can also affect electrophysiological responses.

There is yet another way in which to consider the unexpected large negativity elicited by Related/Sense. It has been argued that N400 is a general response to deviations from expectancy, as opposed to a more specific response to semantic relatedness (Bentin, 1985). According to this position, semantic relatedness is merely one type of expectation that can be created by an experimental situation. It can be argued that at the beginning of the experiment, the wording of the instructions together with the high level of constraint imposed by the sentence contexts created a situation in which Related/Sense was the least expected sentence type, and therefore elicited the greatest negativity. First, for all sentence contexts, the greatest expectation was created for the best completion. Thus, Best Completion words were highly expected. Given the instructions (i.e., to classify sentences as either "sense" or "nonsense"), a strong expectation was also created for words that would be highly discrepant with the sentence contexts, i.e., Unrelated/Nonsense. Since Related/Sense and Related/Nonsense words were related to the best completion, these words were less discrepant with the preceding sentence contexts. However, although empirical evidence is presently unavailable, it is likely that final words of Related/Nonsense were less related (or more unrelated) to best completions relative to final words of

Related/Sense. This would make Related/Nonsense words slightly more discrepant with the preceding context relative to Related/Sense words. The argument here is that Related/Sense was slightly less expected than Related/Nonsense , which could account for the greatest negativity elicited by Related/Sense at the beginning of the experiment. However, after practice with the task and greater familiarity with the stimuli, a different set of expectations developed. Perhaps, when subjects created an expectation for Related/Sense, N400 to these stimuli became less negative, and distinguishable from that of Related/Nonsense and Unrelated/Nonsense (N400 to Related/Sense was statistically distinct only with 10 subjects included in the analysis).

To date, there is only one known study documenting changes in N400 amplitude during the course of an experiment (Fischler, Jin, Boaz, Perry and Childers, 1987). Subjects were assigned an "Assumed Name". When presented with the statement, "My name is X", in which "X" was either the Assumed Name, the subjects "Real Name" or a "False Name", they were instructed to respond "yes" to the Assumed Name, and "no" to the Real Name and the False Name. Fischler, et al. (1987) found that the N400 to the Real Name was larger than that of the Assumed Name, indicating that strategic factors influenced N400 amplitude. Additionally, although the size of the N400 elicited by the Assumed Name and the False Name were unchanged from the first to the second block of trials, N400 amplitude to the Real Name was significantly enhanced by the second half of the experiment. Thus, increased task experience caused subjects to develop a strategy, possibly producing a change in N400 amplitude. These findings are consistent with those of the pilot study, demonstrating that N400 amplitude can vary over a relatively short period of time. The speculation that strategic factors (e.g., a conscious effort to alter original expectations based on task experience) underlie the change in N400 in the present study is consistent with comments made by subjects indicating that they actually did alter their strategy. Again, it is likely that this shift in strategy coincided with the reduction in N400 amplitude elicited by Related/Sense. Moreover, when the initial pilot subjects were retested with additional practice blocks, the absence of an enhanced N400 during this session was accompanied by fewer errors to Related/Sense (across subjects --Session 1: 54 errors vs. Session 2: 18 errors).

E. Methodological Considerations

1. Response requirement

The response requirement was implemented in spite of the potential complications associated with the P300 (i.e, P600) overlapping the N400 measurement window. When P3 and N400 overlap, it can be difficult to disentangle the two components and determine which carries the experimental effect. Other investigators have circumvented this problem by selecting tasks which do not require a decision to stimuli of interest; e.g., Kutas and Hillyard (1980) instructed subjects to read sentences in preparation for a subsequent questionnaire. However, given the predisposition to fatigue and distractibility of the PAD patients (Joynt and Shoulson, 1985), it was important to obtain observable evidence that subjects had engaged in the type of processing that the task was designed to induce.

2. Differentiation among Sentence Types and the N400 Gradient

The N400 pattern in the PAD group suggests that the information required to discriminate among semantically related items is intact. However, it is important to note that the differentiation in N400 amplitude among sentence types could only be as differentiated as the presenting stimuli. It is possible that the degree of differentiation among sentence types was sufficiently large to demonstrate the N400 gradient. However, had the distinctions among related items been more subtle, the ERP gradient might not have been elicited in the PAD group.

It is also worth noting that it was not anticipated that N400 amplitude would differentiate between Related/Sense and Related/Nonsense (i.e., in YN and PAD groups, N400 amplitude elicited by Related/Nonsense was larger compared to that of Related/Sense). The following explanation may account for this outcome. The stimuli for Related/Nonsense were difficult to create, in that most words which were related to the sentences' best completions made sense with respect to the preceding sentence context. In order to compensate for this difficulty, the words selected may have been "less related" to the best completions in comparison to final words of Related/Sense. Thus, the distinction between Related/Sense and Related/Nonsense as a function of semantic relatedness may have been substantial. Since objective ratings were not obtained to assess the "relatedness" of the final words of Related/Sense, Related/Nonsense and Unrelated/Nonsense to their respective best completions, the degree of relatedness among sentence types is unknown.

F. Final comments

The present findings hold implications for semantic organization, semantic processing in the early stages of PAD, and the relationship between ERPs and semantic functioning. The ERP data provide converging evidence that the semantic system is-- or, under certain conditions can function as if it is- associatively structured and is organized according to the nature and strength of semantic relations. Although behavioral findings have supported this view, ERP data not only provide evidence from a different domain, but can be observed prior to and are not contingent upon overt behavioral responding. The fact that the ERP gradient is evident so early in time (the peak latency of N400 was approximately 400 msec, but its onset occured at about 200-250 msec post stimulus) and can appear intact when behavioral evidence argues for its dissolution, suggests that this organization exists at a very basic level, and is resistant to disruption.

Implications for semantic processing in PAD have already been discussed at length. The ERP data support the position that at some level, semantic information is maintained, although access and use of this information is problematic when attentional and retrieval demands are high.

With regard to the PAD group, it should be kept in mind that a number of subjects were excluded from data analysis due to excessive artifactual contamination. It is unknown whether the ERP gradient would have been observed in these patients. With the growth of the elderly population and consequently a rise in the prevalence of dementia, there has been a recent surge in aging and Alzheimer's research. The heterogeneity of findings across studies, in conjunction with the variability in cognitive profiles among individual patients has led a number of investigators to acknowledge that Alzheimer's Disease is not a uniform entity (Albert, et al., 1990; Rosen, 1983; Martin, 1987). Most likely, further research will clarify the distinctions among clinical subtypes of Alzheimer's Disease. With regard to the present investigation, it is possible that these six subjects represent a particular subgroup. The ability to successfully complete the task already distinguished these patients from those who were either excluded from analyses or had to discontinue testing during the experiment. Although excessive artifact was typically the immediate cause for their exclusion, the ability to inhibit these movements could be related to memory and/or attentional capacity or even anxiety level. According to the measures available (e.g., MMS score, WAIS-R scores) there were no obvious differences between patients who were included and those who were not. However, the response pattern of the PAD subjects observed in the current study might not be representative of most patients diagnosed as having (probable) "Alzheimer's Disease". Therefore, it is with caution that the present findings are generalized to the larger population of PAD patients.

Finally, the relationship between semantic priming and N400 was considered well established prior to the current study. However, the anomalous N400 pattern observed in the pilot study warns that this relationship is not immutable, and therefore, interpretations based on this relationship should be offered with caution. Had the protocol been run according to the original design (one practice block), and had the data not been examined for the effects of time on task, the results would have been interpreted very differently.

G. Suggested Future Studies

The current findings have generated new questions which could be explored in the following studies:

1) The data from the pilot study indicated that at some point during the course of three blocks of trials, there was a reduction in N400 amplitude elicited by Related/Sense stimuli. As discussed earlier, there is only one other report of changes in N400 amplitude associated with increased task experience (Fischler and Bloom, 1985). It was unclear from the current data

whether the reduction of N400 amplitude was gradual or sudden, or, whether it reflected an implicit or conscious strategy shift. The ERP data from the three practice blocks of the main experiment were not analyzed since these stimuli had not been normed. With properly normed stimuli and fewer trials per block, the time course of this shift could be explored. Perhaps systematic and detailed debriefing would reveal whether or not subjects are aware of any alterations in strategy.

2a) The NE data, particularly the prolonged RT and enhanced negativity elicited by Related/Sense, was interpreted as an effect of strategy superimposed upon a functionally intact semantic system. A speculation was made which related this unexpected response pattern to impaired word retrieval in elderly subjects when primed by semantically related words (Bowles and Poon, 1985). On the word retrieval task, the removal of correct primes from the stimuli abolished the inhibition previously elicited by the related primes (Bowles, 1989). Thus it was postulated that the superimposed negativity reflected strategy-related inhibitory processes. It would be interesting to observe whether the removal of Best Completion stimuli (similar to corrects primes) would eliminate the negativity associated with Related/Sense.

2b) An alternative means of inducing a strategy change in the NE group would be to explicitly inform NE subjects about the nature of the task, the stimuli, and the strategy they should adopt. Perhaps this manipulation would reduce the negativity associated with Related/Sense in the elderly adults.

2c) Since increasing the number of practice blocks reduced this negativity to Related/Sense in the YN group, it is possible that had more than three practice blocks been administered, the negativity would have also been reduced in the NE group. With additional well-normed stimuli, this question could be explored. In addition, the number of trials that would provide sufficient experience with the task to cause a strategy change and/or negativity reduction could be quantified.

3) Since performance by Wernicke's and Anomic aphasics on semantic/language task resembles the task-dependent performance of PAD patients, it would be interesting to observe whether the aphasics' ERP pattern on the present task would match that of the PAD subjects, or, show a pattern which would match their behavioral performance (as was initially predicted for PAD patients).

4) Although N400 amplitude elicited by Unrelated/Nonsense was greater than that of Related/Nonsense in all groups, the difference between the two was not statistically significant in any group. It was noted earlier that Related/Nonsense was difficult to construct. Using rating scales in which semantic relatedness would be empirically operationalized, it should be possible to create a more precise and reliable set of stimuli. If Related/Nonsense and Unrelated/Nonsense could elicit statistically distinct N400 amplitudes, the full stimulus set (i.e., all Sentence Types) would serve as a powerful instrument, capable of distinguishing four data points along a gradient. This would provide a basis from which to examine 1) the effects of different conditions on responses in normal young adults, 2) other clinical groups (e.g., aphasic and/or dyslexic individuals) under a set of controlled

conditions, and 3) levels of severity in these groups, and possibly longitudinal change.

# Appendix A

## Table A1
Set 1: Best Completion/Sense: Judgement frequencies (n=10)

| | | Yes | No |
|---|---|---|---|
| 1. | The gambler liked to roll the dice. | 10 | 0 |
| 2. | There are twenty-six letters in the alphabet. | 10 | 0 |
| 3. | The scientist studied the bacteria under a microscope. | 10 | 0 |
| 4. | The mouse ate the cheese. | 10 | 0 |
| 5. | The bride's face was covered by a veil. | 10 | 0 |
| 6. | Instead of wine the children drank grape juice. | 10 | 0 |
| 7. | He ate a plate of franks and beans. | 10 | 0 |
| 8. | He chopped down the tree with an axe. | 10 | 0 |
| 9. | The superstitious man would not walk under a ladder. | 10 | 0 |
| 10. | The little girl wanted a piggyback ride. | 10 | 0 |
| 11. | They started down the ski slope. | 10 | 0 |
| 12. | The old man sat on the park bench. | 10 | 0 |
| 13. | The gambler had a streak of bad luck. | 10 | 0 |
| 14. | The messenger bowed to the king and queen. | 10 | 0 |
| 15. | I went to the gym to get some exercise. | 10 | 0 |
| 16. | She bought a new pair of running shoes. | 10 | 0 |
| 17. | The construction worker wore a hard hat. | 10 | 0 |
| 18. | The child was born with a rare disease. | 10 | 0 |
| 19. | The lost dog was returned to his owner. | 10 | 0 |
| 20. | A bicycle has two wheels. | 10 | 0 |
| 21. | He loosened the tie around his neck. | 10 | 0 |
| 22. | The children held hands and formed a circle. | 10 | 0 |
| 23. | He put the ring on her finger. | 10 | 0 |
| 24. | The boat passed easily under the bridge. | 10 | 0 |
| 25. | The immigrant wanted to become a U.S. citizen. | 10 | 0 |
| 26. | He couldn't see without his glasses. | 10 | 0 |
| 27. | The ambulence rushed him to the hospital. | 10 | 0 |
| 28. | The performer bowed to the audience. | 10 | 0 |
| 29. | She put fresh sheets on the bed. | 10 | 0 |
| 30. | The children went to camp for the summer. | 10 | 0 |
| 31. | The paint turned out to be the wrong color. | 10 | 0 |
| 32. | The natives danced around the fire. | 10 | 0 |
| 33. | The open window let in some fresh air. | 10 | 0 |
| 34. | He shouted so loud he lost his voice. | 10 | 0 |
| 35. | The unlucky gambler lost all of his money. | 10 | 0 |
| 36. | The teacher wrote the problem on the board. | 10 | 0 |
| 37. | She called her husband at his office. | 10 | 0 |
| 38. | The medication caused harmful side effects. | 10 | 0 |
| 39. | Look both ways before crossing the street. | 10 | 0 |
| 40. | He went out to warm up the car. | 10 | 0 |

Table A2
Set 2: Related/Sense: Judgement frequencies (n=27)

| | Yes | No |
|---|---|---|
| 1. Let me take your hat and jacket. | 27 | 0 |
| 2. The magician waved his magic stick. | 27 | 0 |
| 3. Abby brushed her teeth after every lunch. | 27 | 0 |
| 4. The wound left a permanent mark. | 27 | 0 |
| 5. The horse collapsed right after the contest. | 27 | 0 |
| 6. The beggar asked for some spare cash. | 27 | 0 |
| 7. She tied up her hair with a yellow string. | 27 | 0 |
| 8. He fried an egg and two strips of pork. | 27 | 0 |
| 9. My father uses too much pepper and garlic. | 27 | 0 |
| 10. She carried her lunch in a picnic box. | 27 | 0 |
| 11. We had bacon and eggs for dinner. | 27 | 0 |
| 12. She liked to eat buttered popcorn at the play. | 27 | 0 |
| 13. She looked at herself in the glass. | 27 | 0 |
| 14. The academic year began in the spring. | 27 | 0 |
| 15. A cleaning woman scrubbed the wall. | 27 | 0 |
| 16. The boy put an apple on the teacher's table. | 27 | 0 |
| 17. My family moved from the country to the town. | 27 | 0 |
| 18. For a headache she takes two pills. | 27 | 0 |
| 19. He liked the scent of her new fragrance. | 26 | 1 |
| 20. The boys played cops and thieves. | 26 | 1 |
| 21. She had to change the baby's shirt. | 26 | 1 |
| 22. The athlete won a gold prize. | 26 | 1 |
| 23. I like hot fudge with ice milk. | 26 | 1 |
| 24. The first graders had milk and apples. | 26 | 1 |
| 25. She won a new washer and stove. | 26 | 1 |
| 26. Please wipe your feet on the towel. | 26 | 1 |
| 27. The magician pulled the rabbit out of the cap. | 26 | 1 |
| 28. The story had a happy beginning. | 26 | 1 |
| 29. The church had a beautiful stained glass door. | 26 | 1 |
| 30. He threw a penny into the wishing pool. | 26 | 1 |
| 31. He flushed the john. | 25 | 2 |
| 32. The unhappy marriage ended in a separation. | 25 | 2 |
| 33. She guessed the answer to the riddle. | 25 | 2 |
| 34. He looked up the word in the thesaurus. | 25 | 2 |
| 35. Most cats see very well at noon. | 25 | 2 |
| 36. In the winter everything is covered with white ice. | 25 | 2 |
| 37. The children watched the cars go down the path. | 25 | 2 |
| 38. I shaved off my mustache and hair. | 25 | 2 |
| 39. Joe hit the baseball with the club. | 25 | 2 |
| 40. Joan fed her baby some warm water. | 25 | 2 |

Table A3
Set 3: Related/Nonsense: Judgement frequencies (n=27)

| | | Yes | No |
|---|---|---|---|
| 1. | To strengthen his arms he lifted times. | 0 | 27 |
| 2. | All her music was on cassette glue. | 0 | 27 |
| 3. | They left the dirty dishes in the faucet. | 0 | 27 |
| 4. | We want to buy a microwave freezer. | 0 | 27 |
| 5. | The cow gave birth to the kitten. | 0 | 27 |
| 6. | The campers pitched a mansion. | 0 | 27 |
| 7. | The dog buried a flesh. | 0 | 27 |
| 8. | He got a tissue and blew his wrist. | 0 | 27 |
| 9. | The murderer received a life paragraph. | 0 | 27 |
| 10. | I fell off my bike and skinned my liver. | 0 | 27 |
| 11. | Don't cross the street when you see a red sound. | 0 | 27 |
| 12. | I look in the phone book to find the age. | 0 | 27 |
| 13. | She went to the grocers to do some flying. | 1 | 26 |
| 14. | John swept the floor with a pole. | 1 | 26 |
| 15. | He was murdered in cold serum. | 1 | 26 |
| 16. | The dentist pulled his wisdom tongue. | 1 | 26 |
| 17. | They rode down to the lobby in the stairs. | 1 | 26 |
| 18. | The cop loaded his knife. | 1 | 26 |
| 19. | The heavy rains caused a drought. | 1 | 26 |
| 20. | They enjoyed looking through the family cassette. | 1 | 26 |
| 21. | She let her daughter pierce her ankles. | 1 | 26 |
| 22. | She sipped the soda through a hay. | 1 | 26 |
| 23. | Bob covered his pancakes with maple trees. | 1 | 26 |
| 24. | The impatient driver honked his flute. | 1 | 26 |
| 25. | She told the brat to go stand in the edge. | 2 | 25 |
| 26. | The bowler knocked down all ten needles. | 2 | 25 |
| 27. | They raised pigs on their apartment. | 2 | 25 |
| 28. | The sleepy man had bloodshot feet. | 2 | 25 |
| 29. | Mail the package at the post home. | 2 | 25 |
| 30. | He was condemned to the electric couch. | 2 | 25 |
| 31. | The interview went well and he got the hobby. | 2 | 25 |
| 32. | They gossiped during their coffee crack. | 2 | 25 |
| 33. | I sewed on the button with a needle and rope. | 3 | 24 |
| 34. | The turtle pulled his head into its rock. | 3 | 24 |
| 35. | The skater fell on the cream. | 3 | 24 |
| 36. | She liked to paint with water shapes. | 3 | 24 |
| 37. | Chew before you swallow your beverage. | 3 | 24 |
| 38. | She was finally accepted to medical books. | 3 | 24 |
| 39. | On her diet she only lost one karat. | 3 | 24 |
| 40. | The shepard watched his flock of frogs. | 4 | 23 |

Table A4
Set 4: Unrelated/Nonsense: Judgement frequencies (n=10)

| | Yes | No |
|---|---|---|
| 1. The landlord came to collect the gaze. | 0 | 10 |
| 2. He mailed a letter without a lace. | 0 | 10 |
| 3. She sang soprano in the church beast. | 0 | 10 |
| 4. Pat blew out all the candles on the fox. | 0 | 10 |
| 5. She was glad they offered her maternity grill. | 0 | 10 |
| 6. The hat was decorated with ostrich thinker. | 0 | 10 |
| 7. The tree died after it was struck by puppet. | 0 | 10 |
| 8. The teller sold him a round-trip echo. | 0 | 10 |
| 9. We wrapped the leftovers in tin ghost. | 0 | 10 |
| 10. The pen ran out of fur. | 0 | 10 |
| 11. He dried his hands on a paper quarrel. | 0 | 10 |
| 12. The cat killed a field graph. | 0 | 10 |
| 13. She ate with a knife and bush. | 0 | 10 |
| 14. He liked lemon and sugar in his plug. | 0 | 10 |
| 15. The records were kept in a filing pioneer. | 0 | 10 |
| 16. She was afraid of the lightning and cocktail. | 0 | 10 |
| 17. She was leaving on a two week guerrilla. | 0 | 10 |
| 18. The butcher delivered an entire side of bath. | 0 | 10 |
| 19. The rude waiter was not given a mud. | 0 | 10 |
| 20. He wondered if the storm had done much protein. | 0 | 10 |
| 21. The typist indented at the begining of a furniture. | 0 | 10 |
| 22. Jane returned the books to the electron. | 0 | 10 |
| 23. The yoyo had a knot in the prince. | 0 | 10 |
| 24. The car had a flat myth. | 0 | 10 |
| 25. The Indian carried a bow and mayor. | 0 | 10 |
| 26. He walked up the steps and rang the egg. | 0 | 10 |
| 27. The house was surrounded by a picket meat. | 0 | 10 |
| 28. The tired runner stopped to catch his lake. | 0 | 10 |
| 29. After dinner they washed the jury. | 0 | 10 |
| 30. Birds fly south for the fashion. | 0 | 10 |
| 31. The girls got into a pillow beach. | 0 | 10 |
| 32. She had a cold and a sore truck. | 0 | 10 |
| 33. The little boy marched like a wooden belief | 0 | 10 |
| 34. He sailed across the seven lips. | 0 | 10 |
| 35. We heard the morning news on the element. | 0 | 10 |
| 36. I'll jot it down on a piece of river. | 0 | 10 |
| 37. The executive dictated a letter to his ground. | 0 | 10 |
| 38. The bill was due at the end of the face. | 0 | 10 |
| 39. He scraped the cold food from his noise. | 1 | 9 |
| 40. There wasn't any toothpaste left in the moon. | 1 | 9 |

# Appendix B

## A. Behavioral analyses for PAD Group without JM

Table B1
Mean RT in milliseconds and Percent Correct by Sentence Type in the PAD group, excluding subject JM

| Sentence Type | PAD (N=5) Reaction Time | PAD (N=5) Percent Correct |
|---|---|---|
| 1 Best Completion | 1075 (238) | 84.6 (15) |
| 2 Related/Sense | 1326 (195) | 66.8 (14) |
| 3 Related/Nonsense | 1365 (200) | 59.8 (13) |
| 4 Unrelated/Nonsense | 1316 (209) | 74.2 (19) |

### A1. Reaction Time

The removal of JM from the behavioral analyses did not alter the RT pattern as a function of Sentence Type. Despite the reduction in the sample size (N=5), the effect of Sentence Type was significant ($F(1/6)=11.92$, $p<.01$). Paired comparisons indicated the following distinctions: Best Completion< Unrelated/Nonsense, Related/Sense, Related/Nonsense ($p<.05$).

### A2. Percent Correct

Similar to RT, the effect of Sentence Type on Percent Correct was reliable despite the small sample size ($F(2/8) = 8.38$, $p<.01$). The ordering and significance levels were unchanged. Post hoc tests revealed the following pattern: Best Completion > Unrelated/Nonsense, Related/Sense> Related/Nonsense ($p<.05$).

A3. ERPs

Unlike the behavioral analyses, none of the ERP effects reached statistical significance with the reduced sample of 5 subjects. Nonetheless, the ERP patterns were maintained (see figures 12 and 15).

# Appendix C

A. Visual N1
A1. MIdline Analyses

Table C1
N1 amplitude in μV averaged across midline electrodes for each group

| | Amp |
|---|---|
| YN | -3.1 |
| NE | 0.1 |
| PAD | 1.2 |

Amplitude values for the visual N1 across midline electrodes are shown in table C1. N1 amplitude showed the greatest negativity in the YN group, and was most positive in the PAD group. The main effect of Group was significant ($F(2/23) = 17.11, p<.001$), but did not interact with any other variables. Although there were no main effects of Sentence Type or Electrode, the two-way interaction of these variables was significant ($F(3/78) = 2.88, p<.03$). Simple effects procedures indicated that the effect of Sentence Type on N1 was significant only at Pz ($F(3/64) = 2.98, p<.04$). N1 amplitudes at Pz for each sentence type are shown in table C2. Post hoc tests showed the following pattern: BC > RN > RS , UN.

There were no other significant interactions for N1 at midline electrodes.

Table C2
N1 amplitude in μV at Pz (collapsed across groups) by Sentence Type.

| Sentence Type | N1 at Pz |
|---|---|
| BC | -1.8 |
| RS | -0.8 |
| RN | -1.3 |
| UN | -0.7 |

A2. Lateral Analyses

N1 amplitudes was maximal at O1 and O2. Again, the main effect of Group (order identical to that shown at the midline) was significant ($F(2,23) = 4.95$, $p<.02$), as were the effects of Sentence Type ($F(2,57) = 3.09$, $p<.04$) and Hemisphere ($F(1,23) = 7.01$, $p<.01$). These effects were modulated by the three-way interaction of Group, Sentence Type and Hemisphere ($F(6,65) = 5.68$, $p<.03$). However, when tested separately within each group, Sentence Type was no longer significant, and only the NE Group showed a significant effect of Hemisphere ($F(1,23) = 9.63$, $p<.005$) such that N1 was larger over the left hemisphere (left: -3.9uv, right: -2.5uv).

In summary, N1 amplitude was largest in the YN group (collapsed across Sentence Types) and largest to Best Completion and Related/Nonsense (collapsed across groups) relative to other Sentence Types. N1 amplitude patterns as a function of Sentence Type did not show differential patterns among Groups.

B. <u>P600</u>

B1. <u>MIDLINE Analyses</u>

As can be see in Figures 13, 14, and 15, P600 amplitude patterns as a function of Sentence Type were different for each group. These values are listed in table C3.

Table C3
P600 amplitude in μV averaged across midline electrodes by Sentence Type

| | Young Normal | Normal Elderly | PAD |
|---|---|---|---|
| BC | 5.0 | 4.7 | 4.0 |
| RS | 7.1 | 2.7 | 1.2 |
| RN | 3.6 | 5.4 | 3.4 |
| UN | 5.8 | 6.9 | 1.0 |

The following descending order of Sentence Types for P600 amplitude by Group was observed:

YN: RS, UN, BC, RN
NE: UN, RN, BC, RS
PAD: BC, RN, RS, UN

There were no main effects of Group, Sentence Type, or Electrode on P600. However, there were significant interactions of Group and Sentence Type, ($F(4/48) = 3.24$, $p < .01$) and Group and Electrode ($F(4/41) = 3.95$, $p < .01$). Tests for simple effects revealed that neither the effects of Electrode nor Sentence Type were reliable in the YN or PAD groups. By contrast, in the NE group, both Sentence Type ($F(2/48) = 3.26$, $p < .04$) and Electrode ($F(2/41) = 6.78$, $p < .004$) were significant, as was the interaction of these variables ($F(3/69) = 4.88$, $p < .003$). Simple effects procedures also indicated that for the

NE group, the effect of Sentence Type on P600 was significant at both Cz (F(2/19) = 3.98, p < .03) and Pz (F(3/24) = 4.31, p < .01) but not Fz, such that: UN, BC, RN > RS (p<.05). The reduced P600 amplitude to Related/Sense may reflect either latency jitter, equivocation, or reduced confidence, (or a combination these) since RT was longest (Related/Sense: 1092 msec compared to Related/Nonsense: 992, Unrelated/Nonsense: 892, and Best Completion: 789) and within-subject SDs were greatest for Related/Sense (261 msec compared to RN: 231, UN:193 , and BC: 216; see Appendix E).

B2. Hemisphere Analyses

All lateral electrodes (F3/F4, C3/C4, P3/P4, T5/T6, and O1/O2) were included in the Hemisphere analyses for P600. Consistent with the midline analyses, there were no main effects of Group or Sentence Type on P600. Similarly, there were no main or interaction effects of Hemisphere.

The interaction of Group and Sentence Type was significant (F(5/52) = 3.11, p<.02). Similar to the midline analyses, tests for simple effects indicated that Sentence Type was significant only in the NE Group (F(2/52) = 3.85, p<.02). This effect interacted with Electrode Site (F(5/126) = 3.71, p<.01), due to increased positivity at frontal electrode sites in the elderly relative to younger adults. This correlate of aging has been described by other investigators, (e.g., Friedman et al., 1989; Picton et al., 1984).

In sum, P600 amplitude was not reliably responsive to experimental variables in the YN and PAD groups. However, in the NE group P600 elicited by Related/Sense was smaller relative to that elicited by other Sentence Types.

C. SW1

C1. Midline Analyses

Table C4
SW1 amplitude in μV at Pz by Sentence Type

| Sentence Type | YN | NE | PAD |
|---|---|---|---|
| Best Completion | 4.4 | 5.2 | 2.0 |
| Related/Sense | 7.9 | 1.4 | -1.9 |
| Related/Nonsense | 2.7 | 4.5 | 2.3 |
| Unrelated/Nonsense | 5.7 | 7.9 | -0.6 |

Amplitude values of SW1 at Pz are shown in table C4. Although there were no main effects of Group, Sentence Type or Electrode on SW1, the interactions of Group and Sentence Type ($F(5/60) = 3.24$, $p < .01$) and Electrode and Sentence Type ($F(3/66) = 5.82$, $p < .001$) were significant. Tests for simple effects indicated that Sentence Type was significant in the YN group ($F(3/60) = 3.77$, $p < .02$). Post hoc test showed the following pattern: RS, UN, BC, RN (similar to ordering of P600) such that only Related/Sense and Related/Nonsense were significantly different from one another. Sentence Type interacted with Electrode in both NE ($F(3/66) = 4.88$, $p < .003$) and PAD groups ($F(3/66) = 4.14$, $p < .01$). Within the PAD group however, Sentence Type did not reach significance at any electrode site. By contrast, in the NE group, Sentence Type was significant at Cz ($F(2/17) = 5.15$, $p < .02$) and Pz ($F(2/18) = 4.50$, $p < .02$) only. Post Hoc tests revealed the following pattern ($p<.05$): UN > BC, RN > RS.

The three-way interaction of Group, Sentence Type and Electrode was also significant ($F(6/66) = 2.59$, $p < .02$), reflecting the differential patterns of Sentence Type and Sentence Type by Electrode effects in each group (detailed above).

C2. Hemisphere Analyses

Amplitude patterns of SW1 at lateral sites were identical to those reported at the midline. Similar to P600, there were no main effects of Group, Sentence Type, or Hemisphere, or interactions of Hemisphere with any variables on SW1. The interaction of Group and Sentence Type was significant ($F(5/57=4.68$, $p<.001$), and as indicated by simple effects procedures, Sentence Type was significant in the YN ($F(2/57 = 3.29$, $p<.04$) and NE ($F(2/57 = 4.03$, $p<.02$) Groups, but not in the PAD group.

D. SW2

D1. Midline Analyses

Table C5
SW2 amplitude in μV at Pz by Sentence Type

| Sentence Type | Young Normal | Normal Elderly | PAD |
|---|---|---|---|
| Best Completion | 1.5 | 4.6 | -0.7 |
| Related/Sense | 5.5 | 3.1 | -3.7 |
| Related/Nonsense | 2.2 | 6.5 | 1.6 |
| Unrelated/Nonsense | 3.7 | 7.2 | -4.7 |

Ordering of Sentence Types for SW2 amplitude was identical to that of P600 and SW1. Amplitude values are shown in table C5. Although there were no main effects of Group, Sentence Type or Electrode on SW2, the interactions of Group and Sentence Type ($F(6/65) = 4.26$, $p < .001$), and Group, Sentence Type and Electrode ($F(8/88) = 3.45$, $p < .002$) were significant. Further examination of these interactions within each group revealed that there were no main effects or interaction effects on SW2 in YN or NE groups. In the PAD group, the interaction of Sentence Type and Electrode was significant ($F(4/88) = 4.57$, $p < .002$). However, when tested separately at each electrode, Sentence Type was not significant. (This may be due to the small sample of 6 subjects in this group and/or increasing variability as the latency interval under observation is further removed from the time of stimulus presentation).

D2. Hemisphere Analyses

Table C6
SW2 amplitude in μV averaged across right and left hemispheres electrodes

| | Left | Right |
|---|---|---|
| YN | 2.0 | 2.9 |
| NE | 2.0 | 3.4 |
| PAD | -2.6 | -2.7 |

Amplitude measures of SW2 collapsed across electrode sites separately for right and left hemispheres for each group are shown in table C6.

Hemisphere was the only main effect on SW2. Although SW2 amplitude was more positive over the right hemisphere only in the YN and NE groups, hemisphere did not interact with any of the experimental variables. However, the interaction of Group and Sentence Type was significant (F(5/62 = 4.29, p<.001). Analyses within each group indicated that the effect of Sentence Type was significant only in the YN (F(2/57) = 3.29, p<.04) and PAD Groups (F(3/62 = 4.79, p<.006).

E. N400 Lateral Analyses (N4b reported in Results)

Analyses of data including all lateral sites, followed by simple effects procedures examining effects separately within each group indicated that there were no interactions of Sentence Type, Hemisphere and Electrode in any group. Since N400 amplitude measures were largest posteriorly, further analyses were performed at T5/T6 and O1/O2 locations only. These latter analyses are presented below.

E1. N4a

N4a amplitude within each Group was generally more negative over the left hemisphere. The effects of Group, Sentence Type (ST) and Hemisphere on N4a, N4c, and N4tot are shown in tables C7.

Table C7
Probability and F-values of experimental effects for N4a, N4c, and N4tot

| N4a | Deg. F. | F-value | p < |
|---|---|---|---|
| Group | 2, 23 | 10.15 | .001 |
| ST | 3, 59 | 3.11 | .04 |
| Hemisphere | 2, 49 | 3.02 | .01 |
| Group x ST | 5, 59 | 3.27 | .01 |
| ST x Hemisphere | 2, 49 | 4.54 | .01 |
| Group x ST x Hemi | 4, 49 | 3.02 | .02 |

| N4c | Deg. F. | F-value | p < |
|---|---|---|---|
| ST | 3, 62 | 10.14 | .001 |
| Group x ST | 5, 62 | 3.12 | .01 |

| N4tot | Deg. F. | F-value | p < |
|---|---|---|---|
| ST | 2, 56 | 11.23 | .001 |
| ST x Hemisphere | 2, 50 | 6.36 | .003 |

Similar to N4b (detailed in the Results), N400tot amplitude increased in negativity from Best Completion to Unrelated/Nonsense except left hemisphere in the NE group.

Further analyses revealed that the effect of Sentence Type collapsed across T5 and T6 was significant in all three groups as shown in table C8.

Table C8
Within group probability and F-values for the effect of Sentence Type across T5 and T6 electrodes on portions of the N400

YN

| | Deg. F. | F-value | p < |
|---|---|---|---|
| N4c | 3, 65 | 13.06 | .001 |
| N4tot | 3, 58 | 7.04 | .001 |

NE

| | Deg. F. | F-value | p < |
|---|---|---|---|
| N4c | 3, 65 | 2.82 | .05 |

PAD

| | Deg. F. | F-value | p < |
|---|---|---|---|
| N4a | 3, 62 | 6.00 | .002 |
| N4tot | 3, 58 | 4.54 | .01 |

Due to the enhanced left hemisphere negativity elicited by Related/Sense in the NE Group, the interaction of Sentence Type and Hemisphere was significant only in this Group. This effect applied to all portions of the N400 (see table C9).

Table C9
Probability and F-values of the interaction of Sentence Type and Hemisphere in the NE group

| | Deg. F. | F-value | p < |
|---|---|---|---|
| N4a | 2, 46 | 6.15 | .004 |
| N4b | 2, 51 | 4.54 | .001 |
| N4c | 2, 56 | 8.01 | .001 |
| N4tot | 3, 69 | 9.86 | .001 |

In sum, the results of the lateral analyses for N400 were similar to those reported in the results section. Generally, N400 increased in amplitude from Best Completion to Unrelated/Nonsense, except for the enhanced negativity elicited by Related/Sense in the NE group,which was largest at left posterior electrode sites.

# Appendix D

## Practice Stimuli

Practice Block 1
(S = "sense", N = "nonsense")

(S) He had to rake the leaves.
(S) He was reading a paperback novel.
(N) He shouted at the top of his steak.
(N) Don't touch the wet flight.
(S) Fred put the worm on a hook.
(N) You can't open the door with the wrong grass.
(S) They rested under a tree in the dark.
(N) Nothing can beat a bowl of hot logs.
(S) The sun is just another star.
(N) The child learned to count to building.
(N) The thief was caught and sent to minerals.
(S) Betsy could never tell a lie.
(S) A dog has a good sense of smell.
(N) The rabbit hid in the tall key.
(S) The earth is shaped like a ball.
(N) Brian poured some sauce on his rare cotton.
(N) David's shirt was made of lungs.
(S) The hunter shot and killed a large deer.
(S) Suzy liked to play with her toy dolls.
(S) The detectives searched for a clue.
(S) The bird of peace is the dove.
(N) The cabin was made of soup.
(N) The dealer shuffled the cards.
(S) Banks keep their money in a vault.
(N) The swimmer dove into the spleen.
(N) The lion is the king of the birthday.
(S) The house was made of red brick.
(N) Success is often just a matter of hard bell.
(S) The actress gave ten percent to her agent.
(N) The teenager talked on the scissor.
(N) Every morning she walked the cake.
(S) She wore a garter on her leg.
(N) The visitors rang the work.
(N) There were only three days until his jungle.
(S) The farm had a red barn.
(S) She hooked a leash to the dog's collar.
(S) The surgeon scrubbed up before the operation.
(N) The cereal was fortified with vitamins and jail.
(S) The pirate buried a treasure chest.
(S) The little lost girl cried for her mother.

Practice Block 2

(BC) He adjusted the antennae to get better reception.
(RN) The best man raised his glass and made a bread.
(RS) His hat was blown away by the breeze.
(BC) The bird hatched her eggs in the nest.
(RN) From the roof they had a bird's eye smell.
(RN) The knight wore a suit of swords.
(RS) The busy man needed some rest and peace.
(UN) He swung at the ball and ran to first deer.
(RS) She covered her toast with strawberry sauce.
(BC) She needed a cookbook to look up the recipe.
(RN) Keep the sherbet in the toaster.
(RN) She put a dime in the parking timer.
(BC) For information I called the operator.
(UN) Banks keep their money in a rehearsal.
(BC) Grandmother knitted a cardigan sweater.
(UN) We marched in the Thanksgiving day collar.
(BC) After their wedding, the newlyweds went on their honeymoon.
(BC) The country was destroyed by an atomic bomb.
(UN) The new homeowner called an interior cards.
(UN) The executive suffered from high blood alley.
(RS) On Valentine's day she received a dozen daisies.
(RS) She liked to go to parties on New Year's afternoon.
(BC) All packaged foods must list their ingredients.
(RN) The bird of peace is the squirrel.
(RN) Instead of glasses, she wore contact slides.
(RS) The prisoners longed for their liberty.
(RN) The case was taken to small claims yard.
(RN) The couple celebrated their silver wedding birthday.
(UN) Every Sunday they went to the bowling pressure.
(RN) Suzy waited for Santa to slide down the ceiling.
(BC) The bird couldn't fly because of its broken wing.
(BC) The couple agreed that it was love at first sight.
(UN) The hunter shot and killed a large base.
(UN) The detectives searched for a lotion.
(UN) She hooked a leash to the dog's parade.
(RS) The tenant wanted a two year contract.
(RN) She weighed herself by stepping on the ruler.
(RS) The dentist told Billy he found only three molars.
(RS) Her parents taught her to always tell the facts.
(UN) The dealer shuffled the decorator.

Practice Block 3

(BC) The Englishman ate fish and chips.
(BC) He used an air pump to inflate the tire.
(UN) The artist painted a self traffic.
(RS) We picked grapes right off the plant.
(UN) The caterpillar spun a small captain.
(RS) Mother told me not to talk to people.
(RS) He challenged his enemy to a fight.
(UN) She sprinkled the baby with talcum quilt.
(BC) He ordered a bottle of red wine.
(RN) Elephants have ivory whiskers.
(RN) Remove the splinter with a wrench.
(BC) The children built a sand castle.
(BC) The palace was surrounded by a moat.
(RS) The fisherman ran out of worms.
(RS) Every night he watched the ten o'clock sports.
(UN) She covered her insect bites with calamine clue.
(UN) The lamp needed a new light property.
(RN) She entered the pig in the county wedding.
(RN) The zebra had black and white ovals.
(RS) He caught his toe in the mouse cage.
(RS) For the interview he wore his best slacks.
(RN) The little girl's dress had polka dashes.
(BC) The little girl had dimples in her cheeks.
(UN) The drunk driver was in the opposing lane of portrait.
(BC) The cat clawed at the scratching post.
(BC) California is on the west coast.
(UN) The crew declared mutiny against their cocoon.
(UN) Betsy sewed a patchwork powder.
(RS) The runner won the relay marathon.
(RN) We ate corn on the petal.
(BC) He opened the oyster and found a pearl.
(BC) The watch stopped because it needed a new battery.
(UN) It is illegal to trespass on private bulb.
(UN) The cast prepared for the dress vault.
(RS) The trapeze artist fell into the mesh.
(RS) She kneeled beside her bed and said her poems.
(RS) The optimist wouldn't give up dreams.
(RN) The surfer liked to ride the sand.
(RN) The guard sounded the lock.
(RN) When the men were introduced they shook toes.

# Appendix E
## Analyses of Within-Subject SD

A. Pilot Study
A1. Session 1

Table E1
Mean RT and Within-subject Standard Deviation by Sentence Type

| | BC | RS | RN | UN |
|---|---|---|---|---|
| RT (msec) | 770 | 1132 | 955 | 840 |
| WSD | 171 | 270 | 208 | 129 |

A1a. Within-subject Standard Deviation of RT (WSD)

Similar to RT, WSD met assumptions of homogeneity of variance among Sentence Types. As shown (see table E1), the greatest variability in RT was elicited by Related/sense, and the least amount of variability was elicited by Unrelated/nonsense. The effect of Sentence Type on WSD was significant ($F(3,16) = 3.54$, $p<.03$). However, paired comparisons indicated that only Related/sense and Unrelated/nonsense (i.e., highest and lowest values) were statistically distinct.

A2. Session 1 versus Session 2

There were no significant changes in WSD from Session 1 to Session 2.

## B. Main Experiment

### B1. Group Effects

Collapsed across Sentence Types, there was a main effect of Group on WSD ($F(2/23) = 10.07$, $p < .001$) As shown in table E2, the PAD group demonstrated the greatest within-subject variabitity in response time. Pairwise comparisons confirmed that across Sentence Types, YN and NE were similar to each other, but distinct from the PAD group ($p < .001$).

Table E2
Within-Subject SD collapsed across Sentence Types within each group

| Group | Mean WSD (SD) |
| --- | --- |
| YN | 182 (74) |
| NE | 225 (72) |
| PAD (6) | 318 (54) |
| PAD (5) | 313 (47) |

Table E3
Mean Within-subject SD by Sentence Type

| ST | Y N | NE | PAD (N=6) |
| --- | --- | --- | --- |
| 1 BC | 177 | 216 | 337 |
| 2 RS | 178 | 261 | 323 |
| 3 RN | 225 | 232 | 316 |
| 4 UN | 149 | 193 | 294 |

B2. Sentence Type

The main effect of Sentence Type ($F(3/64) = 5.36$, $p<.003$) on WSD was significant (values shown in table E2). Although the interaction of Group and Sentence Type was not significant, when examined separately within groups the effect of Sentence Type was reliable only in the YN ($F(2/19) = 6.08$, $p<.01$) and NE Groups ($F(2/20) = 5.19$, $p<.01$). Newman-Keuls post hoc tests showed the following patterns ($p<.05$):

YN: RN > RS,BC > UN
NE: RS > RN,BC > UN

# Appendix F

## Analyses of Homogeneity of Variance

A. Pilot Study

A1. Session 1

A1a. Quartile Analysis

Assumptions of homogeneity of variance for N400 amplitude (F3/56)=2.42, p<.09) and RT (F3/56)=2.04, p<.08) were met.

A1a. New subjects

Homogeneity of variance was indicated for the six addtional subjects for N400: (F(3/68)=1.66, p<.18)

B. Main experiment

B1. Reaction Time

Mean values of behavioral measures for each Group, collapsed across Sentence Types, are shown in table 15 (see Results section). Although variability seemed to increase across groups, a test for homogeneity of variance indicated that differences among groups did not approach significance (F(8/21) = 1.24, critical value = 2.42). Tests for homogeneity within each group among Sentence Types also indicated equality of variances; YN: F(3/36) =.57, p<.63, NE: F(3/36) =.15, p<.9; PAD: F(3/20) =.01, p<.99. Therefore, the RT data were not transformed and are presented in their original form. Additionally, it has been demonstrated that transformation rarely affects statistical outcome, and it is preferable to use the raw RT since it reflects the actual time taken to perform a particular cognitive operation

(Shoben, 1982). Moreover, in the context of the present study, the absolute RT values were less important than the distribution of RT across Sentence Types.

B2. N400

N400 amplitude (all segments) was tested for homogeneity of variance among groups, and among Sentence Types within each group, with the data collapsed across all scalp electrodes. Among groups (collapsed across Sentence Types assumptions of homogeneity were met ($C(3/9)=.5317$, critical value=.6912, $p<.01$), as assessed by Cochran's test (Cochran, 1941, in Kirk, 1969). Although homogeneity was indicated among Sentence Types in the NE group for N4b: ($F(3/516) = 1.46$, $p<.22$), N4c: ($F(3/516) = .82$, $p<.48$), and N4d: ($F(3/516) = 1.17$, $p<.32$), and in the PAD group for N4b: ($F(3/308) = 1.66$, $p<.17$), no portion of the N400 in the YN group showed homogeneity among Sentence Types. Therefore, these data were subjected to log transformation prior to analyses for the effect of Sentence Type (Because logarithmic transformation can only be performed on positive values, a positive value (+23), i.e., the absolute value of the most negative amplitude in the data, was added to each data point. Logarithmic transformations were then performed on these positive data). As a result of these transformations, homogeneity was indicated for log(N4b): ($F(3/516) = 1.47$, $p<.22$) and log(N4c): ($F(3/516) = 2.42$, $p<.07$) among Sentence Types in the YN group. Analyses were conducted on these transformed data, although raw amplitude measures are presented in the text.

## Appendix G

Sentence contexts, completions, and Cloze probabilities for sentence contexts which generated completions with cloze probabilites greater than .84.

Sentence completions are presented in upper-case letters. Cloze probabilities in parentheses are based on completions of 100 volunteers; "*" indicates response given by subject(s) in elderly group only. Illegible responses were omitted.

1. She sipped the soda through a STRAW(100)
2. Joe hit the baseball with the BAT(100)
3. Don't cross the street when you see a red LIGHT(100)
4. We washed our hands with soap and WATER(100)
5. He put the ring on her FINGER(100)
6. The church had a beautiful stained glass WINDOW(100)
7. He fried an egg and two strips of BACON (100)
8. The tired runner stopped to catch his BREATH(100)
9. The story had a happy ENDING(100)
10. The Indian carried a bow and ARROW(100)
11. The impatient driver honked his HORN(100)
12. The dentist pulled his wisdom TOOTH(100)
13. The magician pulled the rabbit out of the HAT(100)
14. The fly got caught in the spider's WEB(100)
15. The pen ran out of INK (100)
16. The old man sat on the park BENCH (100)
17. Birds fly south for the WINTER (100)
18. She looked at herself in the MIRROR (100)
19. The dog buried a BONE (100)
20. She won a new washer and DRYER (100)
21. He looked up the word in the DICTIONARY (100)
22. The medication caused harmful side EFFECTS (100)
23. The open window let in some fresh AIR(100)
24. She put fresh sheets on the BED(100)
25. He got a tissue and blew his NOSE(100)
26. Bob covered his pancakes with maple SYRUP(100)
27. He kept his wallet in his jacket POCKET(100)
28. The car had a flat TIRE(100)
29. There are twenty-six letters in the ALPHABET(99), MAIL(1)
30. The poster was hung with masking TAPE(99), PAPER(1)
31. Look both ways before crossing the STREET(99), ROAD(1) *GUTTER(1)
32. The gift was in a cardboard BOX(99), CONTAINER(1)
33. The boys played cops and ROBBERS(99), THIEF(1)
34. The noise woke her in the middle of the NIGHT(99), AFTERNOON(1)

35. He got a ticket for going over the speed LIMIT(99), ZONE(1)
36. She ate with a knife and FORK(99), SPOON(1)
37. He threw a penny into the wishing WELL(99), FOUNTAIN(1)
38. The house was surrounded by a picket FENCE(99), LINE(1)
39. He chopped down the tree with an AXE(99), BLADE(1)
40. I sewed on the button with a needle and THREAD(99), STRING(1)
41. The messenger bowed to the king and QUEEN(99), LEFT(1)
42. He dried his hands on a paper TOWEL(99), CLOTH(1), *NAPKIN(1)
43. The murderer received a life SENTENCE(99), PRISON(1)
44. I'll jot it down on a piece of PAPER(99),PIECE(1)
45. She solved the crossword PUZZLE(99), GAME(1)
46. He sailed across the seven SEAS(99), LAKES(1)
47. She let her daughter pierce her EARS(98), EARLOBES(2)
48. He was condemned to the electric CHAIR(98), BILL(2)
49. The gambler liked to roll the DICE(98), DECK(1), MONEY(1)
50. He flushed the TOILET(98), JOHN(1)
51. The teller sold him a round-trip TICKET(98), CRUISE(1), TRIP(1),*FARE(1)
52. She polished her finger NAILS(98), BOWLS(1), RING(1)
53. She was afraid of the lightening and THUNDER(98), RAIN(2),
54. The little boy marched like a wooden SOLDIER(98),MARIONETTE(1), TOY(1), GUN(1)
55. The construction worker wore a hard HAT(98), HELMET(2)
56. The bowler knocked down all ten PINS(98), BALLS(1), WICKETS(1)
57. The girls got into a pillow FIGHT(98), SACK(1), TALK(1)
58. The immigrant wanted to become a US CITIZEN (98), SENATOR(1), PRESIDENT(1)
59. On her diet she only lost one POUND(98), INCH(1), OUNCE(1)
60. She sang saprano in the church CHOIR(98), AUDITORIUM(98), CHORUS(1), *BALCONY(1)
61. The little girl wanted a piggyback RIDE(98), RACE(1), BANK(1)
62. The scientist studied the bacteria under a MICROSCOPE(98), LENS(1), SLIDE(1)
63. The superstitious man would not walk under a LADDER(98), PLANK(1), SCAFFOLD(1)
64. She bought a new pair of running SHOES(97), SNEAKERS(1), SHORTS(1),*SKATES(1)
65. The hero was given a twenty-one gun SALUTE(97), SHOT(2), SHELLS(1)
66. He couldn't see without his GLASSES(97), EYEGLASSES(2), EYES(1)
67. She went to the grocers to do some SHOPPING(97), FAVORS(1), MARKETING(1), FOOD(1)

68. They rode down to the lobby in the ELEVATOR(97), CAR(1), HOTEL(1), SHUTTLE(1)
69. They played catch with a tennis BALL(97), RACQUET(2), PLAYER(1)
70. The ambulance rushed him to the HOSPITAL(97), ACCIDENT(1),EMERGENCY ROOM(1), SCENE(1),*CLINIC(1)
71. The magician waved his magic WAND(96), STICK(2), BATON(1), BAT(1)
72. Pat blew out all the candles on the CAKE(96), BIRTHDAY CAKE(3), TABLE(1)
73. They enjoyed looking through the family ALBUM(96),PHOTOS(1), PHOTO ALBUM(1), TREE(1), PICTURES(1)
74. Let me take your hat and COAT(96), GLOVES(3)
75. She put the flowers and water in a VASE(96),GRAVE (1), JAR(1), PITCHER(1), POT(1)
76. The boy put an apple on the teacher's DESK(96),HEAD (3),TABLE(1)
77. The turtle pulled his head into its SHELL(96), NECK(2), BACK(1), BODY(1)
78. The wound left a permanent SCAR(96), MARK(2), DAMAGE(1), LIMP(1),*INJURY(1)
79. For a headache she takes two ASPIRIN(96), ANACIN(1), BUFFERIN(1), PILLS(1), TABLETS(1)
80. She told the brat to go stand in the CORNER(96), DARK(1), LINE(1), ROAD(1), STREET(1)
81. The landlord came to collect the RENT(96), INSURANCE(2), MONEY(1), RENT MONEY(1)
82. The athlete won a gold MEDAL(96), RIBBON(1), GLOVE(1), RING(1), *TROPHY(1), *CHAIN(1), *UNIFORM(1), *STAR(1)
83. He liked the scent of her new PERFUME(96), COLOGNE(1), DEODORANT(1), FLOWERS(1), FRAGRANCE(1), *POWDER (1)
84. The prince awakened her with a KISS(95), JOLT (1), SONG(1), PRINCESS(1), GUN(1), YELL(1)
85. A cleaning woman scrubbed the FLOOR(95), BATHROOM(1), KITCHEN(1), RUG(1), SINK(1), STOVE(1)
86. I look in the phone book to find the NUMBER(95), ADDRESS(3), ANDREWS(1), PHONE#(1)
87. The unhappy marriage ended in a DIVORCE(95), SEPARATION(2), FIGHT(1), BREAK(1)
88. They gossiped during their coffee BREAK(95), HOUR(4), CLOTCH(1)
89. The tree died after it was struck by LIGHTENING(95), CAR(5), *TRUCK(1)
90. She liked to eat buttered popcorn at the MOVIES(95), CINEMA(2), PARTY(2), THEATRE(1), STORE(1), *GAME(1)

91. The brides face was covered by a VEIL(95), FROWN(1), MASK(1), SCAR(1), SMILE(1)
92. He shouted so loud he lost his VOICE(95), HEARING(3), BALANCE(1), HEAD(1), HAIR(1), *BREATH(1)
93. The poor student was expelled from SCHOOL(95), CLASS(4)
94. There wasn't any toothpaste left in the TUBE(95), BATHROOM (2), CABINET(2), HOUSE(1), MEDICINE CHEST(1)
95. A bicycle has two WHEELS(95), PEDALS(4), HANDLEBARS(1)
96. She carried her lunch in a picnic BASKET(95), BOX(4), BAG(1)
97. The sleepy man had bloodshot EYES(92), WOUND(5)
98. Instead of wine the children drank grape JUICE(94), SODA(4), COOLER(1), DRINK(1)
99. The hat was decorated with ostrich FEATHERS(94), COLORS(1), LINING(1), PICTURES (1), SHELLS(1), EGGS(1), *PLUME(1)
100. He was murdered in cold BLOOD(94), NIGHT(2), WEATHER(2), DAYLIGHT(1), RAIN(1)
101. She had to change the baby's DIAPER(94), PAMPERS(4), CLOTHES(1), NAME(1), PANTS(1)
102. The glee club sang a SONG(94), TUNE(4), CHORUS(1), HYMN(1)*CONCERT(1), *BALLAD(1), *MADRIGAL(1), *MELODY(1)
103. Chew before you swallow your FOOD(94), GUM(4), MEAT(2)
104. The first graders had milk and COOKIES(94), CEREAL(2), CRACKERS(2), BISCUITS(1), BREAD(1)
105. We had bacon and eggs for BREAKFAST(94), DINNER(4), LUNCH(2)
106. He went out to warm up the CAR(94), ENGINE(2), BIKE(1), NIGHT(1), BAR-B-QUE(1), COFFEE(1)
107. The mouse ate the CHEESE(93), APPLE(1), CAT(1), FLY(1), POISON(1), CRUMBS(1), ROACH(1), TERMITE(1)
108. The unlucky gambler lost all of his MONEY(93), CHIPS(1), SAVINGS(1), EARNINGS(1), SHIRTS(1), POSSESSIONS(1), *BELONGINGS(1), *WINNINGS(1)
109. The shepherd watched his flock of SHEEP(93), GEESE(2), LAMBS(2), GOATS(1), ANIMALS(1)
110. The cop loaded his GUN(93), PISTOL(4), REVOLVER(3)
111. The beggar asked for some spare CHANGE(93), COINS(2), FOOD(2), MONEY(2), RIBS(1), *DIMES(1)
112. The children went to camp for the SUMMER(93), WEEK(2), WEEKEND(1), DAY(1), MONTH(1), FIRST TIME(1)
113. The fireman unwound the HOSE(93), COIL(1), CLOCK(1), CORD(1), ENGINE(1), ROPE (1), VICTIM(1), FIRE(1), *PIPE(1)
114. The car stopped when I stepped on the BRAKE(93), CURB (2), STREET(2), GAS(1) GAS PEDAL(1), PEDAL(1), *TREADLE(1), *ROADWAY(1)

115. The cow gave birth to the CALF(93), BABY(3), HEFER(2), BABY COW(1), COW(1), *CUB(1)
116. He walked up the steps and rang the BELL(92), DOORBELL(7), GONG(1), BUZZER(1)
117. The heavy rains caused a FLOOD (92), STORM(5), DOWNPOUR(1), PUDDLE(1), FLOODING(1), JAM(1)
118. The executive dictated a letter to his SECRETARY(92), COLLEAGUES(2), DICTAPHONE(2), CLIENT(2), PEOPLE(1), BOSS (1)
119. The cat killed a field MOUSE(92), RAT(7), OF MICE(1), *SPARROW(1)
120. If you ride bareback you don't need a SADDLE(92), SHIRT(5), HORSE(2), CUSHION(1)
121. The interview went well and he got the JOB(91), PROMOTION(3), "A"(1), ANSWER(1), INFORMATION(1), STORY(1), HIRE(1), POSITION(1)
122. He ate a plate of franks and BEANS(91), FRIES(3), HAMBURGERS(2), CHILI(1), CHIPS(1), EGGS(1), ROLLS(1), *SAUERKRAUT(1)
123. The campers pitched a TENT(91), FIRE(2), GAME(1), FISH(1), PARTY(1), WATER(1)
124. She was glad they offered her maternity LEAVE(91), CLOTHES(5), BENEFITS(1), INSURANCE(1), PAYMENTS(1), ROBES(1), SUITS(1)
125. The butcher delivered an entire side of BEEF(90), MEAT(3), COW(1), LAMB(1), PIG(1), RIBS(1), STEED(1), *BACON(1)
126. The skater fell on the ICE(90), FLOOR(3), GROUND(3), SIDEWALK(1), RINK(1), POND(1)
127. The earth spins on its AXIS(90), AXLE(5), ORBIT(2), RINGS(1), SIDE(1)
128. She had a run in her STOCKINGS(90), PANTYHOSE(2), SHORTS(1), MONEY(1), NOSE(1),NYLON(1) SHOES(1), STOMACH(1), TIGHTS(1), YARD(1), * GARDEN(1),
129. The couple applied for a marriage LICENSE(90), CERTIFICATE(8), COUNSELOR(2)
130. The lost dog was returned to his OWNER(90), HOME(3), MASTER(3)
131. We wrapped the leftovers in tin FOIL(90), CANS(5), PAPER(3), CONTAINERS(2), PANS(1)
132. Please wipe your feet on the MAT(89), DOORMAT(4), RUG(4), NEWSPAPER(1), TOWEL(1), FLOOR(1)
133. She had a cold and a sore THROAT(89), NOSE(7), LIP(1), MUSCLE(1)
134. Jane returned the books to the LIBRARY(89), STORE(5), SHELF(2), OWNER(1), TEACHER(1), BOOKCASE(1)

135. She guessed the answer to the QUESTION(88), PROBLEM(4), PUZZLE(2), QUIZ(1), RIDDLE(1), TEST(1), CLASS(1), *GAME(1)
136. The records were kept in a filing CABINET(88), DRAWER(3), BOX(1), BIN(1), CASE(1), FASHION(1), FOLDER(1), SYSTEM(1), ROOM(1)
137. The man closed the curtain of the voting BOOTH(88), MACHINE(3), BOX(1), BLINDS(1), SCREEN(1)
138. The performer bowed to the AUDIENCE(88), CROWD(7), APPLAUSE(3), GROUND(1), PERFORMANCE(1)
139. I went to the gym to get some EXERCISE(88), EXPERIENCE(2), MUSCLES (2), WEIGHTS(2), PRACTICE(1), RELIEF(1), SHORTS(1), SWEAT(1), TRAINING(1), WATER (1), CLOTHES(1), ICE(1), *EQUIPMENT(1), *SNEAKERS(1)
140. He put the cigarette out in the ASHTRAY(87), TRAY(6), SAND(2), WATER(1), GRASS(1), FLOOR(1), TOILET(1), *STREET(1), *LOBBY(1), *CUP(1)
141. The natives danced around the FIRE(87), CIRCLE(2), CLOCK(1), GIRL(1), FLAME(1), FLOOR(1), POT(1), SACRIFICE(1), TOTEM POLE(1), VISITOR(1), CAMP(1), POLE(1), *MAYPOLE(1), *CAPTIVE(1), *BONFIRE(1), *TREE(1)
142. They started down the ski SLOPE(87), HILL(4), JUMP(3), RACING(1), LIFT(1), RAMP(1), RUN(1), TRAIL(1)
143. The yoyo had a knot in the STRING(87), CORD(6), END(4), MIDDLE(2), BOTTOM(1), TWINE(1), *ELASTIC(1), ROPE(1)
144. The ice cream was topped with a red CHERRY(86), SPRINKLES(4), SYRUP(4), STRAWBERRIES(2), BERRIES(1), FROSTING(1), JIMMIES(1), *RASPBERRIES(1), SAUCE(1)
145. Raise the flag up the POLE(86), FLAGPOLE(6), MAST(6), HILL(1), POST(1), SALUTE(1), ROOF(1), *STAFF(1)
146. The witch flew off on her BROOM(86), BROOMSTICK(12), STICK(1)
147. The typist indented at the beginning of a PARAGRAPH(86), SENTENCE(6), LETTER(2), DICTATION(1), MARGIN(1), *LINE(1)
148. Doonesbury is his favorite comic STRIP (86), BOOK(5), CHARACTER(4), STORY(2), ACTOR(1), *CARTOON(1)
149. All her music was on cassette TAPE(85), RECORDER(6), PLAYER(4), DECK(1), DISK(1), DISK PLAYERS(1), *RECORDS(1)
150. Sue put on her shoes and SOCKS(85), COAT(3), DRESS(1), HAT(1), GLOVES(1), LEFT(1), STOCKINGS(1), CLOTHES(1), *HOSE(1)
151. She was leaving on a two week VACATION(85), CRUISE(8), TOUR(2), TRIP(1), HOLIDAY(1), NOTICE(1)
152. She liked to paint with water COLORS(85), PAINTS(13)
153. He cleared his throat to catch their ATTENTION(85), BREATH(5), STORY(2), TRAIN(2), BOUQUET(1), VOICE(1), WORDS(1), QUESTION(1), *EYE(1), *REACTION(1)

154. We heard the morning news on the RADIO(85), TV(10), TELEVISION(4)
155. The wolves howled at the MOON(84), SHEEP(3), NIGHT(2), CAMPERS(1), DOG(1), HUNTERS(1), INTRUDERS(1), NOISE(1), PEOPLE (1), SOUND(1), TIGERS(1), DARK(1), *WIND(1), *MAN(1), *DRAGON(1)
156. For dessert he had apple PIE(84), STRUDEL(5), TART(2), TURNOVER(1), CAKE(1), CRISP(1), CRUMB(1), CUSTARD(1), JUICE(1), SAUCE(1), *COBBLER(1), *DUMPLING(1)
157. He awoke from a sound SLEEP(84), ALARM(4), OUTSIDE(2), SUDDENLY(1), AFAR(1), CRASHING(1), BLAST(1), LAST NIGHT(1), OF THUNDER(1), THUMPING(1), NAP(1)

References

Albert, M.S., Duffy, F.H. and McAnulty, G.B. (1990). Electrophysiologic comparisons between two groups of patients with Alzheimer's Disease. Archives of Neurology, 47, 857-863.

Albert, M.S., Heller, H.S. and Milberg, W. (1988). Changes in naming ability with age. Psychology abnd Aging, 3, 173-178.

Bayles, K.A and Tomeda, C.K. (1983). Confrontation naming impairment in dementia. Brain and Language, 19, 98-114.

Benson, F. (1979). Aphasia, Alexia and Agraphia. Churchill Livingstone, New York.

Bentin, S. (1987). Event-related potentials, semantic processes, and expectancy factors in word recognition. Brain and Language, 31, 308-327.

Bentin, S., McCarthy, G. and Wood, C.C. (1885). Event-related potentials, lexical decision and semantic priming. Electroencephalography and Clinical Neurophysiology, 60: 343-355.

Benton, A.L. (1955). Visual Retention Test. New York: The Psychological Corporation.

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

Bloom, P.A. and Fischler, I. (1980). Completion norms for 329 sentence contexts. Memory and Cognition, 8, 631-642.

Bowles, N.L. (1989). Age and semantic inhibition in word retrieval. Journal of Gerontology, 44, 88-90.

Bowles, N.L. and Poon, L.W. (1985). Age and retrieval of words in semantic memory. Journal of Gerontology, 40 , 71-77.

Brown, A.S. (1979). Priming effects in semantic memory retrieval processes. Journal of Experimental Psychology, Learning, Memory and Cognition, 5, 65-77.

Brown, R. and McNeill, D. (1966). The "tip of the tongue" phenomenon. Journal of Verbal Learning and Verbal Behavior, 5, 325-337.

Burke,D.M., White, H. and Diaz, D.L.(1987). Semantic priming in young and older adults: Evidence for age constancy in automatic and attentional processes. Journal of Experimental Psychology, Learning, Memory and Cognition, 13, 79-88.

Bushke, H. (1973). Selective reminding for analysis of memory and learning. Journal of Verbal Learning and Verbal Behavior, 12, 543-550.

Butters, N.D., Granholm, E., Salmon, D.P. Grant, I. and Wolfe, J. (1987). Episodic and semantic memory: A comparison of amnesic and demented patients. Journal of Clinical and Experimental Neuropsychology, 9, 479-497.

Caplan, D. (1987). Neurolinguistics and Linguistic Aphasiology: An Introduction. Cambridge University Press, Cambridge.

Cerella, J. and Fozard, J.L. (1984). Lexical access and age. Developmental Psychology, 20, 235-243.

Chang, T.M. Semantic memory: Facts and Models. (1986). Psychological Bulletin, 99, 199-220.

Chertow, H. and Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. Brain, 113, 397-417.

Collins, A.M. and Loftus, E.F. (1975). A spreading activation theory of semantic processing. Psychological Review, 87, 407-428.

Collins, A.M. and Quillian, R. (1969). Retrieval time from semantic memory. Journal of Verbal Learning and Verbal Behavior, 8, 240-247.

Collins, A.M. and Quillian, R. (1970). Facilitating retrieval from semantic memory: The effect of repeating part of an inference. Acta Psychologica, 33, 304-314.

Corkin, S., (1982). Some relationships between global amnesias and the memory impairments in Alzheimer's Disease. In Davis, K.L., Growdon, J.H., Usdin, E. and Wurtman, R.J., (Eds.), Alzheimer's Disease: A report of progress in research, Raven Press, New York, New York.

De Groot, A.M. (1983). The range of automatic spreading activation in word priming. Journal of Verbal Learning and Verbal Behavior, 22, 417-436.

Deacon, D. Breton, F. Ritter and Vaughan. (in press). The relationship between N2 and N400: Scalp distribution, stimulus probability and task relevance. Psychophysiology

Dixon, W.D. (1987). The BMDP Biomedical Computer Programs. .Los Angeles: University of California Press.

Fischler, I. and Bloom, P.A. (1979). Automatic and attentional processes in the effects of sentence contexts on word recognition. Journal of Verbal Learning and Verbal Behavior, 18, 1-20.

Fischler, I. and Bloom, P.A. (1985). Effects of constraint and validity of sentence contexts on lexical decisions. Memory and Cognition, 13, 128-139.

Fischler, I., Bloom, P.A., Childers, D.G., Arroyo, A.A. and Perry, N.W. (1984). Brain potentials during sentence verification: Late negativity and long-term memory strength. Neuropsychologia, 22, 559-568.

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

Fischler, I., Jin, Y., Boaz, T.L., Perry, N.W. and Childers, D.G. (1987). Brain potentials related to seeing one's own name. Brain and Language, 30, 245-262.

Flicker, C. Ferris, S.H. and Crook, T. (1987). Implications of memory and language dysfunction in the naming deficit of senile dementia. Brain and Language, 31, 187-200.

Folstein M.F., Folstein S.E. and McHugh, P.R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. Journal of Psychiatric Research., 12, 198-198.

Ford, J., Pfefferbaum, A. and Kopell, B. (1982). Effects of perceptual and cognitive difficulty on P300 and RT in young and old adults. Electroencephalography and Clinical Neurophysiology, 6.

Friedman, D. (1990). ERPs During continuous recognition memory for words. Biological Psychology, 30, 61-87.

Friedman, D., Hamberger, M. (1989). ERPs during continuous recognition memory in young, middle-aged and elderly adults. Psychophysiology, 27, 4A, S27. (Abstract).

Friedman, D., Putnam, L., Ritter, W., Hamberger, M. and Berman, S. (in press). A developmental event-related potential study of picture matching in children, adolescents and young adults: A replication and extension. Psychophysiology.

Gainotti, G., Daniele, A, Nocentini, U. and Silveri, M.C. (1990). The nature of lexical-semantic impairment in Alzheimer's Disease. Journal of Neurolinguistics, 10, in press.

Glass and Holyoak (1974). The effect of some and all on reaction time for semantic decision. Memory and Cogntion, 2, 436-440.

Gratton, G., Coles, M.G.H. and Donchin. (1983). A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55, 468-484.

Grober E., Buschke, H., Kawas, C. and Fuld, P. 1985. Impaired Ranking of Semantic Attributes in Dementia. Brain and Language, 26, 276-286.

Grossberg, S. and Stone G. (1986). Neural Dynamics of word recognition and recall: Attentional priming, learning and resonance. Psychological Review, 93, 46-74.

Gurland, B., Golden, R.R., Teresi, J.A. and Challop, J. (1984). The SHORT CARE: An efficient instrument of the assessment of depression, dementia and disability. Journal of Gerontology, 39, 166-169.

Hamberger, M. and Friedman, D. (in revision). Electrophysiological correlates of repetition priming and stimulus classification in young, middle-aged and elderly adults.

Hasher, L. and Zacks, R.T. (1979). Automatic and effortful processes in memory. Journal of Experimental Psychology: General,108, 356-388.

Hollingshead, A.B. and Redlich. F.C. (1958). Social Class and Mental Illness, New York: John Wiley.

Howard, D.V., McAndrews, M.P. and Lasaga, M.I (1988). Semantic priming of lexical decisions in young and old adults. Journal of Gerontology, 36, 707-714.

Huff, J.F., Corkin, S. and Growdon, J.H. (1986). Semantic impairment and anomia in Alzheimer's Disease. Brain and Language, 28, 235-249.

James, W. (1890). The Principles of Psychology. New York, Holt.

Kintch, W. (1974). The Representation of Meaning in Memory. Hillsdale, New Jersey: Erlbaum.

Kintch, W. (1980). Semantic memory: a tutorial. In R.S. Nickerson, (Ed.), Attention and performance VIII (pp. 595-620) Hillsdale, New Jersey: Erlbaum.

Kintch, W. and Mross, E.F (1985). Context effects in word recognition. Journal of Memory and Language, 24, 336-349.

Kleiman, G.M. (1980). Sentence frame contexts and lexical decisions: sentence acceptability and word relatedness effects. Memory and Cognition, 8, 336-344.

Kopolman, M. (1986). Recall of anomalous sentences in dementia and amnesia. Brain and Language 29, 154-170.

Kutas, M. and Van Petten, C. (1988). Event-related brain potential studies of language. Advances in Psychophysiology, 3,139-178.

Kutas, M. and Hillyard, S.A. (1988). An electrophysiological probe of incidental semantic association. Journal of Cognitive Neuroscience, 1, 38-49.

Kutas, M. and Hillyard, S.A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. Science, 207, 203-205.

Kutas, M. and Hillyard, S.A. (1983). Event-related brain potentials to grammatical errors and semantic anomalies. Memory and Cognition, 11, 539-550.

Kutas, M. and Hillyard, S.A. (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307, 579-590.

Kutas, M. and Hillyard, S.A. (1984). Word expectancy and event-related brain potentials during sentence processing. In Preparatory States and Processes, S. Kornblum and J. Requin (Eds.) Lawrence Erlbaum, New Jersey.

Kutas, M. and Hillyard, S.A. (1983). Event-related brain botentials to grammatical errors and semantic anomalies. Memory and Cognition,11(5), 539-550.

Lakeoff, G. (1987). Women, Fire and Dangerous Things: What categories reveal about the mind. University of Chicago Press, Chicago.

Loftus, E. Activation of semantic memory. (1975). American Journal of Psychology, 86, 331-337.

Lupker, S. (1984). Semantic priming without association: A second look. Journal of Verbal Learning and Verbal Behavior, 23, 709-733.

Luria, A.R. (1973). The Working Brain: An Introduction to Neuropsychology. Basic Books, Inc. New York.

Madden, D.J. (1986). Adult Age differences in visual word recognition: Semantic encoding and episodic retention. Experimental Aging Research, 12, 71-78.

Martin, A. (1987). Representation of semantic and spatial knowledge in Alzheimer's patients: Implications for models of preserved learning in amnesia. Journal of Clinical And Experimental Neuropsychology, 9, 191-224.

Martin, A. and Fedio, P. (1983). Word production and comprehension in Alzheimer's Disease: The breakdown of semantic knowledge. Brain and Language, 19, 124-141.

Martin, A., Brouwers, P., Lalonde, F., Cox, C., Taleska, P., Fedio, P., Foster, N.L., and Chase, T.N. (1986). Towards a behavioral typology of Alzheimer's patients. Journal of Clinical And Experimental Neuropsychology, 8, 594-610.

Martin, A., Brouwers, P., Cox, C. and Fedio, P. (1985). On the nature of the verbal memory deficit in Alzheimer's Disease. Brain and Language 25,323-341.

McCallum, W.C., Farmer, S.F. and Pockock. (1984). The effects of physical and semantic incongruities on auditory event related potentials. Electroencephalography and Clinical Neurophysiology, 59: 477-488.

McCarthy, G. and Donchin, E. (1983). Chronometric analysis of human information processing, In A.W.K. Gaillard and W. Ritter (Eds.), Tutorials in ERP Research: Endogenous Components, (pp. 251-268). Amsterdam: North Holland.

McKhann, G., Drachman, K., Folstein, M., Katzman, R., Price, D. and Stadlan, E.M. (1984). Clinical diagnosis of Alzheimer's Disease: a report of the NINCDS-ADRDA Work Group under the auspices of the Department of Health and Human Services Task Force on Alzheimer's Disease. Neurology, 34, 939-944.

McKlosky, M. and Glucklberg, S.(1979). Decision processes in verifying category membership statements: Implications for models of semantic memory. Cognitive Psychology, 11, 1-37.

Meyer, D., Schvaneveldt, R. and Ruddy, M. (1975). Loci of contextual effects on visual word-recognition. In Rabitt and Dornic (Eds.) Attention and performance V. (pp. 98-116). New York: Academic Press.

Milberg, W. and Blumstein, S.E. (1981). Lexical decision and aphasia: Evidence for semantic processing. Brain and Language, 14, 371-385.

Mitchell, D.B. (1989). How many memory systems? Evidence from aging. Journal of Experimental Psychology: Learning Memory and Cognition, 15, 31-49.

Nebes, R. (1989). Semantic memory in Alzheimer's Disease. Psychological Bulletin, 106, 377-394.

Nebes, R. and Brady, C.B.1990. Preserved organization of semantic attributes in Alzheimer's Disease. Psychology and Aging, 4, 574-579.

Nebes,R., Martin, D., Horn, L. C. (1984). Sparing of semantic memory in Alzheimer's Disease. Journal of Abnormal Psychology,321-330.

Neely, J. (1976). Semantic priming and retrieval from lexical memory: Evidence for facilitory and inhibitory processes. Memory and Cognition, 4, 648-654.

Nelson, K. (1985). Making Sense; The acquisition of shared meanings. Academic Press, New York.

Neville H.J., Kutas M., Chesney G. and Schmidt A.L. (1986). Event-related potentials during initial encoding and recognition memory of congruous and incongruous words. Journal of Memory and Language, 25, 75-92.

Nicholas, M., Obler, L.K., Albert, M.L. and Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's Disease and fluent aphasia. Journal of Speech and Hearing Research, 28, 405-410.

Ober, B.A. and Schenaut, G.K. (1988). Lexical decision and priming in Alzheimer's Disease. Neuropsychologia, 26, 273-286.

Ober, B.A., Dronkers, N.F., Koss, E., Delis, D.C. and Friedland, R.P. (1986). Retrieval from semantic memory in Alzheimer-type Dementia. Journal of Clinical And Experimental Neuropsychology,8 75-92.

Polich, J. (1985).Semantic categorization and event-related potentials. Brain and Language, 26, 304-321.

Riddoch, M.J., Humphreys, G.W., Coltheart, M. Funnell, E. (1988). Semantic systems or system? Neuropsychological evidence re-examined. Cognitive Neuropsychology, 5, 3-27.

Ritter, W., Simson, R. and Vaughan, H. (1988). Effects of the amount of stimulus information processed on negative event-related potentials. Electroencephalography and Clinical Neurophysiology, 69, 2344-258.

Ritter, W., Simson, R., Vaughan, H. and Friedman D.(1979). A brain event to the making of a sensory descrimination. Science, 203, 1358-1361.

Ritter, W., Simson, R. and Vaughan, H. (1972). Association cortex potentials and reaction time in auditory discrimination. Electroencephalography and Clinical Neurophysiology, 33, 547-555.

Roediger, H.L., Neely, J.H. and Blaxton, T.A. (1983). Inhibition from related primes in semantic memory retrieval: A reappraisal of Brown's (1979) Paradigm. Journal of Experimental Psychology, Learning, Memory and Cognition, 9, 478-485.

Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T.E. Moore (Ed.) Cogntive development and acquisition of language. New York: Academic Press.

Rosch, E. and Mervis, C. (1981). Categorization of natural objects. Annual Review of Psychology, 32, 89-115.

Rosen, W.G. (1980). Verbal fluency in aging and dementia. Journal of Clinical Neuropsychology,2, 135-146.

Rosen, W.G. and Mohs, R.C.(1982). Evolution of cognitive decline in dementia. In Davis, K.L., Growdon, J.H., Usdin, E. and Wurtman, R.J., (Eds.) Alzheimer's Disease: A report of progress in research, (pp.183-188) Raven Press, New York, New York.

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

Rugg, M.D. and Nagy, M.E. (1989). Event related potentials and recognition memory for words. Electroencephalography and Clinical Neurophysiology, 72:, 395-406.

Rugg, M.D. (1985). The effect of semantic priming and word repetition on event-related potentials. Psychophysiology, 22, 642-647.

Rugg, M.D. (1987). Dissociation of semantic priming and word and non-word repetition effects by event-related potentials. The Quarterly Journal of Experimental Psychology, 39A, 123-148.

Rugg, M.D., Furda, J. and Lorist, M. (1988). The effects of task on the modulation of event-related potentials by word repetition. Psychophysiology, 25, 55-63.

Salmon, D.P., Shimamura, A.P., Butters, N. and Smith. (1988). Lexical and semantic deficits in patients with Alzheimer's Disease. Journal of Clinical and Experimental Neuropsychology, 10, 477-494.

Salthouse, T.A. (1985). Speed of behavior and its implications for cognition. In J.E.Birren and K.W. Schaie (Eds.) Handbook of the psychology of aging (2nd ed. pp.400-426). New York: Van Nostrand Reinhold.

Schwantes, F.M. (1985). Expectancy, integration and interactional processes: Age differences in the nature of words affected by sentence context. Experimental child psychology, 39, 212-229.

Shoben, E.J. (1982). Semantic and lexical decisions. In Handbook of Research Methods in Human Memory and Cognition. Academic Press, Inc., New York, N.Y. Chapter 10, pp. 287-314.

Smith, E.E., Shoben and Ripps (1974). Structure and process in semantic memory: A featural model for semantic decisions. Psychological Review, 81, 214-241.

Spinnler, H., Della Sala, L., Bandera, R. and Baddely, A. (1988). Dementia, aging and the structure of human memory. Cognitive Neuropsychology, 5, 193-211.

Stuss, D.T., Picton, T.W. and Cerri, A.M. (1988). Electrophysiological manifestations of typicality judgement. Brain and Language, 33, 260-270.

Stuss, D.T., Sarazin, E.E. and Picton, T.W. (1983). Event-related potentials during naming and mental rotation. Electroencephalography and Clinical Neurophysiology, 56, 133-146.

Sutton, S. and John, E.R. (1965). Information delivery and the densory evoked potential. Science, 155, 1436-1439.

Terry, R.D. (1978). Aging, senile dementia and Alzheimer's Disease. In R. Katzman, R.K. Terry and K.L. Bick (Eds), Alzheimer's Disease: Senile dementia and related disorders , 11-15, New York: Raven press.

Terry, R.D. and Davies, P, (1980). Dementia of the Alzheimer type. Annual Review of Neuroscience, 3, 77-95.

Troster, A.I., Salmon,D.P. McCullock, D. and Butter, N.D. (1989). A comparison of category fluency deficits associated with Alzheimer's and Huntington's Disease. Brain and Language, 37, 500-513.

Tulving, E. (1985). How many memory systems are there? American Psychologist, 40, 385-398.

Van Petten, C. and Kutas, M. (1989). Interactions between sentence context and word frequency in event-related potentials. Memory and Cognition.

Verleger, R., Gasser, T. and Mocks, J. (1982). Correction of EOG artifacts in event-related potentials of the EEG: Aspects of reliability and validity. Psychophysiology, 19, 472-480.

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

Winer, B.J. (1971). Statistical principles in experimental design. New York: McGraw Hill.